Supplementary Data

Monitoring Biocatalytic Transformations Mediated by Polyketide Synthase Enzymes in Cell Lysate via Fluorine NMR

Shawn K. Piasecki¹ & Adrian T. Keatinge-Clay¹,²
¹Institute for Cellular and Molecular Biology
²Department of Chemistry and Biochemistry
The University of Texas at Austin, Austin, TX 78712

Contents

I. Synthesis of Fluorinated Substrates 1, 3, and 5
II. Characterization by ¹H NMR, ¹³C NMR, ¹⁹F NMR, and LC-MS
III. Protein Expression
IV. Kinetic Assays Analyzed by Reversed-Phase HPLC
V. Kinetic Assays Analyzed by ¹⁹F NMR
VI. Stereocontrolled Reduction of Substrate 5

I. Synthesis of Fluorinated Substrates 1, 3, and 5

*N-Acetylcysteamine*

*N-Acetylcysteamine* was prepared by combining cysteamine hydrochloride (2 g, 17.6 mmol, 1 eq.), sodium bicarbonate (4.4 g, 52.8 mmol, 3 eq.), and potassium hydroxide (988 mg, 17.6 mmol, 1 eq.) in that order to 100 mL diH₂O at 22 °C. Acetic anhydride (832 μL, 8.8 mmol, 0.5 eq.) was added dropwise over 5 min., and the reaction was stirred for 10 min. Then the reaction was acidified to pH 7. The product was extracted with EtOAc (3 x 200 mL), dried over MgSO₄, filtered, and dried *in vacuo* to yield a clear oil.
3,3,3-Trifluoropropionyl-N-acetylcysteamine thioester (1)

N-Acetylcysteamine (522 mg, 4.4 mmol, 1 eq.) was dissolved in 10 mL dry DCM at 0 °C under N₂. 3,3,3-trifluoropropionyl chloride (452 μL, 4.4 mmol, 1 eq.) and triethylamine (731 μL, 5.3 mmol, 1.2 eq.) were added, and the reaction was left to warm to room temperature while stirring overnight. The reaction was washed with brine (3 x 20 mL) and the brine was back-extracted with DCM (3 x 20 mL). The organic layers were combined and dried in vacuo to yield a brown oil (800 mg, 79.5%). TLC analysis of the fractions was carried out in 100% EtOAc (R_f = 0.40).

3-Oxo-5,5,5-trifluoropentanoyl-N-acetylcysteamine thioester (3)

Meldrum’s acid (1.0 g, 6.9 mmol, 1 eq.) was dissolved in 10 mL dry DCM under N₂. Pyridine (1.2 mL, 14.2 mmol, 2.05 eq.) was added, and the reaction was cooled to 0 °C. 3,3,3-trifluoropropionyl chloride (860 μL, 8.3 mmol, 1.2 eq.) was dissolved in 2 mL dry DCM and added dropwise to the Meldrum’s acid solution over 1 h. The reaction was stirred at 0 °C for an additional hour and warmed to room temperature overnight. The reaction was diluted with 10 mL DCM, washed with 10% aqueous HCl (3 x 50 mL), dried with MgSO₄, filtered, and evaporated to yield a dark orange oil (1.2 g, 72.5%).

The Meldrum’s acid derivative (1.6 g, 6.3 mmol, 1 eq.) was dissolved in 100 mL dry toluene under N₂. N-Acetylcysteamine (750 mg, 6.3 mmol, 1 eq.) was added, and the reaction was refluxed overnight. The reaction was evaporated to yield a dark orange oil. The product was purified via flash chromatography with a ~3 cm Cu²⁺-impregnated silica gel layer on top with a 50-100% EtOAc:hexanes gradient to yield an orange powder after evaporation (572 mg, 33%). TLC analysis of the fractions was carried out in 100% EtOAc (R_f = 0.35).

(2RS)-Methyl-3-oxo-5,5,5-trifluoropentanoyl-N-acetylcysteamine thioester (5)

Thioester 3 (200 mg, 0.74 mmol, 1 eq.) was dissolved in 40 mL dry THF and cooled to 0 °C. Potassium tert-butoxide (99.4 mg, 0.89 mmol, 1.2 eq.) was added, and the flask was placed under N₂. Iodomethane (230.3 μL, 3.7 mmol, 5 eq.) was added dropwise over 5 min., and the reaction was warmed to room temperature while stirring overnight. The pH of the reaction was neutralized, diH₂O was added to create an aqueous layer, and the product was extracted with EtOAc (3 x 200 mL). The organic layer was dried with MgSO₄, filtered, and evaporated to yield a dark orange oil. The product was purified over a silica gel column using a 50-100%
EtOAc:hexanes gradient to yield a dark orange oil (135 mg, 64%). TLC analysis of the fractions was performed in 100% EtOAc ($R_f = 0.37$).

**II. Characterization by $^1$H NMR, $^{13}$C NMR, $^{19}$F NMR, and LC-MS**

NMR characterization was performed on a Varian Mercury 400 MHz instrument. LC-MS was performed on an Agilent Technologies 1200 HPLC with a Gemini C$_{18}$ column (5 μm, 2 x 50 mm; Phenomenex). This was coupled to an Agilent Technologies 6130 quadrupole mass spectrometer system containing an ESI source run in positive mode. The aqueous mobile phase was (A) diH$_2$O with 0.1% formic acid, and the organic mobile phase was (B) acetonitrile with 0.1% formic acid. A 5-95% B gradient was run over 12 min. with a flow rate of 0.7 mL/min.

**3,3,3-Trifluoropropanoyl-N-acetylcysteamine thioester (1)**

$^1$H NMR (400 MHz, CDCl$_3$) δ 1.97 (s, 3H, CH$_3$-C=O), 3.10-3.15 (m, 2H, S-CH$_2$), 3.39 (q, 2H, J = 9.97 Hz, CF$_3$-CH$_2$), 3.45-3.50 (m, 2H, N-CH$_2$), 5.84 (br s, 1H, NH). $^{13}$C NMR (400 MHz, CDCl$_3$) δ 22.10 (s, C=CH$_3$-C=O), 29.10 (s, S-CH$_2$), 38.95 (s, N-CH$_2$), 45.49-46.00 (m, F$_3$C-C=CH$_2$), 122.97 (q, J = 277.53 Hz, CF$_3$), 170.85 (s, CH$_3$=C=O), 189.15 (s, S=C=O). $^{19}$F NMR (400 MHz, CDCl$_3$) δ -62.79 (t, 3F, J = 9.95 Hz, CF$_3$). See also Supplemental Figures 1a and 2a. ESI-MS expected mass: 230.2; observed mass: 230.2.

**3-Oxo-5,5,5-trifluoropentanoyl-N-acetylcysteamine thioester (3)**

$^1$H NMR (400 MHz, CDCl$_3$) δ 1.98 (s, 3H, CH$_3$-C=O), 2.99 (q, 2H, J = 6.77 Hz, CF$_3$-CH$_2$, keto), 3.10-3.15 (m, 2H, S-CH$_2$), 3.40-3.45 (m, 2H, CF$_3$-CH$_2$, enol), 3.45-3.50 (m, 2H, N-CH$_2$), 3.80 (s, 2H, O=C=CH$_2$-C=O), 5.62 (s, 1H, HO-C=CH-C=O, enol), 5.84 (br s, 1H, NH). $^{13}$C NMR (400 MHz, CDCl$_3$) δ 22.12 (s, C=CH$_3$-C=O), 27.10 (s, S-CH$_2$), 38.42 (s, N-CH$_2$), 38.50-38.55 (m, F$_3$C-C=CH$_2$, keto), 45.35-45.40 (m, F$_3$C-C=CH$_2$, enol), 56.19 (s, O=C=CH$_2$-C=O, keto), 101.53 (s, HO-C=CH-C=O, enol), 122.76 (q, J = 277.53 Hz, C=CF$_3$, enol), 122.86 (q, J = 277.53 Hz, C=CF$_3$, keto), 169.68 (s, N=C=O), 190.47 (s, CH$_2$-COH=CH, enol), 191.14 (s, S=C=O), 193.96 (s, CH$_2$-CO-CH$_2$, keto). $^{19}$F NMR (400 MHz, CDCl$_3$) δ -63.57 (t, J = 10.03 Hz, CF$_3$, enol), -62.45 (t, 3F, J =
$10.12 \text{ Hz, CF}_3$, keto). See also Supplemental Figures 1b and 2b. ESI-MS expected mass: 272.3; observed mass: 272.0.

(3R)-Hydroxy-5,5,5-trifluoropentanoyl-N-acetylcysteamine thioester (4)
$^1$H NMR (400 MHz, CDCl$_3$) $\delta$ 1.98 (s, 3H, CH$_3$-C=O), 2.34 (m, 2H, CF$_3$-CH$_2$), 2.81 (m, 2H, J = 4.06 Hz, HO-C-CH$_2$-C=O), 3.02-3.08 (m, 2H, S-CH$_2$), 3.43-3.50 (m, 2H, N-CH$_2$), 4.45 (br s, 1H, OH-CH$_2$), 5.83 (br s, 1H, NH). $^{19}$F-NMR (400 MHz, DMSO-$d_6$) $\delta$ -63.6 (t, J = 10.71 Hz, CF$_3$).

ESI-MS expected mass: 274.3; observed mass: 274.2.

(2RS)-Methyl-3-oxo-5,5,5-trifluoropentanoyl-N-acetylcysteamine thioester (5)
$^1$H NMR (400 MHz, CDCl$_3$) $\delta$ 1.44 (d, 3H, J = 7.06 Hz, O=C-C-CH$_3$, keto), 1.97 (s, 3H, O=C-CH$_3$), 3.11 (q, 2H, J = 6.77 Hz, CF$_3$-CH$_2$), 3.10-3.15 (m, 2H, S-CH$_2$), 3.40-3.45 (m, 2H, N-CH$_2$), 3.47 (s, 3H, HO-C=CH$_3$, enol), 3.85 (q, 1H, J = 7.2, O=C-CH, keto), 6.20 (br s, 1H, NH). $^{13}$C NMR (400 MHz, CDCl$_3$) $\delta$ 12.29 (s, O=C-CH-C$_3$), 22.06 (s, CH$_3$-C=O), 28.13 (s, S-CH$_2$), 36.55-36.60 (m, F$_3$C-CH$_2$, enol), 37.95 (s, N-CH$_2$), 43.60-43.65 (m, F$_3$C-CH$_2$, keto), 60.41 (s, O=C-CH=CH=O, keto), 107.11 (s, HO-C=CH-C=O, enol), 117.90 (q, J = 277.53 Hz, CF$_3$, enol), 127.41 (q, J = 277.53 Hz, CF$_3$, keto), 169.77 (s, HN-C=O), 193.76 (s, HO-C=O, enol), 195.37 (s, S-C=O), 197.46 (s, CH$_2$-CO-C, keto). $^{19}$F NMR (400 MHz, CDCl$_3$) $\delta$ -62.91 (t, 3F, J = 10.12 Hz, CF$_3$, enol), -62.54 (t, 3F, J = 10.03 Hz, CF$_3$, keto). See also Supplemental Figures 1c and 2c.

ESI-MS expected mass: 286.3; observed mass: 286.0.
Supplemental Figure 1. $^1$H NMR spectra of (A) substrate 1, (B) substrate 3 and (C) substrate 5.
Supplemental Figure 2. $^{13}$C NMR spectra of (A) substrate 1, (B) substrate 3 and (C) substrate 5.
III. Protein Expression

The construction of the expression plasmids is detailed in Piasecki, S.K. et al. (Chem. Biol. 2011, 18, 1331). Expression plasmids were transformed into *Escherichia coli* BL21 (DE3) cells. Overnight cultures were inoculated into Luria broth containing 25 mg/L kanamycin and grown at 37 °C to an OD₆₀₀ of 0.6. Cells were then cooled to 15 °C, induced with 1 mM IPTG, and left to grow for an additional 16 h. Cells were spun down at 4,000 x g for 30 min. at 4 °C, and pellets were resuspended in lysis buffer (10% v/v glycerol, 300 mM NaCl, 50 mM Tris, pH 8.0). Cells were lysed via sonication and centrifuged at 30,000 x g for 30 min. at 4 °C. Cell lysate was dialyzed twice into dialysis buffer (10% v/v glycerol, 100 mM NaCl, 30 mM HEPES, pH 7.5) using 10 kDa MWCO cellophane dialysis tubing. Dilutions of each protein were either compared to purified enzyme of known concentration, or BSA as a standard to determine their concentrations within the lysate (Supplemental Figure 3a-c).

IV. Kinetic Assays Analyzed by Reversed-Phase HPLC

HPLC analysis was performed at 235 nm on a Waters 1525 binary pump attached to a 2998 PDA detector on a Microsorb C₁₈ column (300 μm, 250 x 4.6 mm, Varian). A gradient of 5-100% B was employed (solvent A: diH₂O containing 0.1% TFA; solvent B: MeOH containing 0.1% TFA).

*EryTE and Substrate 1*

Assays were set up using 115 μM EryTE cell lysate with either 3.75, 7.5, 15, 30, 60, or 120 mM of 1 at a final volume of 1.5 mL. The substrate was dissolved in 4% v/v DMSO-<sub>d₆</sub> prior to mixing into 10% v/v glycerol, 200 mM HEPES, pH 7.5 (final concentration). 100 μL timepoints were taken every 10 min. for 1 h. by quenching with 500 μL EtOAc. Each timepoint was extracted with EtOAc (3 x 500 μL) and dried in vacuo. Prior to injection on the HPLC each was dissolved in MeOH. The substrate 1 peak was integrated in each chromatograph and compared to
a standard curve in order to determine initial rates from the linear portion of each graph. The standard curve was made by injecting known concentrations of 1 into the HPLC, plotting the resulting peak area by substrate concentration, and fitting to a straight line. Rates were plotted against substrate concentration using GraFit, and curves were fit to the Michaelis equation to extract kinetic constants (Figure 1 in the text). Each assay was performed in duplicate.

_TylKR1 and Substrate 3_  
30 mM substrate 3 was set up in duplicate with 81 μM TylKR1 cell lysate at a final volume of 1 mL. The substrate was dissolved in 4% v/v DMSO-d6 prior to mixing into 300 mM D-glucose, 100 μM NADP⁺, 1 μM GDH, 10% v/v glycerol, 80 mM NaCl, 150 mM Tris, pH 7.5 (final concentration). 100 μL timepoints were taken every 10 min. for 1 h. by quenching with 500 μL EtOAc. Each timepoint was extracted with EtOAc (3 x 500 μL) and dried _in vacuo_. Prior to injection on the HPLC each was dissolved in MeOH. Peak areas were compared to a standard curve in order to extract initial velocities. The standard curve was made by injecting known concentrations of 3 into the HPLC, plotting the resulting peak area by substrate concentration, and fitting to a straight line.

V. Kinetic Assays Analyzed by ¹⁹F NMR

Kinetic assays were performed on a Varian Inova 500 MHz instrument in 10 mm tubes. Prior to adding the enzyme, the magnetic field was shimmed on the FID from DMSO-d₆ used to solubilize the substrate.

_EryTE and Substrate 1_  
The reactions are the same as reported in the previous HPLC section. The first timepoint was taken 15 min. after enzyme addition since it takes time to set up the instrument (this delay was kept constant in all assays). The peaks were integrated using VNMR software, and the areas for each timepoint were compared to a standard curve in order to determine concentration from peak area. The standard curve was prepared by adding a known amount (2 mg, 10.4 mM final) of 3,3,3-trifluoropropionic acid to solutions of substrate 1 (final concentrations of 3.75, 7.5, 15, 30,
and comparing the peak areas to determine actual concentrations of 1 in each array. Arrays were collected at 470 MHz at a probe temperature of 27 °C, sweepwidth = 9400 Hz, 64K data points, pulse angle = 81°, receiver gain = 40, 8 FIDs per timepoint, repetition time of 300 sec. for 21,600 sec. total. 3,3,3-trifluoropropionic acid was also used as a standard to verify product formation (Supplemental Figure 5).

**TylKR1 and Substrate 3**

The reactions are the same as reported in the previous HPLC section. A standard curve was performed in the same manner as with substrate 1 but with 7.5, 15, 30, 60, and 120 mM of 3. All acquisition and processing parameters were as reported above. The first timepoint was taken 15 min. after enzyme addition since it takes time to set up the instrument. ESI-MS was performed on the EtOAc-extracted reaction (3 x 500 μL) after data collection, and an observed ESI-MS mass of 274.2 matched the expected mass (274.3) for the reduced product. 19F and 1H-NMR were also used to verify product formation (Supplemental Figure 6).

**VI. Stereocontrolled Reduction of Substrate 5**

**Reduction assays with AmpKR2, RifKR7, TylKR1, and EryKR1**

KR reduction assays were performed similar to the reduction assays of TylKR1 on substrate 3. 20 mM of 5 was dissolved into 4% v/v DMSO-d6 prior to mixing into 300 mM D-glucose, 100 μM NADP+, 80 mM NaCl, 10% v/v glycerol, 150 mM Tris, pH 7.5, as well as 1 μM GDH and 5 μM of KR (final concentrations) in an 800 μL volume. Reactions were incubated at room temperature, and 19F NMR spectra were acquired on a Varian Mercury 400 MHz instrument after 24 h. by moving the reaction into a 10 mm NMR tube. Since the anti- diastereomers only shift slightly from the unreduced substrate 3, we include an additional 19F NMR spectrum (Supplemental Figure 4) showing the TylKR1 reaction progress after 1 h. in which the substrate and product triplets partially overlap. Once NMR analysis was complete for each of the KR-types, reactions were extracted with EtOAc (3 x 500 μL) and analyzed by ESI-MS. An observed mass of 288.2 verified that the reduced product (expected mass of 288.3) was being generated in each KR reaction.
Supplemental Figure 3. 12% SDS-PAGE gels helped to estimate the concentrations of overexpressed PKS enzymes in the cell lysate through either a comparison with purified enzyme, or a 1 mM BSA standard. Enzymes used in kinetic assays (EryTE and TylKR1) were compared to purified enzyme to more accurately determine their concentrations within lysate. Standards were made by mixing known concentrations of purified enzyme with *E. coli* BL21(DE3) cell lysate. (A) The intensity of the EryTE band from a 1:50 dilution of EryTE-containing lysate equates to a 1:25 dilution of purified EryTE + cell lysate. Since the concentration of undiluted, purified EryTE was 95 μM, the concentration of EryTE in this cell lysate prep was ~190 μM. (B) The concentration of GDH in this cell lysate prep was estimated at ~480 μM through a comparison with BSA standards (C) The concentration of TylKR1 in this cell lysate prep was estimated at ~325 μM through a comparison with TylKR1 standards (undiluted, purified TylKR1 was 163 μM). (D) The concentrations of AmpKR2, RifKR7, and EryKR1 in their cell lysate preps were estimated to be ~16 μM, ~19 μM, and ~200 μM, respectively through a comparison to BSA standards. “L” denotes a Fermentas Page Ruler ladder.
Supplemental Figure 4. A $^{19}$F NMR spectrum showing the transformation of substrate 5 to product 8 by TylKR1. The reaction progress was observed 1 h. after enzyme was added. The substrate and product peaks partially overlap.

Supplemental Figure 5. 3,3,3-trifluoropropionic acid was used as a product standard to confirm EryTE product formation. The standard was added to E. coli BL21(DE3) cell lysate. A shift of -63.71 ppm aligns with a peak observed in the EryTE hydrolysis assays (Figure 1 in manuscript).
Supplemental Figure 6. In addition to ESI-MS, $^{19}$F- and $^1$H-NMR were used to verify formation of 4 by TylKR1. (A) A $^{19}$F-NMR spectrum taken at the completion of the reaction shows a triplet at -63.6 ppm. (B) The product was purified and resuspended in CDCl$_3$ to obtain a $^1$H-NMR spectrum. After the product was again isolated and added to the reaction solution, the same NMR triplet was observed by $^{19}$F-NMR (A).
**Supplemental Figure 7.** No detectable reaction occurs to the substrates of this study when incubated in *E. coli* BL21(DE3) cell lysate lacking PKS biocatalysts (no peaks besides those shown here appear in the spectra).

**Supplemental Figure 8.** The next six pages report the raw data used to determine the kinetic parameters of the EryTE reaction on substrate 1 by $^{19}$F NMR and HPLC.