**Supplementary Information**

**NADPH-dependent Secondary Amine Organocatalysis hosted by a Nucleotide-binding Domain**

Thomas L Williams, Yu-Hsuan Tsai\*, Louis YP Luk\*,#

\*School of Chemistry, Cardiff University, Main Building, Park Place, Cardiff, CF10 3AT

UK

#Cardiff Catalysis Institute, School of Chemistry, Cardiff University, Main Building, Park Place,

Cardiff, CF10 3AT

Table of Contents

[1. Reagents and materials 2](#_Toc68622455)

[2. Molecular cloning 3](#_Toc68622456)

[3. Protein mass spectra 9](#_Toc68622457)

[4. Analytical size exclusion chromatography of LmrR variants (y-axis is A210 and x-axis is retention time in minutes) 11](#_Toc68622458)

[5. Catalytic activity screening reactions 14](#_Toc68622459)

[6. Enzyme kinetic characterisation 16](#_Toc68622460)

[7. Substrate scope 17](#_Toc68622461)

[Table S6. Substrate scope analysis of LmrR-Phe93-1. Conversions determined by 1H NMR. 17](#_Toc68622462)

[LmrR-Phe93-1 substrate scope*.* 17](#_Toc68622463)

[EcDHFR-Ala7-1 substrate scope*.* 18](#_Toc68622464)

[8. Iminium ion intermediate characterisation 19](#_Toc68622465)

[9. NADPH hydride transfer selectivity characterisation 22](#_Toc68622466)

[10. Kinetic isotope effect study 23](#_Toc68622467)

[11. Cofactor recycling assay 25](#_Toc68622468)

[12. Analytical chemistry 26](#_Toc68622469)

[13. Nucleotide and amino acid sequences 28](#_Toc68622470)

[14. References 32](#_Toc68622471)

# Reagents and materials

All α,β-unsaturated aldehyde and ketone substrates, as well as compound **5**, were purchased from commercial retailers. NADPH tetrasodium salt was purchased from Apollo Scientific (#BIB3014). NADP+ disodium salt was purchased from Santa Cruz Biotechnology (#205763). Glucose-6-phosphate dehydrogenase (G6PDH) from *L. mentroides* (#G8259) and glucose-6-phosphate (#10127647001) were purchased from Merck. Isoproponal-D8 was purchased from AcrosOrganics (#174850250).. Synthesis of amino acids **1**, **2** 1 and, **3** 2, BNAH3 and [4*R*-2H]-NADPH (NADPD)4 were performed as previously reported. Gibson Assembly reactions were performed using NEBuilder HiFi DNA assembly (#E2621) following the manufacturer’s protocol. PCR reactions were performed using PrimeSTAR Max from TaKaRaBio (#R045A). Oligonucleotide primers were purchased from Merck, and their sequences are shown in Table S1. Protein extinction coefficients are estimated using ProtParam (Expasy).

# Molecular cloning

*M. bakeri* PylRS/tRNA plasmids

The plasmid encoding MbPylRS and PyltRNA was a kind gift from Jason Chin5. The ThzKRS plasmid was constructed by introducing four mutations by PCR into the MbPylRS plasmid. Firstly introduction of mutation Asp267Ser using primers P1 and P2 and then introduction of mutations Cys313Val, Met315Phe and Asp344Gly using primers P3 and P42.

Expression plasmids

The sfGFP Asp150TAG plasmid is available from Addgene (#133455)6.

The gene encoding the LmrR protein was purchased as a double stranded DNA fragment from GeneArt (Life Technologies). The gene was codon optimised for expression in *E. coli* and contained two mutations, Lys55Asn and Lys59Asp, in comparison to the wild-type sequence. The gene was synthesised with an additional C-terminus hexa-histidine tag and 20 base pair overhangs complimentary to NcoI and BamHI of a linearised pET-28a vector. The gene was cloned into the linearised vector using the Gibson Assembly. Each TAG mutant was generated through PCR. For the Val15TAG mutant, primers used were P5 and P6, the Asp19TAG mutant primers P7 and P8, the Met89TAG mutant primers P9 and P10, and the Phe93TAG mutant primers P11 and P12.

The gene encoding the wild-type *E. coli* DHFR was a kind gift from Rudolf Allemann. The gene was amplified by PCR using primers P13 and P14 and simultaneously a C-terminal hexa-histidine tag and 20 base pair overhangs complimentary to NcoI and BamHI of a linearised pET28a vector were added. The gene was cloned into the linearised vector using the Gibson Assembly. Introduction of the TAG mutations was performed using PCR with primers P15 and P16 for the Ala7TAG mutant, P17 and P18 for the Phe31TAG mutant, and P19 and P20 for the Ser49TAG mutant.

The expression plasmid for the enzyme TbADH used in the synthesis of [4*R*-2H]-NADPH was a kind gift from Rudolf Allemann.

**Table S1**. Table of oligonucleotide primers



1. Protein expression and purification

To test the incorporation of the unnatural amino acids **1**, **2** and **3** the sfGFP Asp150TAG plasmid was co-transformed with either the MbPylRS or ThzKRS plasmid into *E. coli* BL21(DE3) cells and recovered for 1 hour at 37 °C in 1 mL of LB media with constant agitation and then 200 μL was used to inoculate a 10 mL starter culture of LB media supplemented with kanamycin (37.5 μg/mL) and spectinomycin (50 μg/mL). The starter culture was incubated overnight at 37 °C with constant agitation. The following day, 100 μL of the starter culture was used to inoculate 10 mL of LB media supplemented with kanamycin (37.5 μg/mL) and spectinomycin (50 μg/mL) and this culture was incubated at 37 °C until it reached an OD600 of 0.6 – 0.8. The culture was then split into two 5 mL aliquots and IPTG was added to both, to reach a final concentration 0.5 mM, at the same time, to one aliquot was added 1 mM of the required unnatural amino acid. The cultures were then incubated at 20 °C for overnight. The following day the cultures were normalised to an OD600 of 1.0 and unnatural amino acid incorporation was determined by SDS-PAGE (Fig. S1)

A picture containing text

Description automatically generated

**S1.** SDS-PAGE analysis for the expression of sfGFP Asp150TAG with or without 1 mM amino acids **1**, **2** or **3**. **1** was incorporated using plasmid MbPylRS, **2** and **3** using the plasmid ThzKRS. Following overnight expression samples were normalised to OD600 of 1.0. Protein ladder in kDa. Mass of full length sfGFP = approx. 27 kDa.

For LmrR, the plasmid was transformed into *E. coli* BL21(DE3) cells. The cells were recovered for 1 hour at 37 °C in 1 mL of LB media with constant shaking and then 200 µL was used to inoculate 10 mL of LB media supplemented with kanamycin (37.5 μg/mL). The starter culture was incubated overnight at 37 °C with constant shaking. The following day the starter culture was diluted into 1 L of fresh LB media supplemented with kanamycin (37.5 μg/mL). The culture was incubated at 37 °C until it reached an OD600 of 0.8-1.0 at which point gene expression was induced with 0.5 mM of IPTG. The cells were then cultured at 30 °C overnight. The following day the cells were harvested by centrifugation (20,000 rcf, 4 °C, 30 min). The pellet was resuspended in 25 mL PBS buffer (50 mM NaPi, 150 mM NaCl, pH 8.0) and added to the suspension was 1 μg/mL DNase, 100 μg/mL MgCl2 and 5 μg/mL PMSF. The resuspended cells were lysed by sonication on ice. The lysed cells were centrifuged to separate the inclusion bodies (27,000 rcf, 4 °C, 30 min) and the supernatant passed through 22 μm syringe filters. The filtered supernatant was applied to a Ni-NTA column equilibrated in the above buffer. The column was washed twice with 30 mL PBS buffer (50 mM NaPi, 150 mM NaCl, 20 mM imidazole, pH 8.0) and then the protein was eluted in elution buffer (50 mM NaPi, 150 mM NaCl, 300 mM imidazole, pH 8.0). The elution was dialysed overnight into the reaction buffer (50 mM NaPi, 150 mM NaCl, pH 7.0) at 4 °C. The following day the protein was concentrated with 10 kDa cut-off centrifugal concentrators and the concentration determined using a NanoDrop instrument by measuring the UV absorbance 280 nm using the extinction coefficient of 19940 M-1 cm-1 for one LmrR monomer. Purified proteins were stored in the reaction buffer at -80 °C.

For the LmrR variants containing an unnatural amino acid, *E. coli* BL21(DE3) cells were co-transformed with the relevant LmrR plasmid and either the MbPylRS plasmid or the ThzKRS plasmid. In addition to kanamycin, spectinomycin (50 μg/mL) was added to each starter and expression culture and upon induction with IPTG, the relevant unnatural amino acid was added to a final concentration of 1 mM. The variants were purified as the wild-type enzyme. Each protein was analysed by SDS-PAGE (Fig. S2) and unnatural amino acid incorporation determined by mass spectrometry (Fig. S4). Dimer formation was analysed by analytical size exclusion chromatography (SI, Pg S7-S8).

A picture containing text, electronics

Description automatically generated

**S2**. SDS-PAGE analysis of LmrR TAG variants isolated by Ni-NTA chromatography. Protein ladder in kDa. Mass of a LmrR monomer = approx. 14 kDa.

For DHFR, the plasmid was transformed into *E. coli* BL21(DE3) cells. The cells were recovered for 1 hour at 37 °C in 1 mL of LB media with constant shaking and then 200 µL was used to inoculate 10 mL of LB media supplemented with kanamycin (37.5 μg/mL). The starter culture was incubated overnight at 37 °C with constant shaking. The following day the starter culture was diluted into 1 L of fresh LB media supplemented with kanamycin (37.5 μg/mL). The culture was incubated at 37 °C until it reached an OD600 of 0.6-8.0 at which point gene expression was induced with 0.5 mM of IPTG. The cells were then cultured at 20 °C overnight. The following day the cells were harvested by centrifugation (5,000 rcf, 4 °C, 30 min). The dry pellet was resuspended in 25 mL PBS buffer (50 mM NaPi, 150 mM NaCl, pH 8.0) and added to the suspension was 5 μg/mL PMSF. The resuspended cells were lysed by sonication on ice. The lysed cells were centrifuged to separate the inclusion bodies (27,000 rcf, 4 °C, 30 min) and the supernatant passed through 22 μm syringe filters. The filtered supernatant was applied to a Ni-NTA column equilibrated in the above buffer. The column was washed twice with 30 mL PBS buffer (50 mM NaPi, 150 mM NaCl, 20 mM Imidazole, pH 8.0) and then the protein was eluted in elution buffer (50 mM NaPi, 150 mM NaCl, 300 mM Imidazole, pH 8.0). The elution was dialysed overnight into the reaction buffer (50 mM NaPi, 150 mM NaCl pH 7.0) at 4 °C. The following day the protein was concentrated with 10 kDa cut-off centrifugal concentrators and the concentration determined using a NanoDrop instrument by measuring the UV absorbance at 280 nm using the extinction coefficient of 33585 M-1 cm-1. Purified proteins were stored in the reaction buffer at -80 °C.

For the DHFR TAG mutants the transformation was performed by co-transforming the relevant DHFR plasmid with the MbPylRS plasmid. In addition to kanamycin, spectinomycin (50 μg/mL) was added to each starter and expression culture and upon induction with IPTG, amino acid **1** was added to a final concentration of 1 mM. The variants were purified as the wild-type enzyme. Each protein was analysed by SDS-PAGE (Fig. S3) and unnatural amino acid incorporation determined by mass spectrometry (Fig. S5).

Following the preliminary screening reactions **LmrR-Phe93-1** and **DHFR-Ala7-1** were further purified using fast protein liquid chromatography (FPLC). FPLC was performed on an ÄKTA purifier (GE Healthcare) system at room temperature using a GE Healthcare BSD75 10/300 SEC column. Protein elution was monitored by UV absorbance at 280 nm and the elution buffer was PBS (50 mM NaPi, 150 mM NaCl, pH 7.0). Fractions containing **DHFR-Ala7-1** were combined and concentrated using a 10 kDa cut off centrifugal concentrator (Millipore). Protein concentration was determined using nanodrop. Purified proteins were stored in the reaction buffer at -80 °C.

A picture containing text

Description automatically generated

**S3.** SDS-PAGE analysis of DHFR and variants isolated by Ni-NTA chromatography. Protein ladder in kDa. Mass of DHFR = approx. 19 kDa.

Expression and purification of TbADH was performed as previously reported and was confirmed by SDS-PAGE4.

# 3. Protein mass spectra

A picture containing graphical user interface

Description automatically generated

**S4.** Deconvoluted ESI mass spectra for each LmrR variant (without methionine).

A picture containing chart

Description automatically generated

**S5.** Deconvoluted ESI mass spectra for each DHFR variant (with methionine included).

# 4. Analytical size exclusion chromatography of LmrR variants *(y-axis is A210 and x-axis is retention time in minutes)*

Wt-LmrR

A screenshot of a social media post

Description automatically generated

Val15-**1**

A screenshot of a social media post

Description automatically generated

Val15-**2**

A screenshot of a social media post

Description automatically generated

Val15-**3**

A screenshot of a social media post

Description automatically generated

Asp19-**1**

A screenshot of a social media post

Description automatically generated

Asp19-**2**

A screenshot of a social media post

Description automatically generated

Asp19-**3**

A screenshot of a social media post

Description automatically generated

Met89-**1**

A screenshot of a cell phone

Description automatically generated

Met89-**2 (**found to be unstable and prone to precipitation**)**

A screenshot of a social media post

Description automatically generated

Met89 **3**

A screenshot of a social media post

Description automatically generated

Phe93-**1**

A screenshot of a social media post

Description automatically generated

Phe93-**2**

A screenshot of a social media post

Description automatically generated

Phe93-**3**

A screenshot of a social media post

Description automatically generated

# 5. Catalytic activity screening reactions

Catalysis screening reactions for the LmrR variants were performed with the enzymes (>95% purity by SDS PAGE analysis) in total reaction volumes of 100 µL. To 50 µL (100 µg, 6.8 nmol) of enzyme solution in the reaction buffer (50 mM NaPi, 150 mM NaCl, pH 7.0) in a microcentrifuge tube was added a 5 µL solution of cinnamaldehyde **4** (8.9 µg, 68 nmol, 1 eq) in methanol, followed by 5 µL solution of BNAH (28 µg, 136 nmol, 2 eq) in methanol. The reaction buffer was added to reach a final volume of 100 µL.

Catalysis screening reactions for DHFR were performed with the enzymes (>95% purity by SDS PAGE analysis) in total reaction volumes of 100 µL. The concentration of NADPH was determined by measuring the UV absorbance at 340 nm using the extinction coefficient of 6200 M-1 cm-1 7. To 90 µL (100 µg, 5.12 nmol) of enzyme solution in the reaction buffer (50 mM NaPi, 150 mM NaCl, pH 7.0) in a microcentrifuge tube was added a 5 µL solution of cinnamaldehyde **4** (8.9 µg, 52 nmol, 1 eq) in methanol, followed by 5 µL solution of NADPH (77 µg, 104 nmol, 2 eq) in the reaction buffer.

The reactions were placed in a thermomixer at 25 °C and 500 rpm for 18 hours. The reactions were halted by adding 200 µL of DCM and vortexed vigorously, then centrifuged to separate the layers (20,000 rcf, 3 min, rt) and 100 µL of the organic layer was removed and subjected to analysis by GC-MS. In addition to the enzyme variants, catalyst free controls and reactions with the wild-type protein were also performed. **1** was also used with the same catalyst loading as the enzymes. Each reaction was performed in triplicate. (Tables S2 and S3).

**Table S2**. Yield of product hydro-cinnamaldehyde **5** by LmrR and variants after 18 h determined by comparison of the peak area against a standard calibration curve by GC-MS.



|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Catalyst | Yield 1 (%) | Yield 2 (%) | Yield 3 (%) | Mean (%) | | |
| Val15-**1** | 31 | 14 | 17 | 21 | ± | 9 |
| Val15-**2** | 15 | 11 | 11 | 12 | ± | 3 |
| Val15-**3** | 7 | 8 | 8 | 7 | ± | 1 |
| Asp19-**1** | 18 | 17 | 15 | 17 | ± | 2 |
| Asp19-**2** | 4 | 3 | 3 | 4 | ± | 1 |
| Asp19-**3** | 33 | 33 | 28 | 31 | ± | 3 |
| Met89-**1** | 12 | 8 | 1 | 7 | ± | 6 |
| Met89-**2** | 5 | 4 | 5 | 5 | ± | 1 |
| Met89-**3** | 9 | 11 | 9 | 10 | ± | 1 |
| Phe93-**1** | 68 | 57 | 49 | 58 | ± | 9 |
| Phe93-**2** | 3 | 2 | 2 | 2 | ± | 0 |
| Phe93-**3** | 13 | 10 | 13 | 12 | ± | 2 |
| Wt | 2 | 2 | 2 | 2 | ± | 0 |
| UAA **1** | 20 | 21 | 18 | 20 | ± | 2 |
| (-) control | 0 | 0 | 0 | 0 | ± | 0 |

**Table S3**. Yield of product hydro-cinnamaldehyde **5** by DHFR and variants after 18 h determined by comparison of the peak area against a standard calibration curve by GC-MS.



|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Catalyst | Yield 1 (%) | Yield 2 (%) | Yield 3 (%) | Mean (%) | | |
| Ala7-**1** | 89 | 88 | 99 | 92 | ± | 6 |
| Phe31-**1** | 18 | 18 | 18 | 18 | ± | 0 |
| Ser49-**1** | 0 | 0 | 0 | 0 | ± | 0 |
| Wt | 0 | 0 | 0 | 0 | ± | 0 |
| UAA **1** | 0 | 0 | 0 | 0 | ± | 0 |
| (-) control | 0 | 0 | 0 | 0 | ± | 0 |

# 6. Enzyme kinetic characterisation

Enzyme kinetic assays were performed under saturating substrate concentrations. Analysis was performed using the software Prizm with the non-linear regression Michaelis-Menten output.

For **LmrR-Phe93-1**, to 50 µM of the enzyme in the reaction buffer (90 µL) was added 1 mM of cinnamaldehyde **4** dissolved in 5 µL of methanol. The reaction was then initiated by adding BNAH dissolved in 5 µL of methanol in increasing concentrations. The reactions were placed in a thermomixer at 25 °C and 500 rpm for 2 hours.

For **DHFR-Ala7-1**, to 20 µM of the enzyme in the reaction buffer (90 µL) was added 1 mM of cinnamaldehyde **4** dissolved in 5 µL of methanol. The reaction was then initiated by adding NADPH dissolved in 5 µL of the reaction buffer in increasing concentrations. The reactions were placed in a thermomixer at 25 °C and 500 rpm for 2 hours.

The reactions were halted by adding 200 µL of DCM and vortexed vigorously, then centrifuged to separate the layers (20,000 rcf, 3 min, rt) and 100 µL of the organic layer was removed and analysed by GC-FID. Reactions were performed in triplicate (Table S4 and S5).

**Table S4**. Enzyme kinetic characterisation of **LmrR-Phe93-1***.*

|  |  |  |  |
| --- | --- | --- | --- |
| [BNAH] (µM) | Mean Rate (x10-4 s-1) | | |
| 200 | 1.70 | ± | 0.18 |
| 400 | 2.61 | ± | 0.31 |
| 600 | 2.98 | ± | 0.94 |
| 800 | 4.58 | ± | 0.75 |
| 1000 | 5.54 | ± | 0.41 |
| 2000 | 7.30 | ± | 0.26 |
| 3000 | 9.97 | ± | 0.90 |
| 4000 | 10.06 | ± | 0.13 |
| 5000 | 9.22 | ± | 0.97 |

**Table S5**. Enzyme kinetic characterisation of **EcDHFR-Ala7-1***.*

|  |  |  |  |
| --- | --- | --- | --- |
| [NADPH] (µM) | Mean Rate (x10-4 s-1) | | |
| 40 | 0.84 | ± | 0.44 |
| 50 | 1.68 | ± | 0.20 |
| 60 | 1.88 | ± | 0.45 |
| 75 | 2.29 | ± | 0.24 |
| 100 | 3.20 | ± | 0.27 |
| 125 | 3.51 | ± | 0.61 |
| 150 | 3.74 | ± | 0.36 |
| 250 | 4.61 | ± | 0.30 |
| 300 | 4.41 | ± | 0.41 |
| 500 | 3.77 | ± | 0.09 |

The results were fitted to the Michaelis-Menten equation using Prism and the *V*max, *k*cat and *K*M determined from the curve.

# 7. Substrate scope

Conversion of the substates was determined by 1H NMR. This was performed on a Bruker Advance 600 MHz NMR system equipped with a He cooled cryoprobe using 128 scans.

For **LmrR-Phe93-1**,to a microcentrifuge tube was added 500 µL of a stock solution of the enzyme (1 mg, 68 nmol) in PBS (50 mM NaPi, 150 mM NaCl, pH 7.0). To this was added 400 µL of PBS buffer and then 50 µL of a stock solution of the α,β-unsaturated carbonyl substrate (680 nmol, 1 eq) in methanol followed by 50 µL of a stock solution of BNAH (725 µg, 3.4 mmol, 5 eq) in methanol. For **DHFR-Ala7-1**, to a microcentrifuge tube was added 500 µL of a stock solution of the enzyme (100 µg, 5.12 nmol) in PBS (50 mM NaPi, 150 mM NaCl, pH 7.0). To the enzyme was added 400 µL PBS buffer and then 50 µL of a stock solution of the α,β-unsaturated carbonyl substrate (512 nmol, 1 eq) in methanol followed by 50 µL of a stock solution of NADPH (725 µg, 3.4 mmol, 5 eq) in the reaction buffer.

The reactions were placed in a thermoshaker and shook at 500 rpm for 48 hours at 25 °C. Following the completion of the reactions, DCM was added (500 µL) and the sample vortexed vigorously. The phases were separated by centrifugation (20,000 rcf, 3 min, rt). The organic phase was removed and placed in a new microcentrifuge tube. This was repeated a second time and the two extractions combined. The solvent was removed under nitrogen and the remaining residue re-dissolved in CDCl3 (600 µL). Each sample was then subjected to 1H NMR and the conversion determined by integration of the aldehyde protons. Reactions were performed in triplicate (Tables S6 and S7).

## **Table S6**. Substrate scope analysis of **LmrR-Phe93-1.** Conversions determined by 1H NMR.

## **LmrR-Phe93-1** substrate scope*.*



|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Substrate | Replica 1 | Replica 2 | Replica 3 | Product Conversion per mol% catalyst Mean (%) | | |
| R1 = Cl, R2 = H | 9 | 9.2 | 7.2 | 8.5 | ± | 1 |
| R1 = F, R2 = H | 8.6 | 7.6 | 8.6 | 8.3 | ± | 0.6 |
| R1 = Br, R2 = H | 9.2 | 10 | 10 | 9.7 | ± | 0.5 |
| R1 = OMe, R2 = H | 10 | 10 | 9.5 | 9.8 | ± | 0.3 |
| R1 = H, R2 = H | 7.9 | 7.7 | 6.1 | 7.2 | ± | 1.0 |
| R1 = Me, R2 = H | 8.0 | 9.5 | 8.0 | 8.5 | ± | 0.9 |
| R1 = NO2, R2 = H | 8.1 | 10 | 8.6 | 8.9 | ± | 1.0 |
| R1 = Cl, R2 = CH3 | 0 | 0 | 0 | 0 | ± | 0 |

## **EcDHFR-Ala7-1** substrate scope*.*

**Table S7**. Substrate scope analysis of **EcDHFR-Ala7-1.** Conversions determined by 1H NMR.



|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Substrate | Replica 1 | Replica 2 | Replica 3 | Product Conversion per mol% catalyst Mean (%) | | |
| R1 = Cl, R2 = H | 33 | 60 | 54 | 49 | ± | 14 | |
| R1 = F, R2 = H | 30 | 18 | 25 | 24 | ± | 6 | |
| R1 = Br, R2 = H | 40 | 60 | 49 | 50 | ± | 10 | |
| R1 = OMe, R2 = H | 25 | 28 | 28 | 27 | ± | 2 | |
| R1 = H, R2 = H | 38 | 37 | 35 | 37 | ± | 2 | |
| R1 = Me, R2 = H | 33 | 43 | 36 | 37 | ± | 5 | |
| R1 = NO2, R2 = H | 36 | 24 | 37 | 32 | ± | 7 | |
| R1 = Cl, R2 = CH3 | 0 | 0 | 0 | 0 | ± | 0 | |

# 8. Iminium ion intermediate characterisation

To determine the formation of an iminium ion, the intermediate was trapped by reduction of the iminium ion species to a tertiary amine and subsequent analysis by mass spectrometry. The reaction was performed as follows. To a microcentrifuge tube was added 100 µL of either **LmrR-Phe93-1** or **DHFR-Ala7-1** (100 µg) in the reaction buffer PBS (50 mM NaPi, 150 mM NaCl, pH 7.0) and to this was added 20 eq. of cinnamaldehyde (1 mM) in 10 µL of methanol. The reaction was placed in a thermoshaker for 2 hours at 25 °C with agitation at 500 rpm. After 2 hours, 50 eq. of NaCNBH3 dissolved in methanol was added to the reactions and they were left overnight under the above reaction conditions.

The following day, the protein samples were buffer exchanged through micro centrifuge ultrafiltration 10 kDa cut-off columns (Millipore), simultaneously removing the small molecules and exchanging into tris buffer (50 mM, pH 8.0). 10 µL of the sample was removed and analysed by protein mass spectrometry. The remaining samples were then digested with 1:20 (chrymotrypsin:protein substrate) overnight at 25 °C. The digested samples were analysed by LC-MS (Fig. S7 and S8).

Chart, histogram

Description automatically generated

**S6**. Deconvoluted ESI mass spectra for the active proteins following the reaction with cinnamaldehyde and sodium cyanoborohydride. **A LmrR-Phe93-1** cinnamaldehyde adduct calculated mass = 14851 Da, measured mass = 14851 Da. **B DHFR-Ala7-1** cinnamaldehyde adduct calculated mass = 19555 Da, measured mass = 19555 Da. Mass of 19439 Da equates to the parent **DHFR-Ala7-1**. Mass of 19341 Da equates to the removal of the prolyl-group of unnatural amino acid **1**, a mass not observed in the protein prior to the sodium cyanoborohydride reaction.

Shape

Description automatically generated

**S7.** High-resolution mass analysis of the peptide fragment containing UAA **1** covalently bound to the substrate cinnamaldehyde **4** produced from the chymotrypsin digest of **LmrR-Phe93-1.**

A picture containing dark, night sky

Description automatically generated

**S8**. High-resolution mass analysis of the peptide fragment containing UAA **1** covalently bound to the substrate cinnamaldehyde **4** produced from the chymotrypsin digest of **EcDHFR-Ala7-1.**

# 9. NADPH hydride transfer selectivity characterisation

NADPD concentration was determined by measuring the UV absorbance at 340 nm using the extinction coefficient of 6200 M-1 cm-1 7. Hydride transfer selectivity reaction was performed as follows. To 50 µL (100 µg, 6.8 nmol) of **DHFR-Ala7-1** in the reaction buffer (50 mM NaPi, 150 mM NaCl, pH 7.0) in a microcentrifuge tube was added a 5 µL solution of cinnamaldehyde **4** (8.9 µg, 68 nmol, 1 eq) in methanol, followed by 5 µL solution of NADPD (28 µg, 136 nmol, 2 eq) in the reaction buffer. The reactions were placed in a thermomixer at 25 °C and 500 rpm for 18 hours. The reactions were halted by adding 200 µL of DCM and vortexed vigorously, then centrifuged to separate the layers (20,000 rcf, 3 min, rt) and 100 µL of the organic layer was removed and subjected to analysis by GC-MS. Control reactions with NADPH were performed alongside. Each reaction was performed in triplicate (Fig. S9).

A picture containing text

Description automatically generated

**S9**. Mass spectrometry analysis of the product produced from the reaction between cinnamaldehyde **6** and NADPH (top) and NADPD (Bottom) catalysed by **EcDHFR-Ala7-1**. Resulting ions were produced by electron ionisation.

# 10. Kinetic isotope effect study

To determine if there is a kinetic isotope effect at the step of the transfer hydrogenation reaction, the oxidation of NADPH and NADPD under saturating conditions was measured comparatively using a UV plate reader assay. To 90 µL of the enzyme (final concentration 25 µM) on a 96 well plate, in the reaction buffer (50 mM NaPi, 150 mM NaCl, pH 7.0) was added 5 µL of NADPH or NADPD dissolved in the same buffer (final concentration 250 µM). The samples were placed in the plate reader (BMG Labtech FLUOstar OPTIMA microplate reader) at a fixed temperature of 25 °C for five minutes. After the allotted time the plate was ejected and 5 µL of cinnamaldehyde (final concentration 1 mM) in methanol was added to each reaction. The oxidation of NADPH/D was monitored at a wavelength of 340 nm for two hours. The reactions were performed in triplicate. Measurements were taken every minute over 2 hours (Fig S10). Following completion of the reaction the slope was determined for both NADPH and NADPD. The kinetic isotope effect, if any, was determined using **equation 1**.

Equation 1.

Graphical user interface, chart

Description automatically generated

**S10.** Kinetic isotope effect (KIE) determination by UV absorbance.

# 11. Cofactor recycling assay

To a microcentrifuge tube was added 100 µM of the enzyme in the reaction buffer, 50 nM of glucose-6-phosphate dehydrogenase, 2 mM of glucose-6-phosphate and 1 mM of cinnamaldehyde in methanol. NADPH was then added in varying concentrations dissolved in the reaction buffer (50 mM NaPi, 150 mM NaCl, pH 7.0). The final reaction volume was made up to 100 µL with the reaction buffer. The reactions were placed in a thermomixer at 25 °C and 500 rpm for 18 hours. The reactions were halted by adding 200 µL of DCM and vortexed vigorously, then centrifuged to separate the layers (20,000 rcf, 3 min, rt) and 100 µL of the organic layer was removed and subjected to analysis by GC-FID. Reaction yield and total turnover number were determined from an average of three repeats. Total turnover number was defined as moles of product formed divided by moles of NADPH added (Table S8).

**Table S8.** Yield of product and total turnover number (TTN) with and without glucose-6-phosphate dehydrogenase as the recycling enzyme.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| [NADPH] (µM) | Yield of product (%) | | | | | | TTN |
| (+) G6PDH | | | (-) G6PDH | | |
| 0.01 | 10 | ± | 0 | 0 | ± | 0 | 10460 |
| 0.1 | 19 | ± | 0 | 0 | ± | 0 | 1880 |
| 1 | 63 | ± | 2 | 0 | ± | 0 | 632 |
| 5 | 76 | ± | 6 | 1 | ± | 0 | 153 |
| 10 | 91 | ± | 1 | 1 | ± | 0 | 91 |
| 50 | 72 | ± | 3 | 4 | ± | 0 | 14 |
| 125 | 80 | ± | 9 | 11 | ± | 1 | 6 |
| 250 | 71 | ± | 3 | 22 | ± | 1 | 3 |

# 12. Analytical chemistry

Protein liquid chromatography-mass spectrometry

Protein liquid chromatography mass spectrometry was acquired on a Waters Acquity H-Class UPLC system coupled to a Waters Synapt G2-Si quadrupole time of flight mass spectrometer. The column used was a Waters Acquity UPLC Protein C4 BEH column 300 Å, 1.7 mm (2.1 x 100 mm) held at 60 °C. The flow rate was 0.2 mL/min and the gradient employed is highlighted in the table below. Mass spectrometry data was collected in positive electrospray ionisation mode and the data analysed using Waters MassLynx 4.1. Deconvoluted mass spectra were generated using the maximum entropy 1 (MaxEnt1) software.

**Table S9**. Protein liquid chromatography-mass spectrometry chromatography parameters

|  |  |  |
| --- | --- | --- |
| **Time (min)** | **H2O (0.1% CHOOH)** | **ACN (0.1% CHOOH)** |
| 0 | 95 | 5 |
| 3 | 95 | 5 |
| 50 | 35 | 65 |
| 52 | 3 | 97 |
| 54 | 3 | 97 |
| 56 | 95 | 5 |
| 60 | 95 | 5 |

Analytical size exclusion chromatography

Analytical size exclusion chromatography was performed on an Agilent infinity 1260 HPLC. The column used was an Agilent Bio SEC-3 150 Å, 3 mm (4.6 x 300 mm) held at 20 °C. The elution was isocratic using PBS (50 mM NaPi, 150 mM NaCl, pH 7.0) and the flow rate was 1 mL/min. Detection was performed at 210 nm.

Gas chromatography – mass spectrometry (GC-MS)

GC-MS was performed on a Perkin Elmer Clarus 680 GC system coupled to a Perking Elmer Clarus SQ 8C quadrupole mass spectrometer in electron impact ionisation mode. The column used was a Perkin Elmer Elite-1 30 m (0.25 mm x 0.25 mm). The inlet temperature was set to 150 °C with a split ratio of 19:1. An injection volume of 1 µL was used. The temperature programme started at 50 °C and held for 1 minute, then ramped up to 220 °C at 15 °C / min and held for 2 minutes. Data was analysed using Perkin and Elmer TurboMass software.

Gas chromatography – flame ionisation detection (GC-FID)

GC-FID analysis. Enzyme kinetic characterisation was performed using a GC-FID assay measuring product yield of hydro-cinnamaldehyde **5** over time. GC-FID analysis was performed as follows: The instrument used was an Agilent 7890A GC system equipped with a flame ionisation detector, the column used was a Restek RE-bDEXsm 30 m (0.32 mm x 0.25 mm). The inlet temperature was set to 200 °C with a split ratio of 10:1. An injection volume of 3 µL was used. The temperature programme started at 80 °C and held for 2 minutes, then ramped up to 200 °C at 15 °C/min and held for 2 minutes. Data was analysed using Agilent ChemStation software.

Liquid chromatography-mass spectrometry (LC-MS)

Liquid chromatography-mass spectrometry (LC-MS) was acquired on a Waters Acquity H-Class UPLC system coupled to a Waters Synapt G2-Si quadrupole time of flight mass spectrometer. The column used was a Waters Acquity UPLC C18 BEH column 75 Å, 1.7 μm (2.1 x 100 mm) held at 40 °C. The flow rate was 0.3 mL/min and the gradient employed is highlighted in the table below. Mass spectrometry data was collected in positive electrospray ionisation mode and the data analysed using Waters MassLynx 4.1.

**Table S10.** Liquid chromatography-mass spectrometry chromatography parameters.

|  |  |  |
| --- | --- | --- |
| Time (min) | H2O (0.1% CHOOH) | ACN (0.1% CHOOH) |
| 0 | 97 | 3 |
| 1 | 97 | 3 |
| 30 | 40 | 60 |
| 31 | 5 | 95 |
| 33 | 5 | 95 |
| 34 | 97 | 3 |
| 40 | 97 | 3 |

# 13. Nucleotide and amino acid sequences

**MbPylRS**

Nucleotide Sequence:

ATGGATAAAAAACCGCTGGATGTGCTGATTAGCGCGACCGGCCTGTGGATGAGCCGTACCGGCACCCTGCATAAAATCAAACATCATGAAGTGAGCCGCAGCAAAATCTATATTGAAATGGCGTGCGGCGATCATCTGGTGGTGAACAACAGCCGTAGCTGCCGTACCGCGCGTGCGTTTCGTCATCATAAATACCGCAAAACCTGCAAACGTTGCCGTGTGAGCGATGAAGATATCAACAACTTTCTGACCCGTAGCACCGAAAGCAAAAACAGCGTGAAAGTGCGTGTGGTGAGCGCGCCGAAAGTGAAAAAAGCGATGCCGAAAAGCGTGAGCCGTGCGCCGAAACCGCTGGAAAATAGCGTGAGCGCGAAAGCGAGCACCAACACCAGCCGTAGCGTTCCGAGCCCGGCGAAAAGCACCCCGAACAGCAGCGTTCCGGCGTCTGCGCCGGCACCGAGCCTGACCCGCAGCCAGCTGGATCGTGTGGAAGCGCTGCTGTCTCCGGAAGATAAAATTAGCCTGAACATGGCGAAACCGTTTCGTGAACTGGAACCGGAACTGGTGACCCGTCGTAAAAACGATTTTCAGCGCCTGTATACCAACGATCGTGAAGATTATCTGGGCAAACTGGAACGTGATATCACCAAATTTTTTGTGGATCGCGGCTTTCTGGAAATTAAAAGCCCGATTCTGATTCCGGCGGAATATGTGGAACGTATGGGCATTAACAACGACACCGAACTGAGCAAACAAATTTTCCGCGTGGATAAAAACCTGTGCCTGCGTCCGATGCTGGCCCCGACCCTGTATAACTATCTGCGTAAACTGGATCGTATTCTGCCGGGTCCGATCAAAATTTTTGAAGTGGGCCCGTGCTATCGCAAAGAAAGCGATGGCAAAGAACACCTGGAAGAATTCACCATGGTTAACTTTTGCCAAATGGGCAGCGGCTGCACCCGTGAAAACCTGGAAGCGCTGATCAAAGAATTCCTGGATTATCTGGAAATCGACTTCGAAATTGTGGGCGATAGCTGCATGGTGTATGGCGATACCCTGGATATTATGCATGGCGATCTGGAACTGAGCAGCGCGGTGGTGGGTCCGGTTAGCCTGGATCGTGAATGGGGCATTGATAAACCGTGGATTGGCGCGGGTTTTGGCCTGGAACGTCTGCTGAAAGTGATGCATGGCTTCAAAAACATTAAACGTGCGAGCCGTAGCGAAAGCTACTATAACGGCATTAGCACGAACCTGTAA

Amino Acid Sequence:

MDKKPLDVLISATGLWMSRTGTLHKIKHHEVSRSKIYIEMACGDHLVVNNSRSCRTARAFRHHKYRKTCKRCRVSDEDINNFLTRSTESKNSVKVRVVSAPKVKKAMPKSVSRAPKPLENSVSAKASTNTSRSVPSPAKSTPNSSVPASAPAPSLTRSQLDRVEALLSPEDKISLNMAKPFRELEPELVTRRKNDFQRLYTNDREDYLGKLERDITKFFVDRGFLEIKSPILIPAEYVERMGINNDTELSKQIFRVDKNLCLRPMLAPTLYNYLRKLDRILPGPIKIFEVGPCYRKESDGKEHLEEFTMVNFCQMGSGCTRENLEALIKEFLDYLEIDFEIVGDSCMVYGDTLDIMHGDLELSSAVVGPVSLDREWGIDKPWIGAGFGLERLLKVMHGFKNIKRASRSESYYNGISTNL

**ThzKRS**

Nucleotide Sequence:

ATGGATAAAAAACCGCTGGATGTGCTGATTAGCGCGACCGGCCTGTGGATGAGCCGTACCGGCACCCTGCATAAAATCAAACATCATGAAGTGAGCCGCAGCAAAATCTATATTGAAATGGCGTGCGGCGATCATCTGGTGGTGAACAACAGCCGTAGCTGCCGTACCGCGCGTGCGTTTCGTCATCATAAATACCGCAAAACCTGCAAACGTTGCCGTGTGAGCGGTGAAGATATCAACAACTTTCTGACCCGTAGCACCGAAAGCAAAAACAGCGTGAAAGTGCGTGTGGTGAGCGCGCCGAAAGTGAAAAAAGCGATGCCGAAAAGCGTGAGCCGTGCGCCGAAACCGCTGGAAAATAGCGTGGGCGCGAAAGCGAGCACCAACACCAGCCGTAGCGTTCCGAGCCCGGCGAAAAGCACCCCGAACAGCAGCGTTCCGGCGTCTGCGCCGGCACCGAGCCTGACCCGCAGCCAGCTGGATCGTGTGGAAGCGCTGCTGTCTCCGGAAGATAAAATTAGCCTGAACATGGCGAAACCGTTTCGTGAACTGGAACCGGAACTGGTGACCCGTCGTAAAAACGATTTTCAGCGCCTGTATACCAACGATCGTGAAGATTATCTGGGCAAACTGGAACGTGATATCACCAAATTTTTTGTGGATCGCGGCTTTCTGGAAATTAAAAGCCCGATTCTGATTCCGGCGGAATATGTGGAACGTATGGGCATTAACAACGACACCGAACTGAGCAAACAAATTTTCCGCGTGGATAAAAACCTGTGCCTGCGTCCGATGCTGAGCCCGACCCTGTATAACTATCTGCGTAAACTGGATCGTATTCTGCCGGGTCCGATCAAAATTTTTGAAGTGGGCCCGTGCTATCGCAAAGAAAGCGATGGCAAAGAACACCTGGAAGAATTCACCATGGTTAACTTTGTGCAATTTGGCAGCGGCTGCACCCGTGAAAACCTGGAAGCGCTGATCAAAGAATTCCTGGATTATCTGGAAATCGACTTCGAAATTGTGGGCGGTAGCTGCATGGTGTATGGCGATACCCTGGATATTATGCATGGCGATCTGGAACTGAGCAGCGCGGTGGTGGGTCCGGTTAGCCTGGATCGTGAATGGGGCATTGATAAACCGTGGATTGGCGCGGGTTTTGGCCTGGAACGTCTGCTGAAAGTGATGCATGGCTTCAAAAACATTAAACGTGCGAGCCGTAGCGAAAGCTACTATAACGGCATTAGCACGAACCTGTAA

Amino Acid Sequence:

MDKKPLDVLISATGLWMSRTGTLHKIKHHEVSRSKIYIEMACGDHLVVNNSRSCRTARAFRHHKYRKTCKRCRVSGEDINNFLTRSTESKNSVKVRVVSAPKVKKAMPKSVSRAPKPLENSVGAKASTNTSRSVPSPAKSTPNSSVPASAPAPSLTRSQLDRVEALLSPEDKISLNMAKPFRELEPELVTRRKNDFQRLYTNDREDYLGKLERDITKFFVDRGFLEIKSPILIPAEYVERMGINNDTELSKQIFRVDKNLCLRPMLSPTLYNYLRKLDRILPGPIKIFEVGPCYRKESDGKEHLEEFTMVNFVQFGSGCTRENLEALIKEFLDYLEIDFEIVGGSCMVYGDTLDIMHGDLELSSAVVGPVSLDREWGIDKPWIGAGFGLERLLKVMHGFKNIKRASRSESYYNGISTNL

**MbPyl-tRNA**

GGGAACCTGATCATGTAGATCGAATGGACTCTAAATCCGTTCAGCCGGGTTAGATTCCCGGGGTTTCCGCCA

**Wild-type LmrR**

Nucleotide Sequence:

ATGGGTGCCGAAATCCCGAAAGAAATGCTGCGTGCTCAAACCAATGTCATCCTGCTGAATGTCCTGAAACAAGGCGATAACTATGTGTATGGCATTATCAAACAGGTGAAAGAAGCGAGCAACGGTGAAATGGAACTGAATGAAGCCACCCTGTATACGATTTTTGATCGTCTGGAACAGGACGGCATTATCAGCTCTTACTGGGGTGATGAAAGTCAAGGCGGTCGTCGCAAATATTACCGTCTGACCGAAATCGGCCATGAAAACATGCGCCTGGCGTTCGAATCCTGGAGTCGTGTGGACAAAATCATTGAAAATCTGGAAGCAAACAAAAAATCTGAAGCGATCAAAGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCTAA

Amino Acid Sequence:

MGAEIPKEMLRAQTNVILLNVLKQGDNYVYGIIKQVKEASNGEMELNEATLYTIFDRLEQDGIISSYWGDESQGGRRKYYRLTEIGHENMRLAFESWSRVDKIIENLEANKKSEAIKGSSHHHHHHSSG

**LmrR-Val15-TAG**

Nucleotide Sequence:

ATGGGTGCCGAAATCCCGAAAGAAATGCTGCGTGCTCAAACCAATTAGATCCTGCTGAATGTCCTGAAACAAGGCGATAACTATGTGTATGGCATTATCAAACAGGTGAAAGAAGCGAGCAACGGTGAAATGGAACTGAATGAAGCCACCCTGTATACGATTTTTGATCGTCTGGAACAGGACGGCATTATCAGCTCTTACTGGGGTGATGAAAGTCAAGGCGGTCGTCGCAAATATTACCGTCTGACCGAAATCGGCCATGAAAACATGCGCCTGGCGTTCGAATCCTGGAGTCGTGTGGACAAAATCATTGAAAATCTGGAAGCAAACAAAAAATCTGAAGCGATCAAAGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCTAA

Amino Acid Sequence:

MGAEIPKEMLRAQTN\*ILLNVLKQGDNYVYGIIKQVKEASNGEMELNEATLYTIFDRLEQDGIISSYWGDESQGGRRKYYRLTEIGHENMRLAFESWSRVDKIIENLEANKKSEAIKGSSHHHHHHSSG

**LmrR-Asp19-TAG**

Nucleotide Sequence:

ATGGGTGCCGAAATCCCGAAAGAAATGCTGCGTGCTCAAACCAATGTCATCCTGCTGTAGGTCCTGAAACAAGGCGATAACTATGTGTATGGCATTATCAAACAGGTGAAAGAAGCGAGCAACGGTGAAATGGAACTGAATGAAGCCACCCTGTATACGATTTTTGATCGTCTGGAACAGGACGGCATTATCAGCTCTTACTGGGGTGATGAAAGTCAAGGCGGTCGTCGCAAATATTACCGTCTGACCGAAATCGGCCATGAAAACATGCGCCTGGCGTTCGAATCCTGGAGTCGTGTGGACAAAATCATTGAAAATCTGGAAGCAAACAAAAAATCTGAAGCGATCAAAGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCTAA

Amino Acid Sequence:

MGAEIPKEMLRAQTNVILL\*VLKQGDNYVYGIIKQVKEASNGEMELNEATLYTIFDRLEQDGIISSYWGDESQGGRRKYYRLTEIGHENMRLAFESWSRVDKIIENLEANKKSEAIKGSSHHHHHHSSG

**LmrR-Met89-TAG**

Nucleotide Sequence:

ATGGGTGCCGAAATCCCGAAAGAAATGCTGCGTGCTCAAACCAATGTCATCCTGCTGAATGTCCTGAAACAAGGCGATAACTATGTGTATGGCATTATCAAACAGGTGAAAGAAGCGAGCAACGGTGAAATGGAACTGAATGAAGCCACCCTGTATACGATTTTTGATCGTCTGGAACAGGACGGCATTATCAGCTCTTACTGGGGTGATGAAAGTCAAGGCGGTCGTCGCAAATATTACCGTCTGACCGAAATCGGCCATGAAAACTAGCGCCTGGCGTTCGAATCCTGGAGTCGTGTGGACAAAATCATTGAAAATCTGGAAGCAAACAAAAAATCTGAAGCGATCAAAGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCTAA

Amino Acid Sequence:

MGAEIPKEMLRAQTNVILLNVLKQGDNYVYGIIKQVKEASNGEMELNEATLYTIFDRLEQDGIISSYWGDESQGGRRKYYRLTEIGHEN\*RLAFESWSRVDKIIENLEANKKSEAIKGSSHHHHHHSSG

**LmrR-Phe93-TAG**

Nucleotide Sequence:

GGTGCCGAAATCCCGAAAGAAATGCTGCGTGCTCAAACCAATGTCATCCTGCTGAATGTCCTGAAACAAGGCGATAACTATGTGTATGGCATTATCAAACAGGTGAAAGAAGCGAGCAACGGTGAAATGGAACTGAATGAAGCCACCCTGTATACGATTTTTGATCGTCTGGAACAGGACGGCATTATCAGCTCTTACTGGGGTGATGAAAGTCAAGGCGGTCGTCGCAAATATTACCGTCTGACCGAAATCGGCCATGAAAACATGCGCCTGGCGTAGGAATCCTGGAGTCGTGTGGACAAAATCATTGAAAATCTGGAAGCAAACAAAAAATCTGAAGCGATCAAAGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCTAA

Amino Acid Sequence:

MGAEIPKEMLRAQTNVILLNVLKQGDNYVYGIIKQVKEASNGEMELNEATLYTIFDRLEQDGIISSYWGDESQGGRRKYYRLTEIGHENMRLA\*ESWSRVDKIIENLEANKKSEAIKGSSHHHHHHSSG

**Wild-type EcDHFR**

Nucleotide Sequence:

ATGATCAGTCTGATTGCGGCGTTAGCGGTAGATCGCGTTATCGGCATGGAAAACGCCATGCCGTGGAACCTGCCTGCCGATCTCGCCTGGTTTAAACGCAACACCTTAAATAAACCCGTGATTATGGGCCGCCATACCTGGGAATCAATCGGTCGTCCGTTGCCAGGACGCAAAAATATTATCCTCAGCAGTCAACCGGGTACGGACGATCGCGTAACGTGGGTGAAGTCGGTGGATGAAGCCATCGCGGCGTGTGGTGACGTACCAGAAATCATGGTGATTGGCGGCGGTCGCGTTTATGAACAGTTCTTGCCAAAAGCGCAAAAACTGTATCTGACGCATATCGACGCAGAAGTGGAAGGCGACACCCATTTCCCGGATTACGAGCCGGATGACTGGGAATCGGTATTCAGCGAATTCCACGATGCTGATGCGCAGAACTCTCACAGCTATTGCTTTGAGATTCTGGAGCGGCGGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCTAA

Amino Acid Sequence:

MISLIAALAVDRVIGMENAMPWNLPADLAWFKRNTLNKPVIMGRHTWESIGRPLPGRKNIILSSQPGTDDRVTWVKSVDEAIAACGDVPEIMVIGGGRVYEQFLPKAQKLYLTHIDAEVEGDTHFPDYEPDDWESVFSEFHDADAQNSHSYCFEILERRGSSHHHHHHSSG

**EcDHFR-Ala7-TAG**

Nucleotide Sequence:

ATGATCAGTCTGATTGCGTAGTTAGCGGTAGATCGCGTTATCGGCATGGAAAACGCCATGCCGTGGAACCTGCCTGCCGATCTCGCCTGGTTTAAACGCAACACCTTAAATAAACCCGTGATTATGGGCCGCCATACCTGGGAATCAATCGGTCGTCCGTTGCCAGGACGCAAAAATATTATCCTCAGCAGTCAACCGGGTACGGACGATCGCGTAACGTGGGTGAAGTCGGTGGATGAAGCCATCGCGGCGTGTGGTGACGTACCAGAAATCATGGTGATTGGCGGCGGTCGCGTTTATGAACAGTTCTTGCCAAAAGCGCAAAAACTGTATCTGACGCATATCGACGCAGAAGTGGAAGGCGACACCCATTTCCCGGATTACGAGCCGGATGACTGGGAATCGGTATTCAGCGAATTCCACGATGCTGATGCGCAGAACTCTCACAGCTATTGCTTTGAGATTCTGGAGCGGCGGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCTAA

Amino Acid Sequence:

MISLIA\*LAVDRVIGMENAMPWNLPADLAWFKRNTLNKPVIMGRHTWESIGRPLPGRKNIILSSQPGTDDRVTWVKSVDEAIAACGDVPEIMVIGGGRVYEQFLPKAQKLYLTHIDAEVEGDTHFPDYEPDDWESVFSEFHDADAQNSHSYCFEILERRGSSHHHHHHSSG

**EcDHFR-Phe31-TAG**

Nucleotide Sequence:

ATGATCAGTCTGATTGCGGCGTTAGCGGTAGATCGCGTTATCGGCATGGAAAACGCCATGCCGTGGAACCTGCCTGCCGATCTCGCCTGGTAGAAACGCAACACCTTAAATAAACCCGTGATTATGGGCCGCCATACCTGGGAATCAATCGGTCGTCCGTTGCCAGGACGCAAAAATATTATCCTCAGCAGTCAACCGGGTACGGACGATCGCGTAACGTGGGTGAAGTCGGTGGATGAAGCCATCGCGGCGTGTGGTGACGTACCAGAAATCATGGTGATTGGCGGCGGTCGCGTTTATGAACAGTTCTTGCCAAAAGCGCAAAAACTGTATCTGACGCATATCGACGCAGAAGTGGAAGGCGACACCCATTTCCCGGATTACGAGCCGGATGACTGGGAATCGGTATTCAGCGAATTCCACGATGCTGATGCGCAGAACTCTCACAGCTATTGCTTTGAGATTCTGGAGCGGCGGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCTAA

Amino Acid Sequence:

MISLIAALAVDRVIGMENAMPWNLPADLAW\*KRNTLNKPVIMGRHTWESIGRPLPGRKNIILSSQPGTDDRVTWVKSVDEAIAACGDVPEIMVIGGGRVYEQFLPKAQKLYLTHIDAEVEGDTHFPDYEPDDWESVFSEFHDADAQNSHSYCFEILERRGSSHHHHHHSSG

**EcDHFR-Ser49-TAG**

Nucleotide Sequence:

ATGATCAGTCTGATTGCGGCGTTAGCGGTAGATCGCGTTATCGGCATGGAAAACGCCATGCCGTGGAACCTGCCTGCCGATCTCGCCTGGTTTAAACGCAACACCTTAAATAAACCCGTGATTATGGGCCGCCATACCTGGGAATAGATCGGTCGTCCGTTGCCAGGACGCAAAAATATTATCCTCAGCAGTCAACCGGGTACGGACGATCGCGTAACGTGGGTGAAGTCGGTGGATGAAGCCATCGCGGCGTGTGGTGACGTACCAGAAATCATGGTGATTGGCGGCGGTCGCGTTTATGAACAGTTCTTGCCAAAAGCGCAAAAACTGTATCTGACGCATATCGACGCAGAAGTGGAAGGCGACACCCATTTCCCGGATTACGAGCCGGATGACTGGGAATCGGTATTCAGCGAATTCCACGATGCTGATGCGCAGAACTCTCACAGCTATTGCTTTGAGATTCTGGAGCGGCGGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCTAA

Amino Acid Sequence:

MISLIAALAVDRVIGMENAMPWNLPADLAWFKRNTLNKPVIMGRHTWE\*IGRPLPGRKNIILSSQPGTDDRVTWVKSVDEAIAACGDVPEIMVIGGGRVYEQFLPKAQKLYLTHIDAEVEGDTHFPDYEPDDWESVFSEFHDADAQNSHSYCFEILERRGSSHHHHHHSSG

# 14. References

1 Polycarpo, C. R. et al. Pyrrolysine analogues as substrates for pyrrolysyl-tRNA synthetase. *FEBS Lett.* **580**, 6695-6700, (2006).

2 Nguyen, D. P. et al. Genetically encoded 1,2-aminothiols facilitate rapid and site-specific protein labeling via a bio-orthogonal cyanobenzothiazole condensation. *J. Am. Chem. Soc.* **133**, 11418-11421, (2011).

3 Ismail, M. et al. Straightforward Regeneration of Reduced Flavin Adenine Dinucleotide Required for Enzymatic Tryptophan Halogenation. *ACS Catal.* **9**, 1389-1395, (2019).

4 Loveridge, E. J. & Allemann, R. K. Effect of pH on hydride transfer by Escherichia coli dihydrofolate reductase. *ChemBioChem* **12**, 1258-1262, (2011).

5 Wang, K. et al. Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET. *Nat. Chem.* **6**, 393-403, (2014).

6 Williams, T. L. et al. Transferability of N-terminal mutations of pyrrolysyl-tRNA synthetase in one species to that in another species on unnatural amino acid incorporation efficiency. *Amino Acids* **53**, 89-96, (2021).

7 Swanwick, R. S., Maglia, G., Tey, L. H. & Allemann, R. K. Coupling of protein motions and hydrogen transfer during catalysis by Escherichia coli dihydrofolate reductase. *Biochem. J* **394**, 259-265, (2006).