Supplementary Information

Combined biosynthesis and site-specific incorporation of phenylalanine derivatives from aryl aldehydes or carboxylic acids

Authors: Shelby R. Anderson, Abigail P. Spangler, Amanda M. Forti, Michaela A. Jones, Roman M. Dickey, Aditya M. Kunjapur*

Department of Chemical & Biomolecular Engineering, University of Delaware, Newark DE 19716

*Correspondence: Correspondence to A.M.K. at kunjapur@udel.edu

Table of Contents

I. Supplementary Tables *5*

Supplementary Table 1. Estimated β-OH nsAA intermediate concentrations on the basis of the corresponding L-Phe absorbances. 5

Supplementary Table 2. In vivo desired nsAA titers (product e) from carboxylic acid or aldehyde precursors. 5

Supplementary Table 3. Oligonucleotides used in this study. 6

II. Supplementary Figures *7*

Supplementary Figure 1. Reactions catalyzed by engineered live *E. coli* strains that were transformed for heterologous expression of L-threonine transaldolases (TTAs), which catalyze the first step of the pathway starting from aldehyde substrates. 7

Supplementary Figure 2. A candidate beta-hydroxy nsAA, beta-hydroxy O-methyl-tyrosine (β-OH oMeTyr), is not readily accepted by engineered tyrosyl-tRNA synthetases 8

Supplementary Figure 3. Improvement in expression and product titer with appendage of SUMO-tag to PSDH. 9

Supplementary Figure 4. Evaluation of effect of overexpression of the aminotransferase and/or cofactor supplementation. 9

Supplementary Figure 5. Aldehyde stability of 1b in live cells. 10

Supplementary Figure 6. Aldehyde stability of 2b in live cells. 10

Supplementary Figure 7. Aldehyde stability of 3b in live cells. 11

Supplementary Figure 8. Aldehyde stability of 4b in live cells. 11

Supplementary Figure 9. Peak area of the β-OH nsAA intermediate. 12

Supplementary Figure 10. Relative concentrations of carboxylic acid, alcohol, aldehyde and desired nsAA (product 1e) 20 h after supplementation of 1 mM of the aldehyde (1b) to strains containing the biosynthesis pathway. 12

Supplementary Figure 11. Relative concentrations of carboxylic acid, alcohol, aldehyde and desired nsAA (product 2e) 20 h after supplementation of 1 mM of the aldehyde (2b) to strains containing the biosynthesis pathway. 13

Supplementary Figure 12. Relative concentrations of carboxylic acid, alcohol, aldehyde and desired nsAA (product 3e) 20 h after supplementation of 1 mM of the aldehyde (3b) to strains containing the biosynthesis pathway. 13

Supplementary Figure 13. Relative concentrations of carboxylic acid, alcohol, aldehyde and desired nsAA (product 4e) 20 h after supplementation of 1 mM of the aldehyde (4b) to strains containing the biosynthesis pathway. 14

Supplementary Figure 14. Demonstration of biosynthesis of encodable phenylalanine derivatives from carboxylic acids supplemented to live cells. 14

Supplementary Figure 15. Poor growth and conversion with two-plasmid system with both containing a medium—high copy number origin of replication. 15

Supplementary Figure 16. Improvement in cell growth with an updated plasmid design. 16

Supplementary Figure 17. Incorporation assay screen with 0.5 mM nsAA supplementation. . 17

Supplementary Figure 18. MS/MS spectra of ubiquitin-GFP fusion with peptide fragment corresponding to the site-specific incorporation of biosynthesized BipA (4e) from the aldehyde precursor 4b. 18

Supplementary Figure 19. MS/MS spectra of ubiquitin-GFP fusion with peptide fragment corresponding to the site-specific incorporation of biosynthesized pAzF (1e) from the carboxylic acid precursor 1a. 19

Supplementary Figure 20. Effect of different plasmid construct designs on desired nsAA (product e) titers. 20

Supplementary Figure 21. Expected size of the reporter protein containing desired biosynthesized nsAA via SDS-PAGE. 20

Supplementary Figure 22. In vivo biosynthesis of the nsAA 5e from aldehyde precursor (5b), as confirmed by MS. 21

Supplementary Figure 23. In vivo biosynthesis of the nsAA 6e from aldehyde precursor (6b), as confirmed by MS. 22

Supplementary Figure 24. In vivo biosynthesis of the nsAA 7e from aldehyde precursor (7b), as confirmed by MS. 23

Supplementary Figure 25. In vivo biosynthesis of the nsAA 8e from aldehyde precursor (8b), as confirmed by MS. 24

Supplementary Figure 26. In vivo biosynthesis of the nsAA 9e from aldehyde precursor (9b), as confirmed by MS. 25

Supplementary Figure 27. In vivo biosynthesis of the nsAA 10e from aldehyde precursor (10b), as confirmed by MS. 26

Supplementary Figure 28. In vivo biosynthesis of the nsAA 11e from aldehyde precursor (11b), as confirmed by MS. 27

Supplementary Figure 29. In vivo biosynthesis of the nsAA 12e from aldehyde precursor (12b), as confirmed by MS. 28

Supplementary Figure 30. In vivo biosynthesis of the nsAA 5e from carboxylic acid precursor (5a), as confirmed by MS. 29

Supplementary Figure 31. In vivo biosynthesis of the nsAA 6e from carboxylic acid precursor (6a), as confirmed by MS. 30

Supplementary Figure 32. In vivo biosynthesis of the nsAA 7e from carboxylic acid precursor (7a), as confirmed by MS. 31

Supplementary Figure 33. In vivo biosynthesis of the nsAA 8e from carboxylic acid precursor (8a), as confirmed by MS. 32

Supplementary Figure 34. In vivo biosynthesis of the nsAA 9e from carboxylic acid precursor (9a), as confirmed by MS. 33

Supplementary Figure 35. In vivo biosynthesis of the nsAA 11e from carboxylic acid precursor (11a), as confirmed by MS. 34

Supplementary Figure 36. In vivo biosynthesis of the nsAA 12e from carboxylic acid precursor (12a), as confirmed by MS. 35

Supplementary Figure 37. Monitoring the growth of the engineered strains containing the genetic constructs for biosynthesis and incorporation of nsAAs. 36

Supplementary Figure 38. C.f.u.s ml^{-1} values of DEP.e5 transformed with pSA-2. 37

Supplementary Figure 39. Monitoring for escape of the synthetic auxotroph transformed with the biosynthetic pathway following provision with 4b. 37

Supplementary Figure 40. Monitoring for escape of the synthetic auxotroph transformed with the biosynthetic pathway. 38

Supplementary Figure 41. Monitoring the OD_{600} of the film for growth. 38

III. Supplementary References *39*

I. Supplementary Tables

Supplementary Table 1. Estimated β-OH nsAA intermediate concentrations on the basis of the corresponding absorbances for the non-hydroxylated analogs.

Strain	β-OH nsAA titer (μM)			
	1c	2c	3c	4c
MG1655	8 ± 2	33 ± 0.1	0 ± 0	2 ± 0.2
RARE	22 ± 1	2 ± 4	87 ± 8	5 ± 0.6
RARE.Δ16	20 ± 2	0 ± 0	87 ± 20	4 ± 0
ROAR	16 ± 2	0 ± 0	88 ± 8	0 ± 0

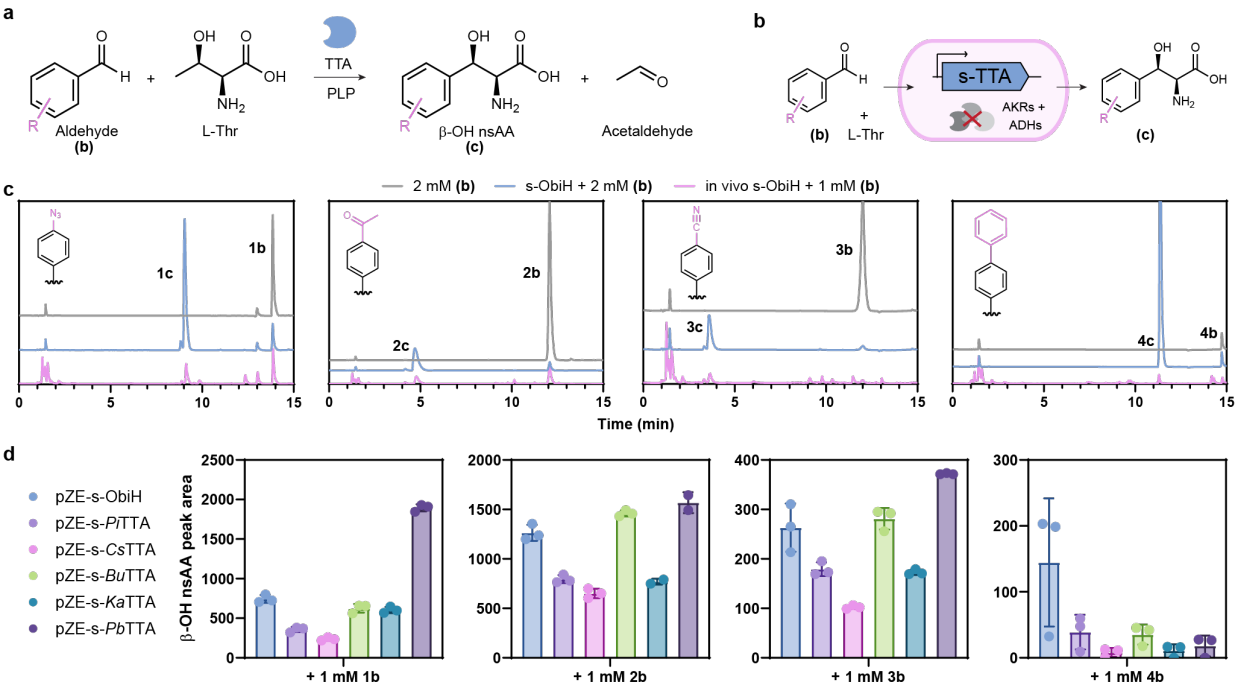
Supplementary Table 2. In vivo desired nsAA titers (product e) from carboxylic acid or aldehyde precursors.

Carboxylic acid precursor		Aldehyde precursor	
Chemistry	nsAA titer (μM)	Chemistry	nsAA titer (μM)
5a	67 ± 10	5b	141 ± 7
6a	170 ± 5	6b	272 ± 52
7a	62 ± 1	7b	54 ± 15
9a	3 ± 1	9b	36 ± 14
10a	0 ± 0	10b	10 ± 1
11a	19 ± 2	11b	101 ± 3
12a	13 ± 1	12b	117 ± 58

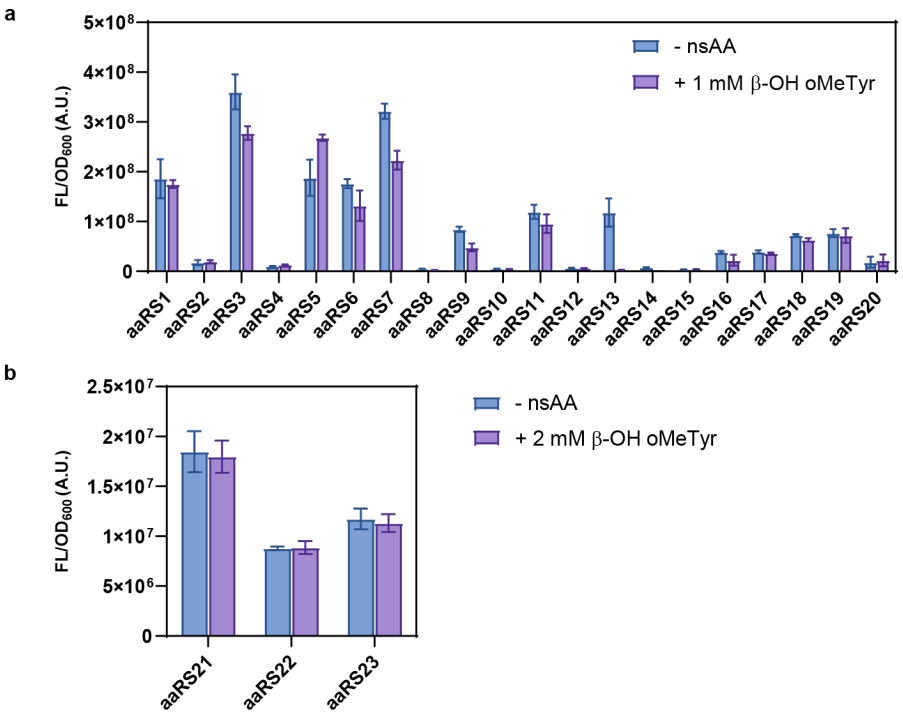
Supplementary Table 3. Oligonucleotides used in this study.

Oligo Name	Sequence
pACYC bbone F	AAGCTTGATGGGGGATC
pACYC bbone R	GGTATATCTCCTTATTAAAGTTAAAC
s-TTA pACYC ins F	CTTTAATAAGGAGATATACCATGGGCAGCAGCCATCA
s-PbTTA pACYC ins R	GGATCCCCCATCAAGCTTTTAGAATAACTTCTCGTAGATCTCGT
s-ObiH pACYC ins R	GGATCCCCCATCAAGCTTTTAACGTTGGGCTCCTTGGT
pACYC-CAR bbone R	GCCCATGGTATATCTCCTTATT
pACYC-CAR bbone F	ACCATGAAAATCTATGGCATTT
pACYC-CAR ins R	GGACGATCCATGTAAATGCCATAG
pACYC seq F	CTCCCTTATGCGACTCC
pZE seq F	TTTCGTCTTCACCTCGAG
pZE-His-SUMO2-PbTTA ins F	CCCATCGCGAGCAAATCGGGGGCGAAACCTCCCTGAAG
pZE-His-SUMO2-TTA ins R	CCATGGGATCCCCCATCAAG
pZE-His bbone F	TAACTTGATGGGGGATCCC
pZE-His-SUMO2 bbone R	CCCGATTTGCTCGCGAT
pZE-His-SUMO2-ObiH ins F	CCCATCGCGAGCAAATCGGGGGCTCCAATGTCAAGCAACA
pZE-spacer-s-RpPSDH F	GACGTCTAATAAAAGGAGATATACCATGGGCAGCAGCCATCA
pZE-s-RpPSDH R	CCATGGGATCCCCCATCAAGTTATCCCAGCCCAAGGTGAC
pZE-SUMO2-ObiH-spacer R	GGTATATCTCCTTTTATTAGACGTCTTAACGTTGGGCTCCTTG
pZE-SUMO2-PbTTA-spacer R	GGTATATCTCCTTTTATTAGACGTCTTAGAATAACTTCTCGTAGATCTCG
pCola-s-RpPSDH F	TATAAGAAGGAGATATACATATGGGCAGCAGCCATCA
pCola-s-RpPSDH R	GTTTCTTTACCAGACTCGAGTTATCCCAGCCCAAGGTGAC
pCola bbone F	ACTCGAGTCTGGTAAAGAAAC
pCola bbone ins1 R	ATGTATATCTCCTTCTTATACTTAAC
pCola bbone ins2 F	TCGAGATCCCGGTGCCTAATGTGAGCTAACTTACATTAATTGC
pCola bbone ins2 R	TCGCGGTATGGCATGATAGCGGGCAGGGTCGTTAAATAG
pCola bbone F (2)	GCTATCATGCCATACCGC
pCola bbone R (2)	ATTAGGCACCGGGATCTC
pCola seq R	GACGAGTCCATGTGCTG
pZE-spacer-sRpPSDHins2 F	GACGTCTAATAAAAGGAGATATACCATGGGCAGCAGCCATCA
pZE-sRpPSDH ins2 R	CCATGGGATCCCCCATCAAGTTATCCCAGCCCAAGGTGAC
pZE-sObaG-spacer R	GGTATATCTCCTTTTATTAGACGTCTTAACGTTGGGCTCCTTG
pZE-sPbTTA-spacer R	GGTATATCTCCTTTTATTAGACGTCTTAGAATAACTTCTCGTAGATCTCG
pZE-J23100 F	TTGACGGCTAGCTCAGTCCTAGGTACAGTGCTAGCGAATTCATTAAAGAGGAGAAAGGT
pZE-J23100 R	GCTAGCACTGTACCTAGGACTGAGCTAGCCGTCAACTCGAGGTGAAGACGAAAGG
pZE-J23108 F	CTGACAGCTAGCTCAGTCCTAGGTATAATGCTAGCGAATTCATTAAAGAGGAGAAAGGT
pZE-J23108 R	GCTAGCATTATACCTAGGACTGAGCTAGCTGTCAGCTCGAGGTGAAGACGAAAGG
CloDF_strR_F	TTATTTGCCGACTACCTTGG
CloDF_strR_R	GATCAAAGGATCTTCTTGAGATC
pZEbbonetoCloDFstr_F	CTCAAGAAGATCCTTTGATCGCGAAACGATCCTCATCCTG
pZEbbonetoCloDFstr_R	CCAAGGTAGTCGGCAAATAAAACGCCAGCAACGCG
pZE-sRpPSDH no His F	TTAAAGAGGAGAAAGGTACCATGTCCCTGCAGGACTCG
pZE-sRpPSDH no His R	CCATGGGATCCCCCATCAAGTTATCCCAGCCCAAGGTGAC
pZE-sObiH no His (orthog) F	TTGGTGCCCCGTGGGT
pZE-sObiH-sRp no His spacer R	CATGGTATATCTCCTTTTATTAGACGTCTTAACG
pZE-sObiH-sRpPSDH bbone F (no His)	CTAATAAAAGGAGATATACCATGTCCCTGCAGGACTC
pZE-SUMO bbone R	GCAGACCCACGGGGCACCAACATGGTACCTTTCTCCTCTTT
pZE-sPbTTA-sRp no His spacer R	CATGGTATATCTCCTTTTATTAGACGTCTTAGAAT
carb,ColA pRepSynth ins F	CAATAACTGCCTTAAAAAAAATTTCTGGCGGCACGATG
carb,ColA pRepSynth ins R	CAGGTGCTACATTTGAAGAGTGGTGTCGGGAATCCGTAAA

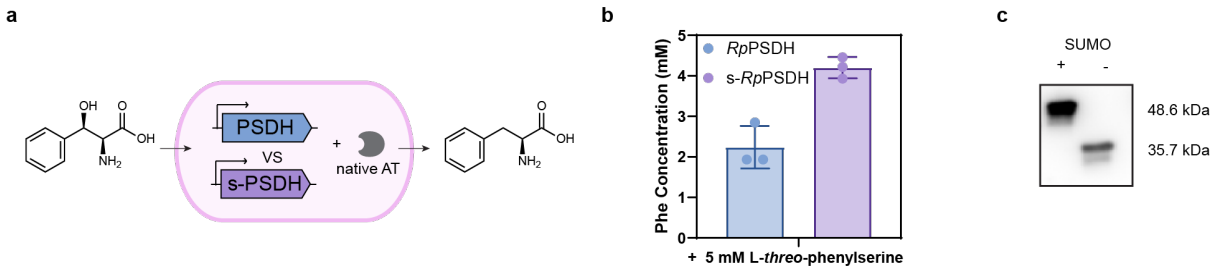
II. Supplementary Figures



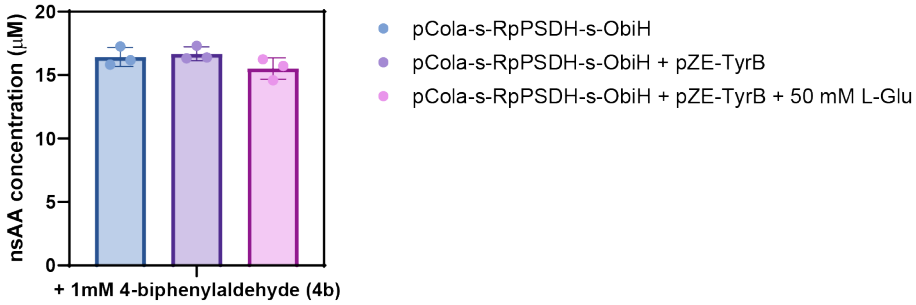
Supplementary Figure 1. Reactions catalyzed by engineered live *E. coli* strains that were transformed for heterologous expression of L-threonine transaldolases (TTAs), which catalyze the first step of the pathway starting from aldehyde substrates. a. Reaction catalyzed by TTA. **b.** Illustration of the reaction format with a live strain engineered both for aldehyde retention and for expression of one of six SUMO-tagged TTA variants (s-TTA). **c.** HPLC traces obtained when supplying either just the aldehyde substrate (2 mM **1b**-**4b**) as a standard (gray), an endpoint reaction catalyzed by the purified s-ObiH after provision of 2 mM aldehyde substrate (to determine the RT of the expected β-OH nsAA product **1c**-**4c**, shown in blue), or an endpoint reaction catalyzed by cells that express s-ObiH and were supplied 1 mM aldehyde substrate (pink). Four different aldehyde substrates were tested (#**b**), each corresponding to a commonly used nsAA for genetic code expansion. **d.** Endpoint HPLC measurements demonstrating how the peak area of the expected β-OH nsAA product varies based on the TTA homolog that is chosen for heterologous expression in live cells. Sample sizes are n = 3 using biological replicates and data shown are mean ± s.d. Source data are provided as a Source Data file.



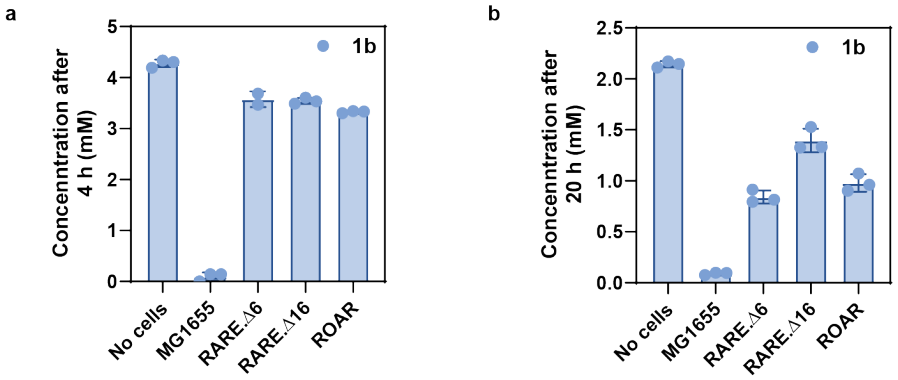
Supplementary Figure 2. A candidate beta-hydroxy nsAA, beta-hydroxy O-methyl-tyrosine (β-OH oMeTyr), is not readily accepted by engineered tyrosyl-tRNA synthetases. a. Standard incorporation assay in which positive incorporation is measured by an increase in the fluorescence normalized by optical density of cells supplemented nsAA compared to cells in the absence of nsAA. No significant increase in FL/OD was observed for cells supplemented 1mM β-OH oMeTyr when screened against a panel of 20 *Mj*TyrRSs. **b.** No significant increase in FL/OD was observed for cells supplemented 2mM β-OH oMeTyr when screened against a panel of three additional PylRSs. Sample sizes are $n = 3$ using biological replicates. Data are the mean calculated as the ratio of the average of the replicates in a condition compared to a baseline average of replicates lacking nsAA supplementation (fold change) ± s.e.m. Source data are provided as a Source Data file.



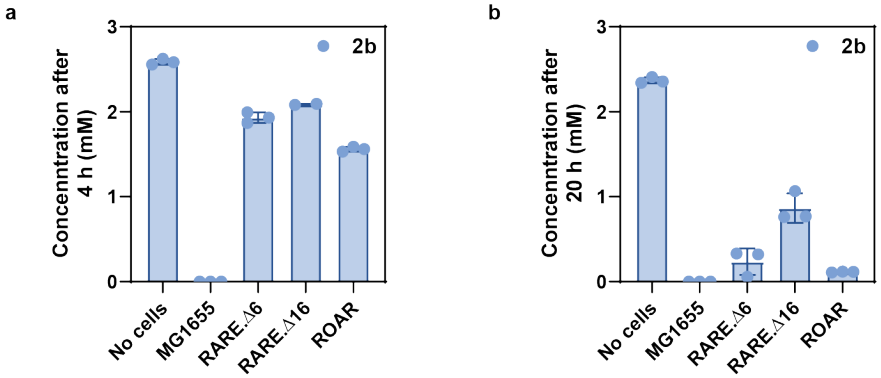
Supplementary Figure 3. Improvement in expression and product titer with appendage of SUMO-tag to PSDH. a. Comparison of the performance of cells that are transformed to express either *Rp*PSDH or SUMO-tagged *Rp*PSDH (s-*Rp*PSDH) at converting supplemented phenylserine to phenylalanine. **b.** Endpoint phenylalanine concentration in culture media is measured using HPLC. **c.** Relative PSDH expression comparing variants with or without appendage of a SUMO-tag as analyzed by Western blot. Sample sizes are $n = 3$ using biological replicates and data shown are mean ± s.d. Source data are provided as a Source Data file.



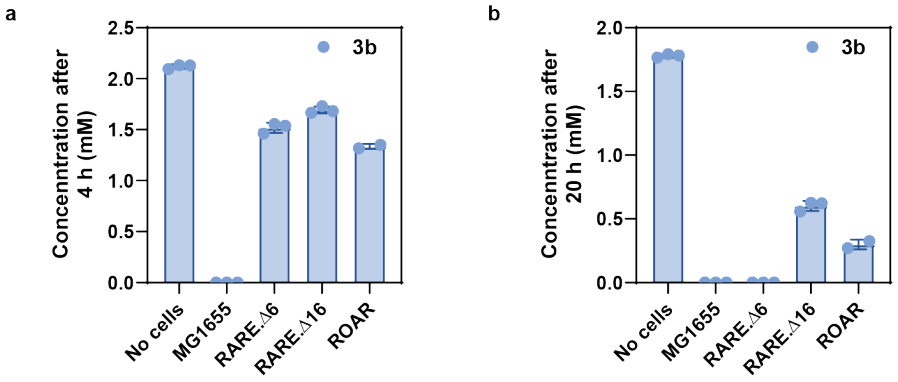
Supplementary Figure 4. Evaluation of effect of overexpression of the aminotransferase and/or cofactor supplementation. We co-transformed the pCola vector containing the genes for s-TTA and s-*Rp*PSDH with a pZE vector containing the gene for TyrB expression in the RARE.Δ16 strain and measured nsAA titers in the presence and absence of excess L-Glu 20 h after 1 mM aldehyde addition. However, we did not observe a benefit from TyrB overexpression or L-Glu supplementation under the conditions tested. See Strains S10 and S15 in **Table S1** for further strain details. Sample sizes are $n = 3$ using biological replicates and data shown are mean ± s.d. Source data are provided as a Source Data file.



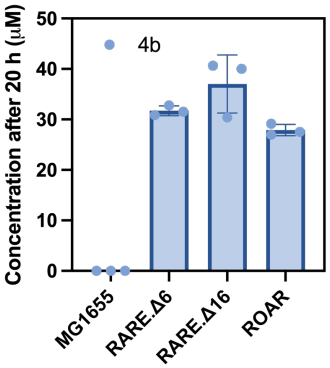
Supplementary Figure 5. Aldehyde stability of 1b in live cells. Stability of 5 mM aldehyde externally supplemented at mid-exponential growth phase to the wild-type MG1655, RARE.Δ6 (or RARE), RARE.Δ16, or ROAR strains in MOPS EZ Rich Defined Media + 2% glucose after incubation at 30 °C for **a.** 4 h and **b.** 20 h, as measured by HPLC. A "no cells" condition is provided as a control to account for aldehyde volatility. Sample sizes are $n = 3$ using biological replicates and data shown are mean ± s.d. Source data are provided as a Source Data file.



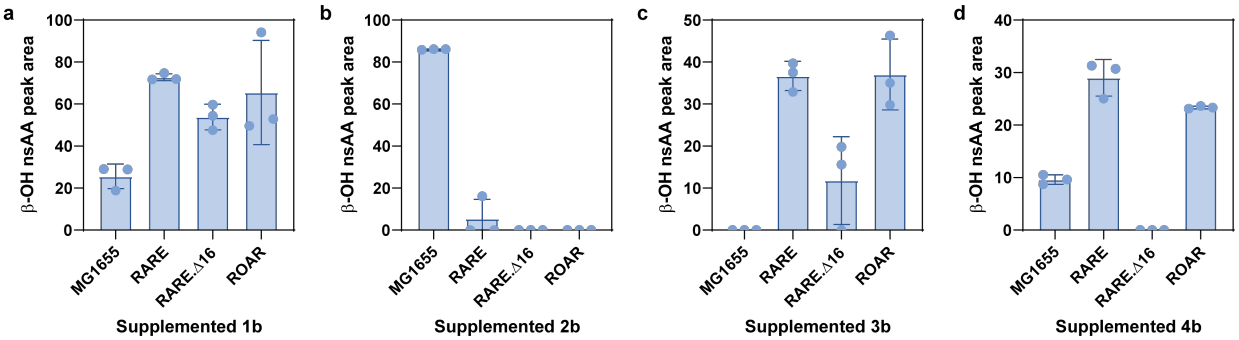
Supplementary Figure 6. Aldehyde stability of 2b in live cells. Stability of 5 mM aldehyde externally supplemented at mid-exponential growth phase to the wild-type MG1655, RARE.Δ6 (or RARE), RARE.Δ16, or ROAR strains in MOPS EZ Rich Defined Media + 2% glucose after incubation at 30 °C for **a.** 4 h and **b.** 20 h, as measured by HPLC. A "no cells" condition is provided as a control to account for aldehyde volatility. Sample sizes are $n = 3$ using biological replicates and data shown are mean ± s.d. Source data are provided as a Source Data file.



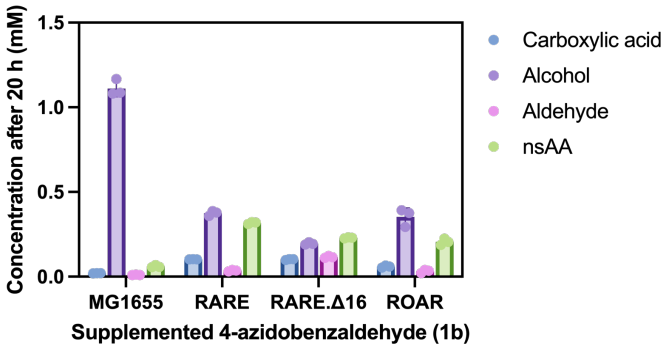
Supplementary Figure 7. Aldehyde stability of 3b in live cells. Stability of 5 mM aldehyde externally supplemented at mid-exponential growth phase to the wild-type MG1655, RARE.Δ6 (or RARE), RARE.Δ16, or ROAR strains in MOPS EZ Rich Defined Media + 2% glucose after incubation at 30 °C for **a.** 4 h and **b.** 20 h, as measured by HPLC. A "no cells" condition is provided as a control to account for aldehyde volatility. Sample sizes are $n = 3$ using biological replicates and data shown are mean ± s.d. Source data are provided as a Source Data file.



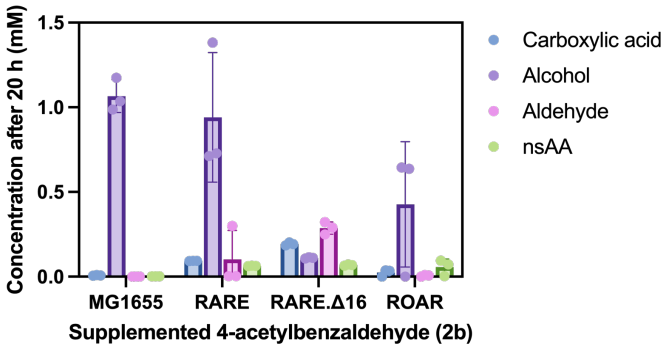
Supplementary Figure 8. Aldehyde stability of 4b in live cells. Stability of 0.5 mM aldehyde externally supplemented at mid-exponential growth phase to the wild-type MG1655, RARE.Δ6 (or RARE), RARE.Δ16, or ROAR strains in MOPS EZ Rich Defined Media + 2% glucose after incubation at 30 °C for 20 h, as measured by HPLC. Sample sizes are $n = 3$ using biological replicates and data shown are mean ± s.d. Source data are provided as a Source Data file.



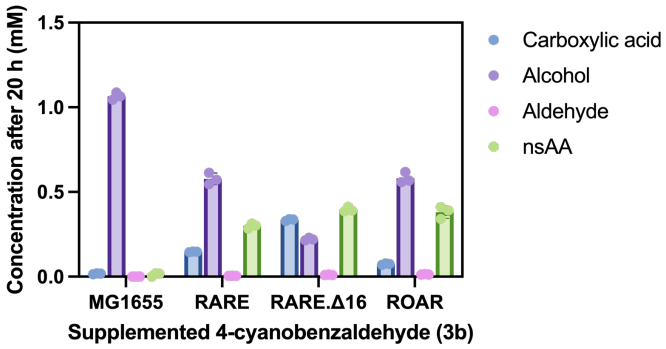
Supplementary Figure 9. Peak area of the β-OH nsAA intermediate. Analysis by HPLC 20 h after supplementation of 1 mM of the aldehyde to strains containing the biosynthesis pathway (TTA, PSDH, AT). **a.** Peak area of the β-OH nsAA intermediate **1c** after supplementation of **1b**. **b.** Peak area of **2c** after supplementation of **2b**. **c.** Peak area of **3c** after supplementation of **3b**. **d.** Peak area of **4c** after supplementation of **4b**. Sample sizes are $n = 3$ using biological replicates and data shown are mean ± s.d. Source data are provided as a Source Data file. There is some minor observation of β-OH nsAA intermediate, but at less than 5% of the peak area relative to that of the RARE.Δ16 strain expressing s-PbTTA only (no PSDH), as shown in **Supplementary Figure 1d**.



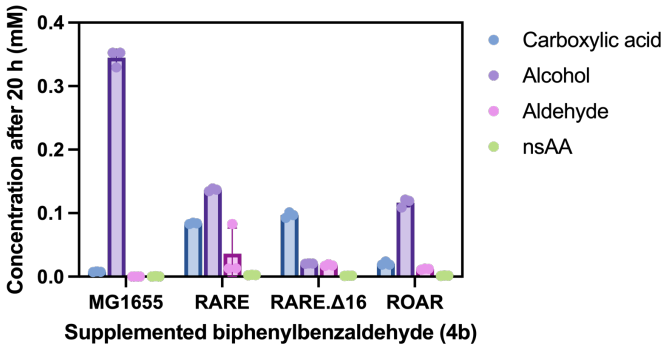
Supplementary Figure 10. Relative concentrations of carboxylic acid, alcohol, aldehyde and desired nsAA (product 1e) 20 h after supplementation of 1 mM of the aldehyde (1b) to strains containing the biosynthesis pathway. The carboxylic acid byproduct is a result of oxidation, and the alcohol byproduct is a result of reduction of the supplemented aldehyde in a cellular environment. The RARE and RARE.Δ16 strains are engineered to prevent aldehyde reduction, and the ROAR strain is additionally designed to prevent aldehyde oxidation. Sample sizes are $n = 3$ using biological replicates and data shown are mean ± s.d. Source data are provided as a Source Data file.



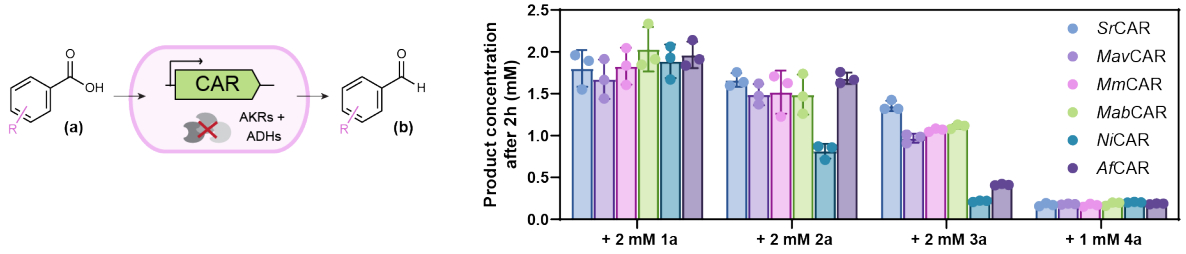
Supplementary Figure 11. Relative concentrations of carboxylic acid, alcohol, aldehyde and desired nsAA (product 2e) 20 h after supplementation of 1 mM of the aldehyde (2b) to strains containing the biosynthesis pathway. The carboxylic acid byproduct is a result of oxidation, and the alcohol byproduct is a result of reduction of the supplemented aldehyde in a cellular environment. The RARE and RARE.Δ16 strains are engineered to prevent aldehyde reduction, and the ROAR strain is additionally designed to prevent aldehyde oxidation. Sample sizes are $n = 3$ using biological replicates and data shown are mean ± s.d. Source data are provided as a Source Data file.



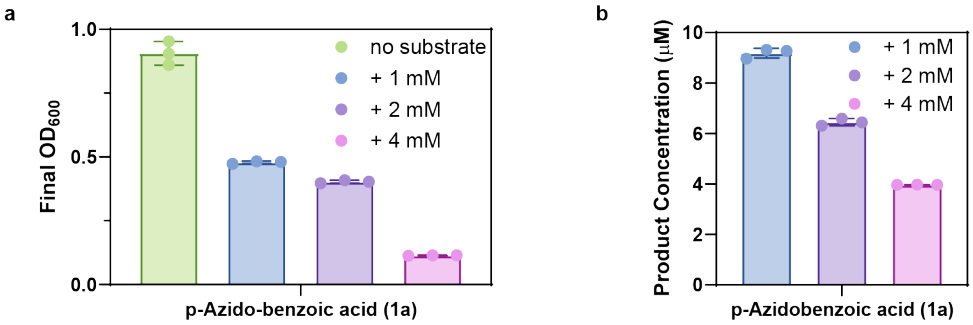
Supplementary Figure 12. Relative concentrations of carboxylic acid, alcohol, aldehyde and desired nsAA (product 3e) 20 h after supplementation of 1 mM of the aldehyde (3b) to strains containing the biosynthesis pathway. The carboxylic acid byproduct is a result of oxidation, and the alcohol byproduct is a result of reduction of the supplemented aldehyde in a cellular environment. The RARE and RARE.Δ16 strains are engineered to prevent aldehyde reduction, and the ROAR strain is additionally designed to prevent aldehyde oxidation. Sample sizes are $n = 3$ using biological replicates and data shown are mean ± s.d. Source data are provided as a Source Data file.



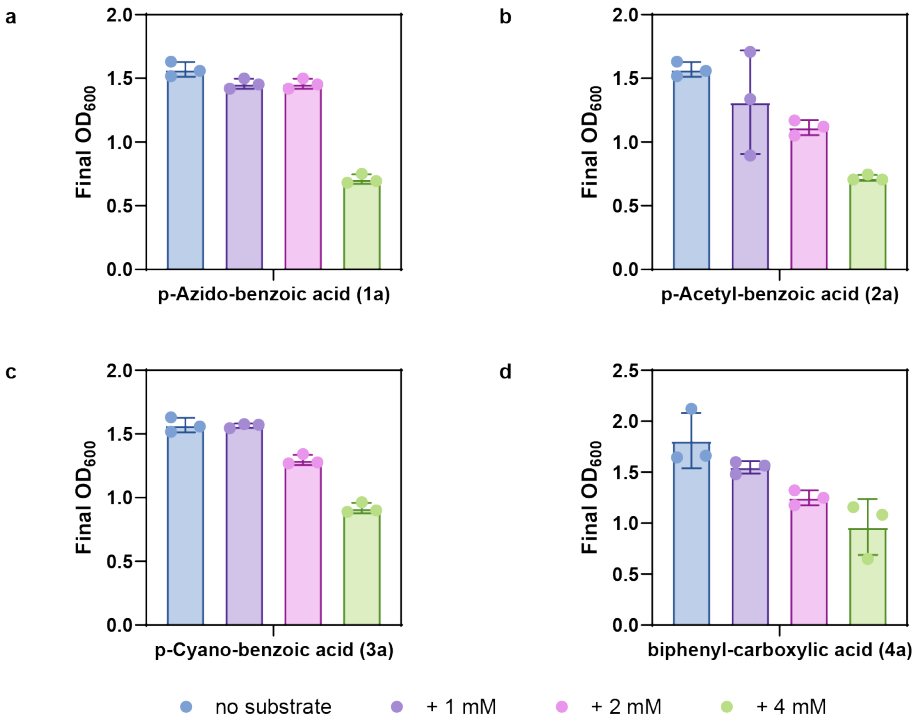
Supplementary Figure 13. Relative concentrations of carboxylic acid, alcohol, aldehyde and desired nsAA (product 4e) 20 h after supplementation of 1 mM of the aldehyde (4b) to strains containing the biosynthesis pathway. The carboxylic acid byproduct is a result of oxidation, and the alcohol byproduct is a result of reduction of the supplemented aldehyde in a cellular environment. The RARE and RARE.Δ16 strains are engineered to prevent aldehyde reduction, and the ROAR strain is additionally designed to prevent aldehyde oxidation. Sample sizes are $n = 3$ using biological replicates and data shown are mean ± s.d. Source data are provided as a Source Data file.



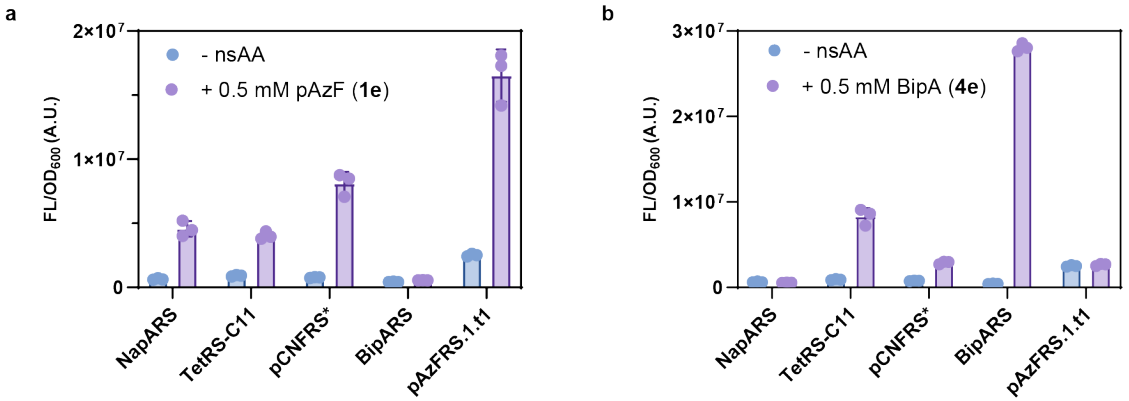
Supplementary Figure 14. Demonstration of biosynthesis of encodable phenylalanine derivatives from carboxylic acids supplemented to live cells. Depiction of the construct design in which the RARE.Δ16 strain was transformed to express a carboxylic acid reductase (CAR) for production of aryl aldehydes from carboxylic acids, as shown on the left. Comparison of the performance of cells that are transformed to express one of six CAR homologs for converting supplemented carboxylic acids **1a-4a** to their corresponding aldehydes, as shown in the bar graph on the right. An endpoint of 2 h is chosen to identify fast-acting variants. Sample sizes are $n = 3$ using biological replicates and data shown are mean ± s.d. Source data are provided as a Source Data file.



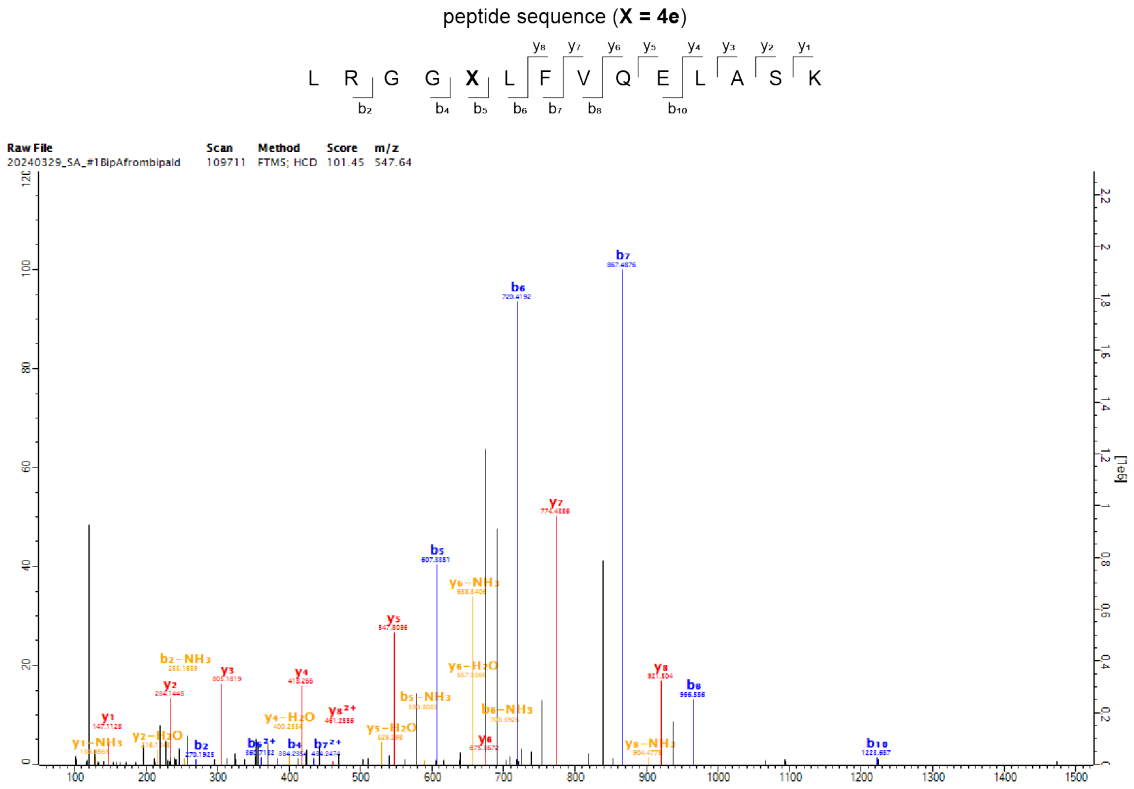
Supplementary Figure 15. Poor growth and conversion with two-plasmid system with both containing a medium—high copy number origin of replication. Initially, we utilized a two-plasmid system, both containing a medium-high (15-40) copy number origin of replication (ori) for both the *Mav*CAR with *Bs*Sfp (pZE vector, ColE1 ori) and the s-TTA with s-*Rp*PSDH (pCOLA vector, ColA ori) constructs. However, this initial design resulted in impacts on cell growth, as monitored by final OD_{600} in panel **a** and by poor titers, as monitored by HPLC in panel **b**. **a.** Growth of Strain S24 on externally supplemented **1a** (0 - 4 mM) as compared to a lack of substrate. **b.** Product titers of **1e** after 20 h at varying concentrations of supplemented carboxylic acid **1a** (0 – 4 mM). Sample sizes are $n = 3$ using biological replicates and data shown are mean ± s.d. Source data are provided as a Source Data file.



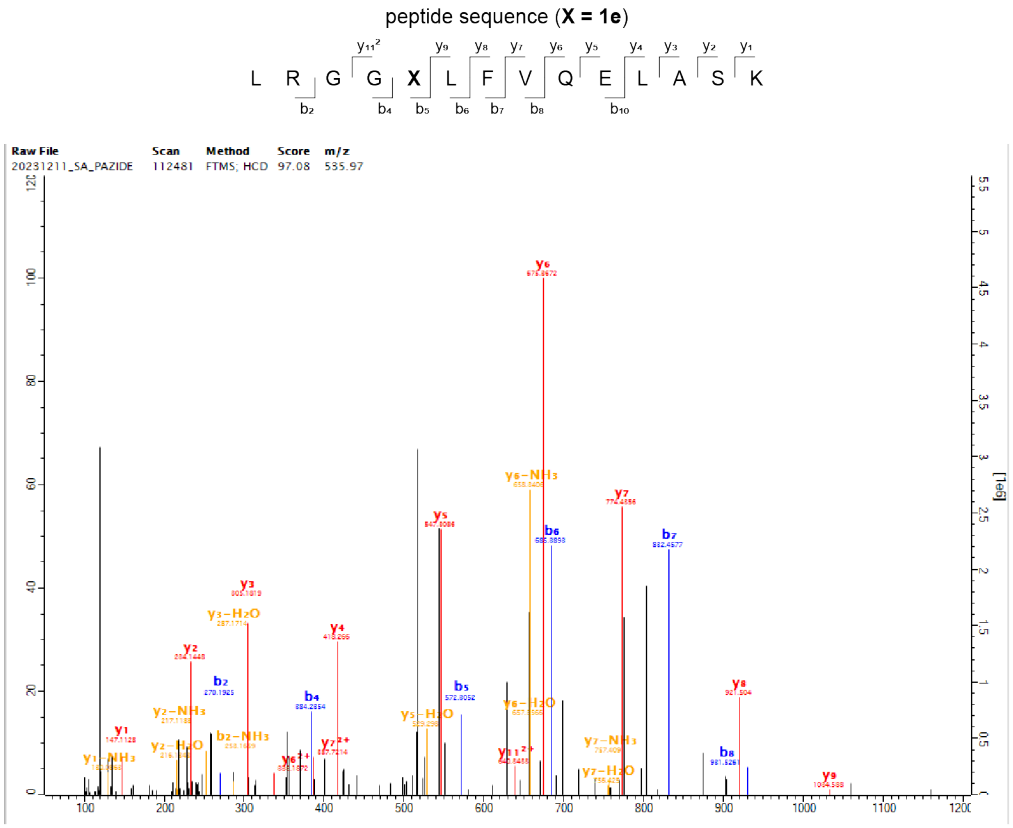
Supplementary Figure 16. Improvement in cell growth with an updated plasmid design. With the second design, consisting of the original pCOLA vector containing the genes s-TTA and s-*Rp*PSDH co-transformed with the pACYC vector containing mavCAR and *Bs*Sfp, we observed more robust growth, as monitored by final OD_{600}. In panel **a.** Growth of Strain S26 on externally supplemented **1a**, panel **b.** growth of Strain S26 on externally supplemented **2a**, panel **c**. growth of Strain S26 on externally supplemented **3a**, and panel **d.** growth of Strain S25 on externally supplemented **4a** at varying concentrations of each carboxylic acid (0 – 4 mM). Sample sizes are $n = 3$ using biological replicates and data shown are mean ± s.d. Source data are provided as a Source Data file.



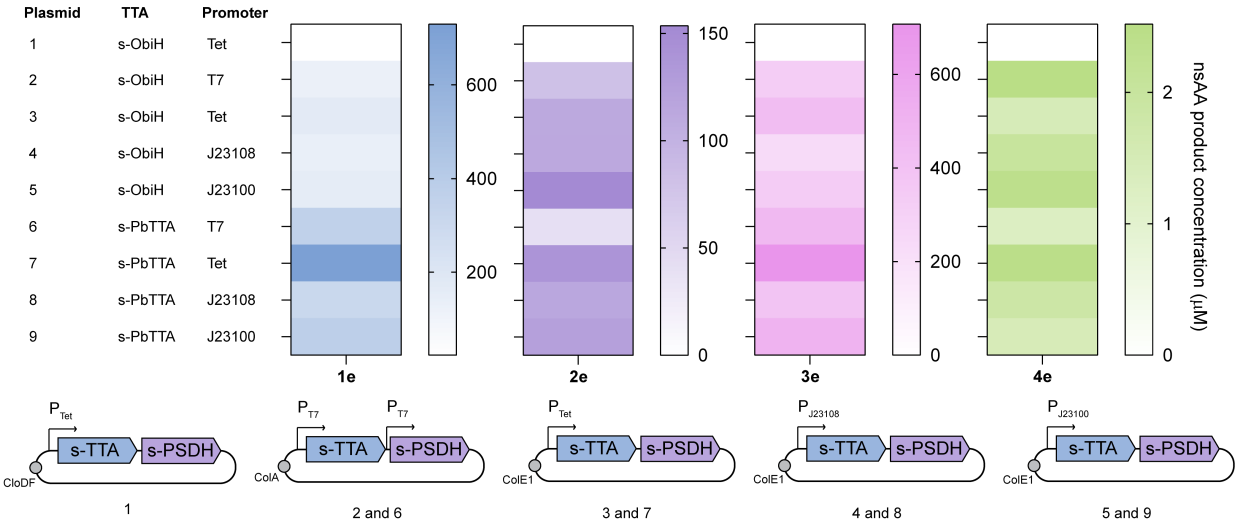
Supplementary Figure 17. Incorporation assay screen with 0.5 mM nsAA supplementation. We began by performing a well-established fluorescence assay with commercially sourced nsAAs **1e** and **4e** (0.5 mM) or no nsAA supplemented to cells transformed to express five previously reported OTSs derived from the *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase (*Mj*TyrRS)/tRNA pair and a green fluorescent protein (GFP) reporter containing an in-frame UAG codon targeted for suppression (**Fig. 3a**). Unsurprisingly, different aminoacyl-tRNA synthetases resulted in varying fold changes in fluorescence normalized to optical density (FL/OD) as a function of the substrate. **a.** FL/OD for cells supplemented nsAA **1e** compared to the no nsAA supplemented control. **b.** FL/OD for cells supplemented nsAA **4e** compared to the no nsAA supplemented control. Strains S27-S31 were used for this study. Sample sizes are n = 3 using biological replicates and data shown are mean ± s.d. Source data are provided as a Source Data file.



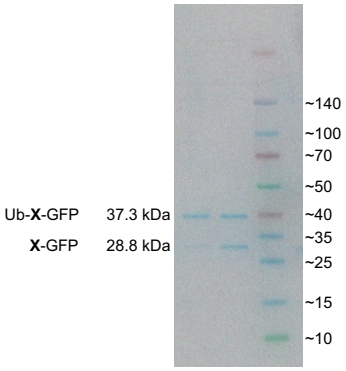
Supplementary Figure 18. MS/MS spectra of ubiquitin-GFP fusion with peptide fragment corresponding to the site-specific incorporation of biosynthesized BipA (4e) from the aldehyde precursor 4b. MS/MS of the peptide LRGG(**4e**)LFVQELASK. Mass spectra data were collected in $n = 1$.



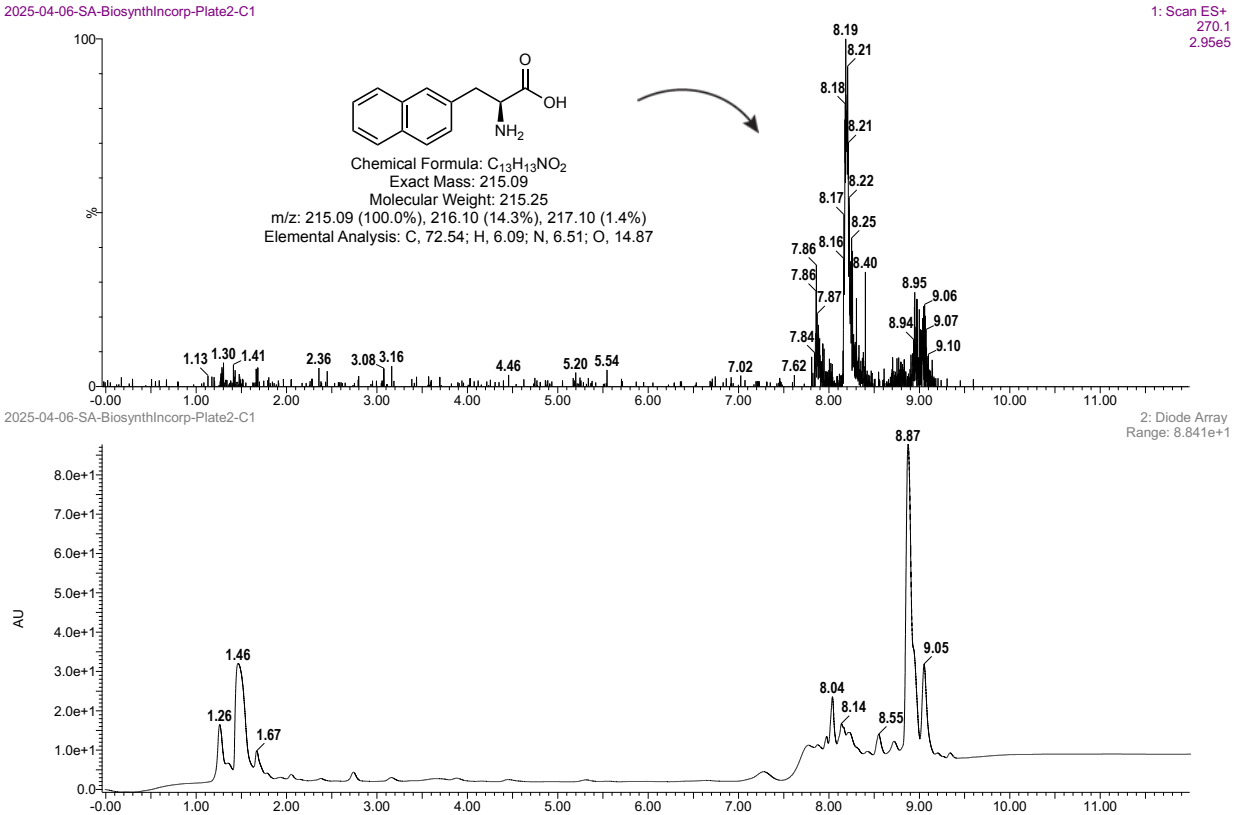
Supplementary Figure 19. MS/MS spectra of ubiquitin-GFP fusion with peptide fragment corresponding to the site-specific incorporation of biosynthesized pAzF (1e) from the carboxylic acid precursor 1a. MS/MS of the peptide LRGG(**1e**)LFVQELASK. Mass spectra data were collected in $n = 1$.



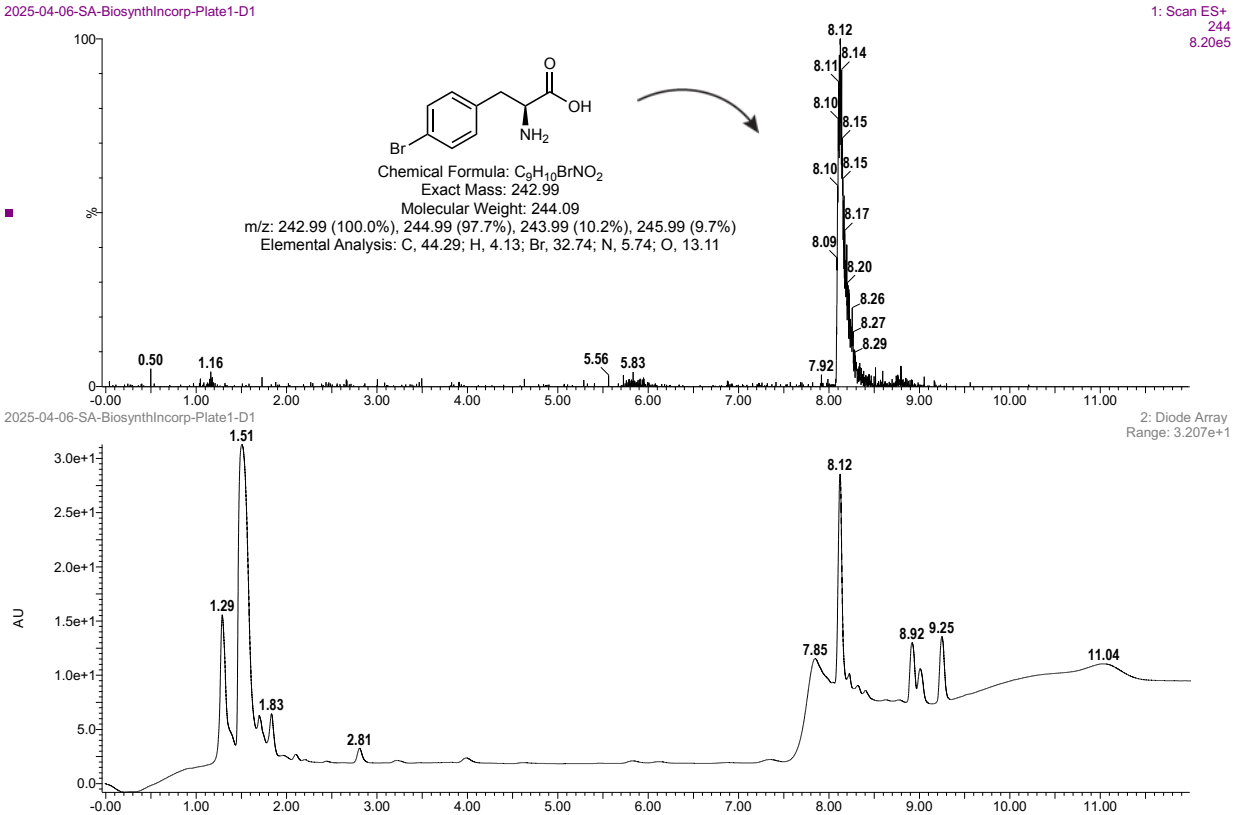
Supplementary Figure 20. Effect of different plasmid construct designs on desired nsAA (product e) titers. With the tested genetic constructs, we varied the ori, the TTA variant, as well as the promoter. Ultimately we observed poor titers with the use of a plasmid containing the high copy number ori, CloDF (1). We observed higher titers when utilizing a vector containing a ColE1 ori, a Tet or J21300 strong promoter, and one of the two TTAs screened, dependent on the chemistry. Sample sizes are n = 3 using biological replicates and data shown are mean ± s.d. Source data are provided as a Source Data file.



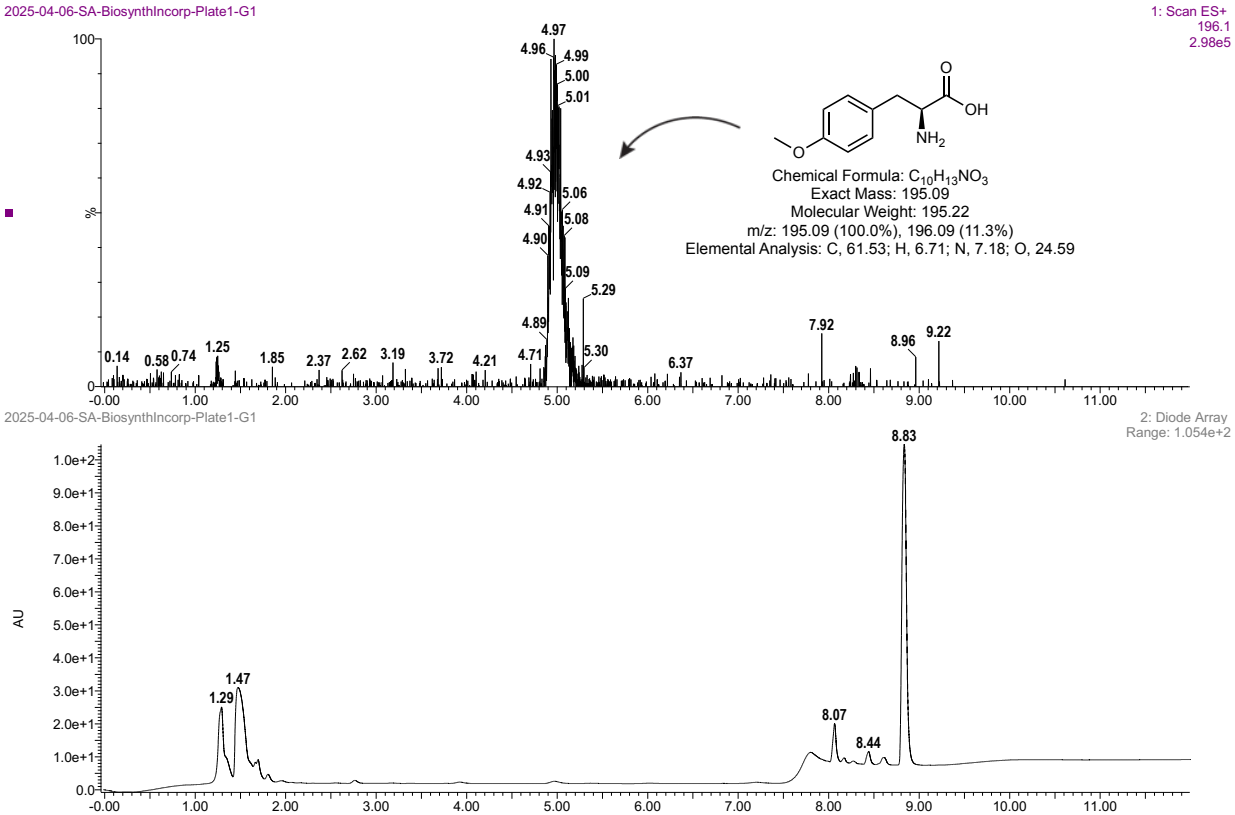
Supplementary Figure 21. Expected size of the reporter protein containing desired biosynthesized nsAA via SDS-PAGE. Lane 1 contains protein with biosynthesized 4-acetyl-phenylalanine (2e) and lane 2 contains protein with biosynthesized 4-cyano-phenylalanine (3e), with the Spectra Multicolor Broad Range Protein Ladder in lane 3. In both lanes, we observe a major band at 37.3 kDa, reflective of the fully translated Ub-X-GFP, where X is the position of the nsAA. Additionally, we observe a band at 28.8 kDa, indicative of Ub cleavage, a phenomenon that we have observed occurs in *E. coli*.



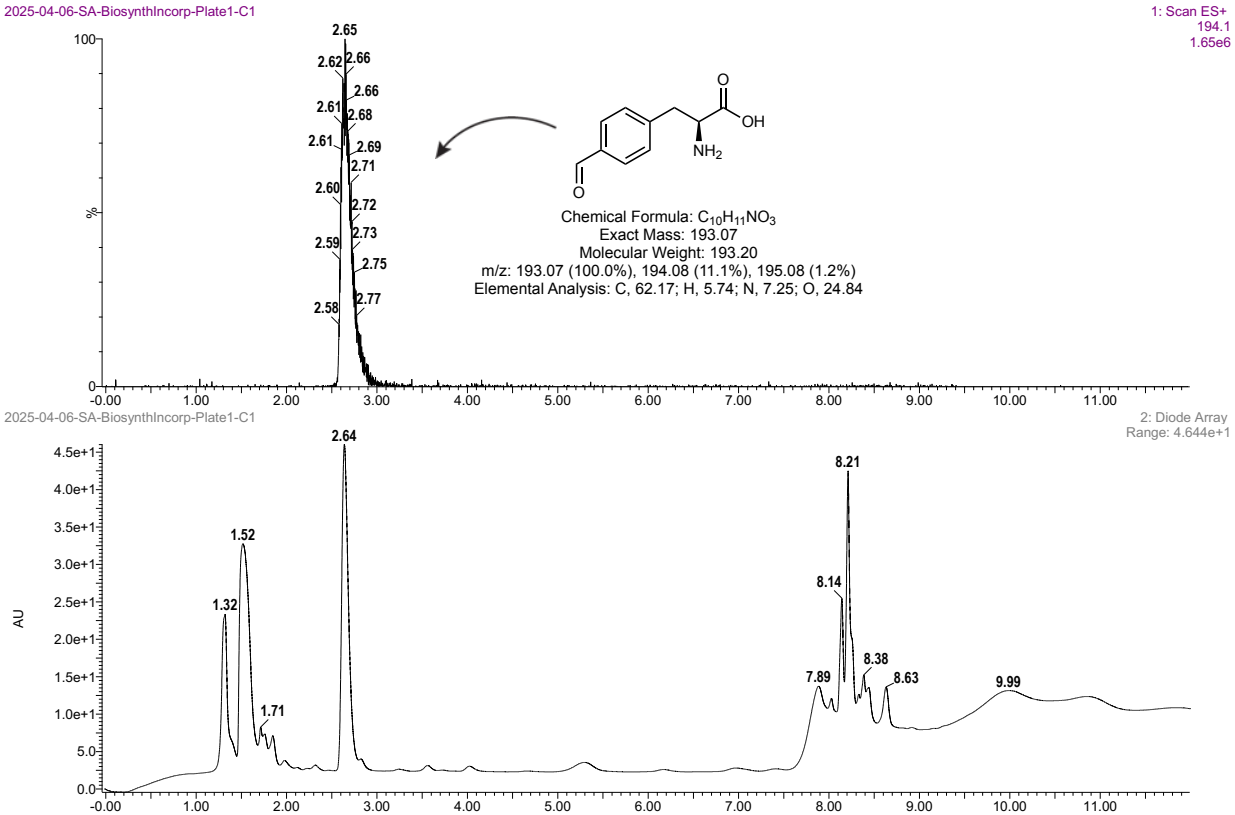
Supplementary Figure 22. In vivo biosynthesis of the nsAA 5e from aldehyde precursor (5b), as confirmed by MS. Mass spectra data were collected in $n = 1$.



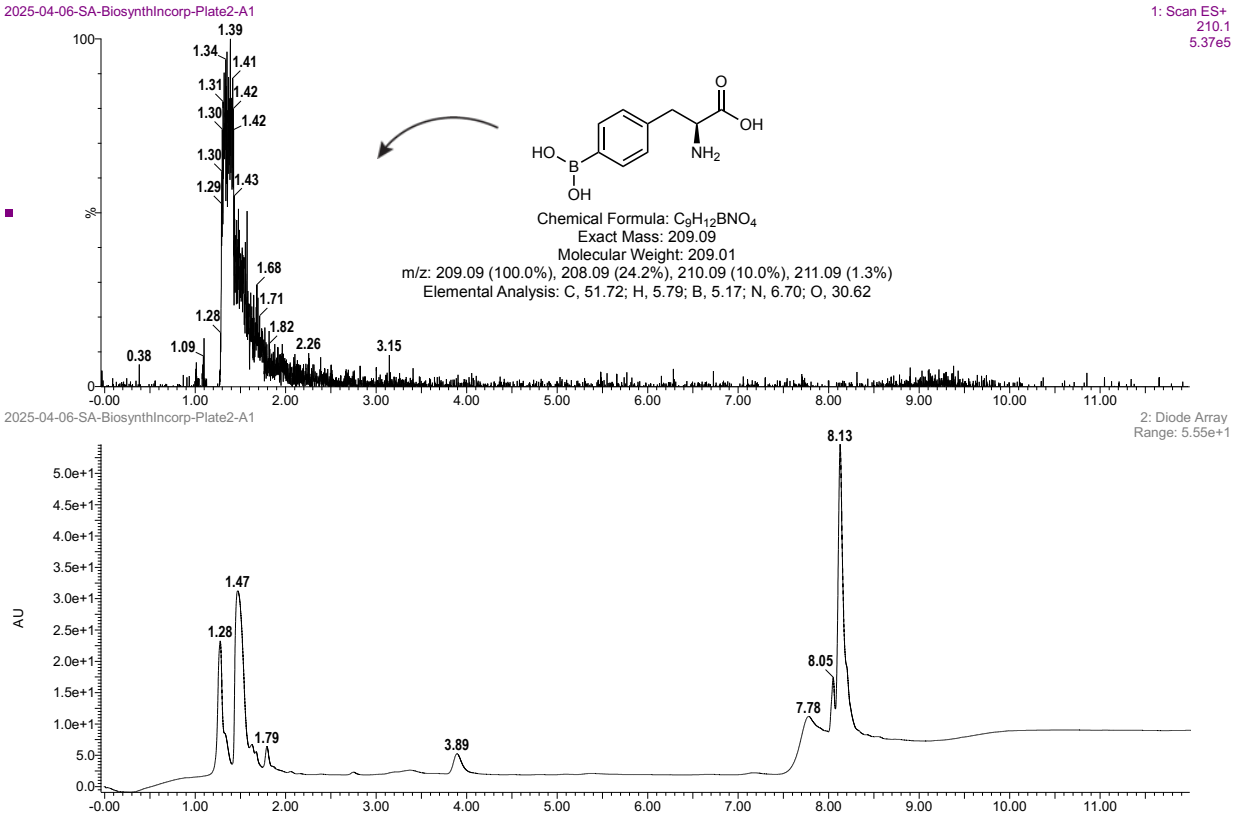
Supplementary Figure 23. In vivo biosynthesis of the nsAA 6e from aldehyde precursor (6b), as confirmed by MS. Mass spectra data were collected in $n = 1$.



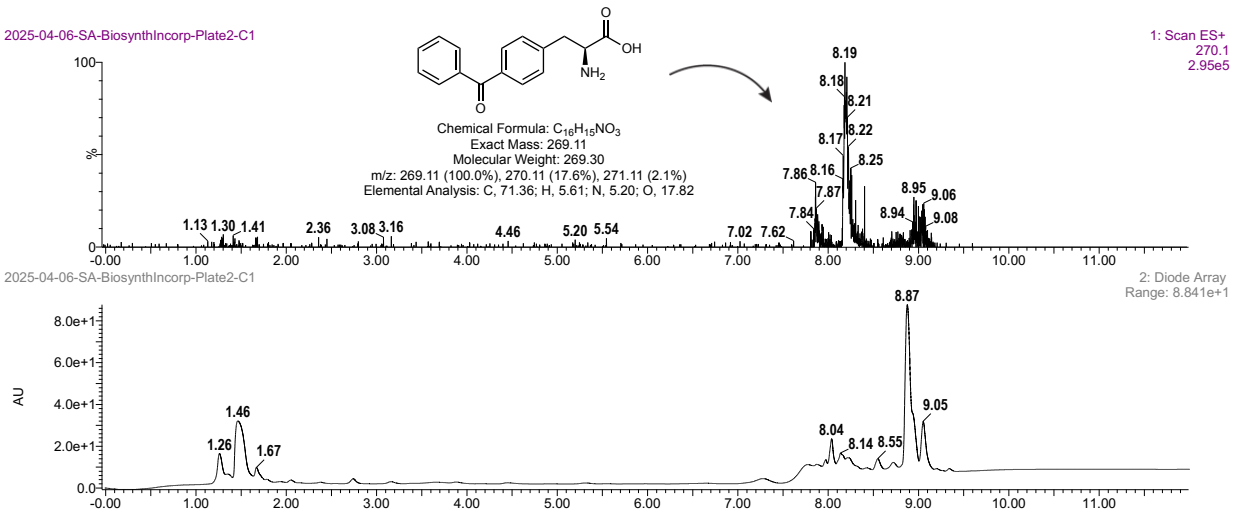
Supplementary Figure 24. In vivo biosynthesis of the nsAA 7e from aldehyde precursor (7b), as confirmed by MS. Mass spectra data were collected in $n = 1$.



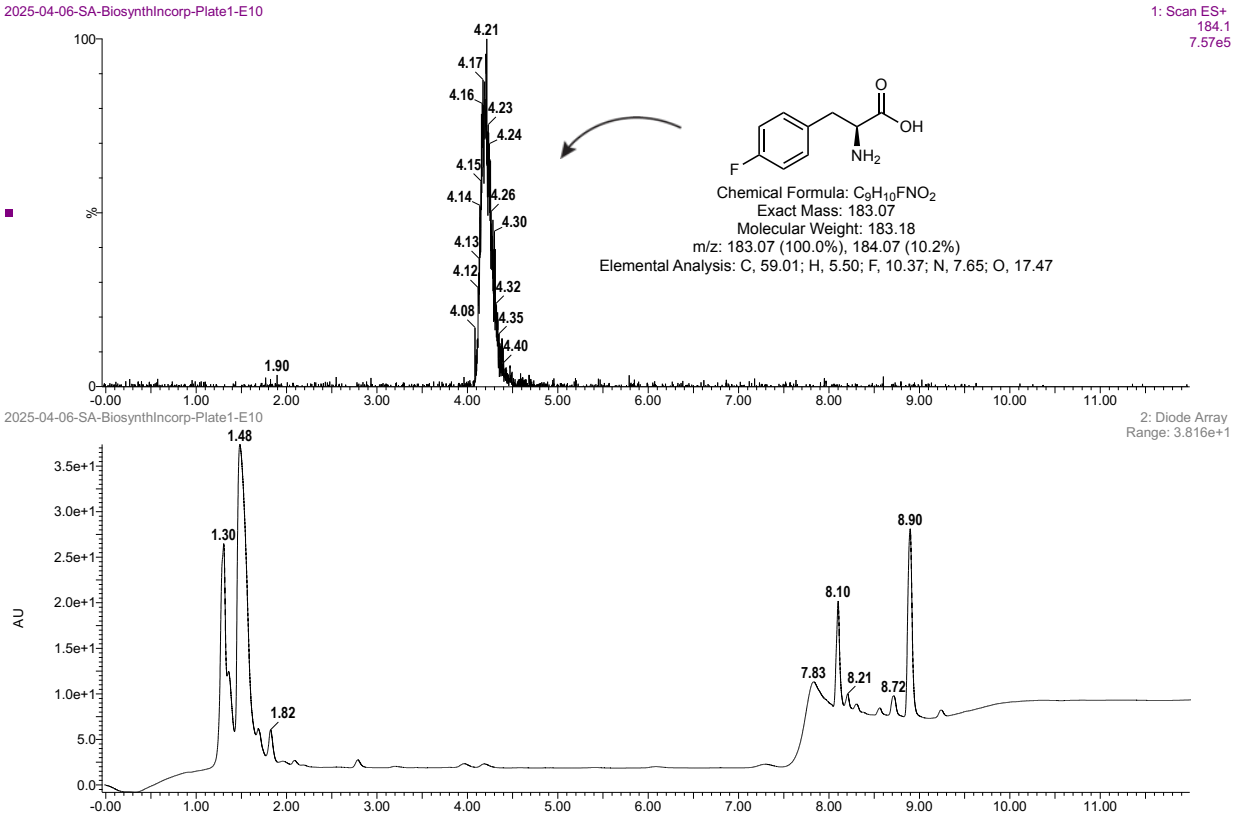
Supplementary Figure 25. In vivo biosynthesis of the nsAA 8e from aldehyde precursor (8b), as confirmed by MS. Mass spectra data were collected in $n = 1$.



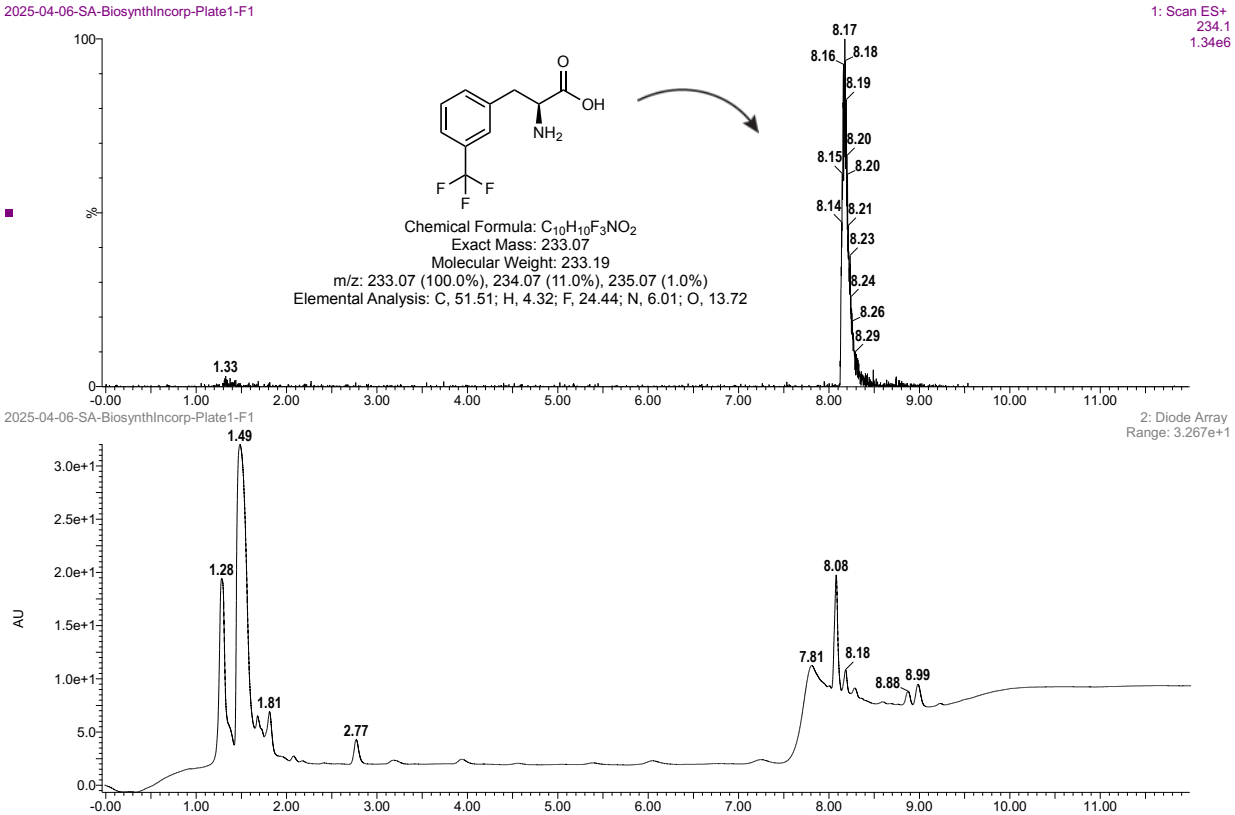
Supplementary Figure 26. In vivo biosynthesis of the nsAA 9e from aldehyde precursor (9b), as confirmed by MS. Mass spectra data were collected in $n = 1$.



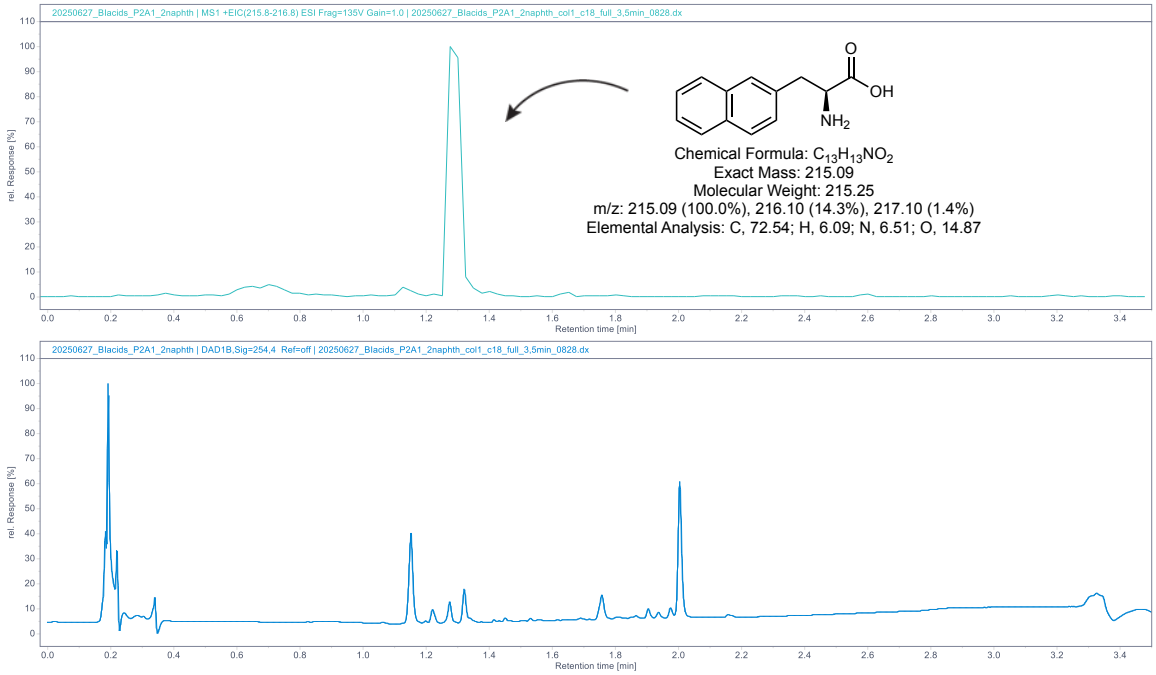
Supplementary Figure 27. In vivo biosynthesis of the nsAA 10e from aldehyde precursor (10b), as confirmed by MS. Mass spectra data were collected in $n = 1$.



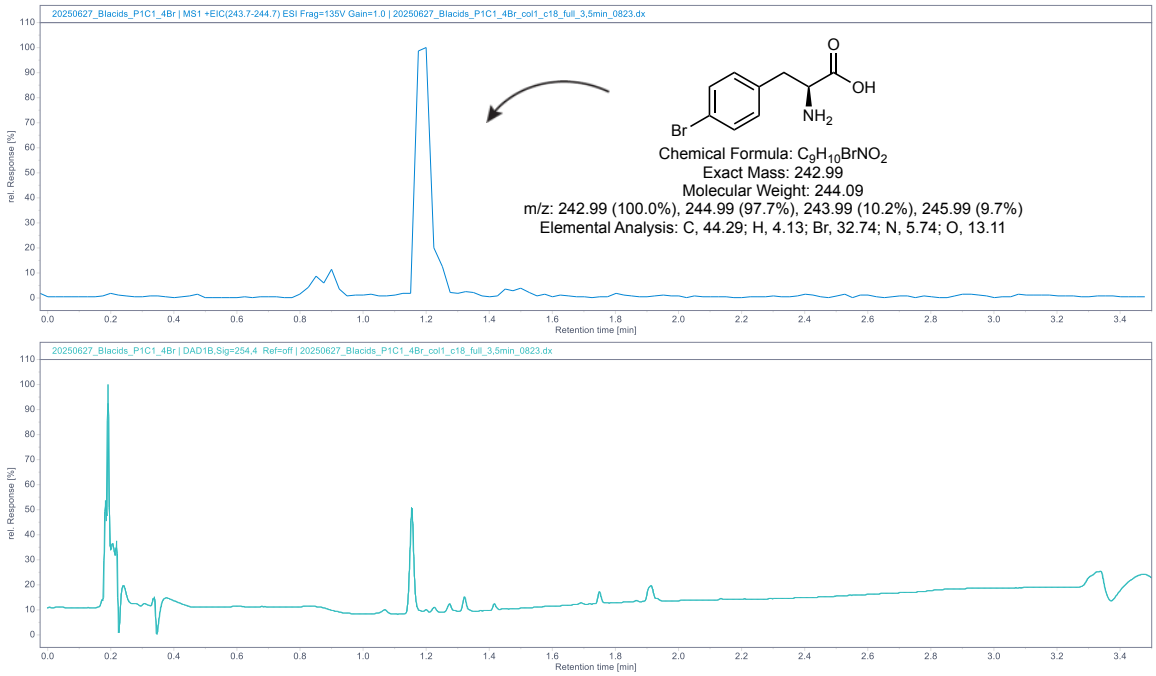
Supplementary Figure 28. In vivo biosynthesis of the nsAA 11e from aldehyde precursor (11b), as confirmed by MS. Mass spectra data were collected in $n = 1$.



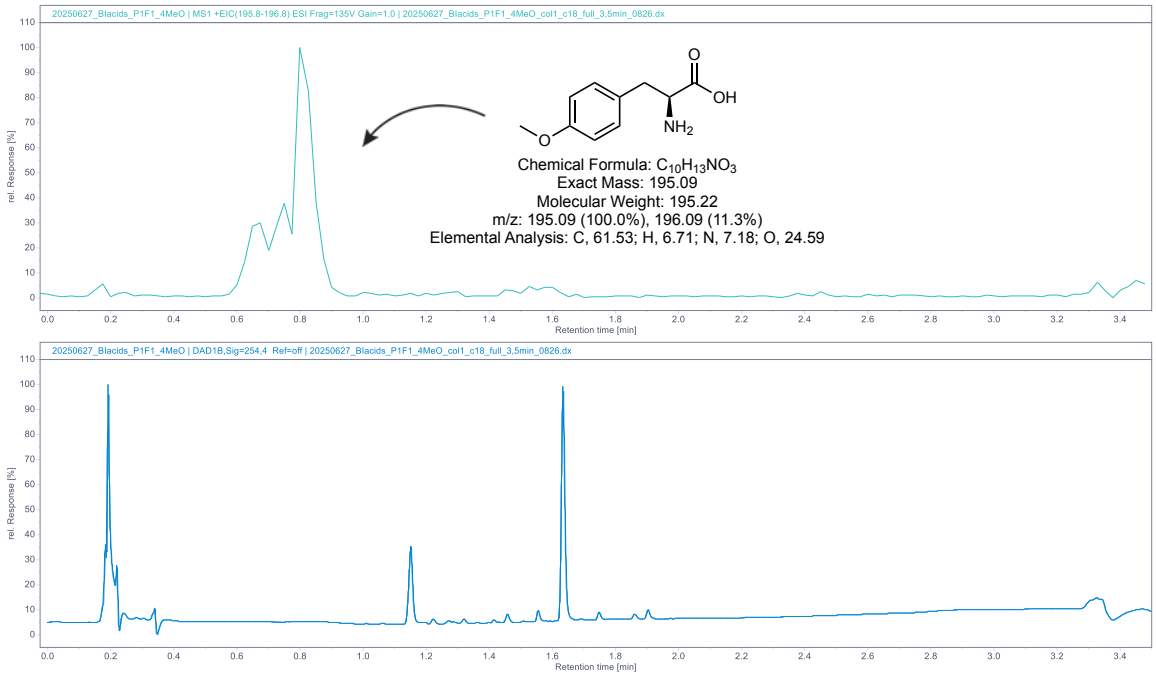
Supplementary Figure 29. In vivo biosynthesis of the nsAA 12e from aldehyde precursor (12b), as confirmed by MS. Mass spectra data were collected in $n = 1$.



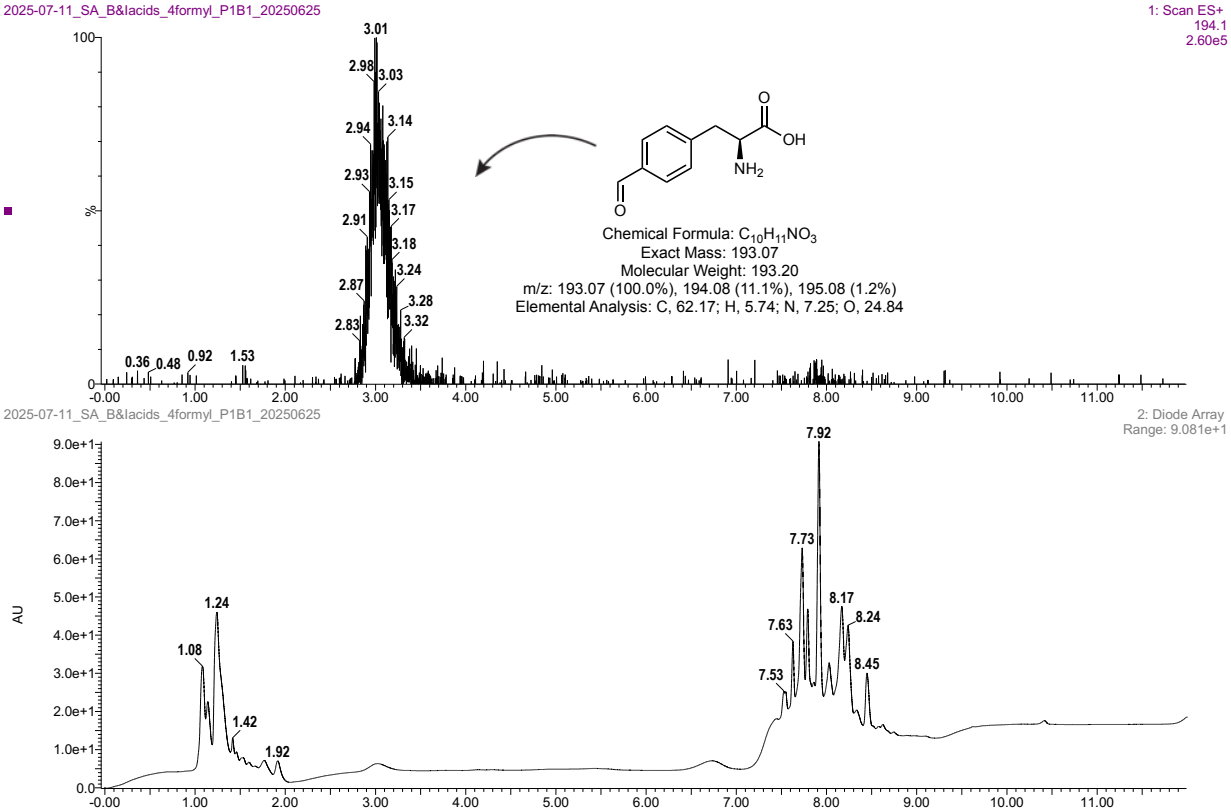
Supplementary Figure 30. In vivo biosynthesis of the nsAA 5e from carboxylic acid precursor (5a), as confirmed by MS. Mass spectra data were collected in $n = 1$.



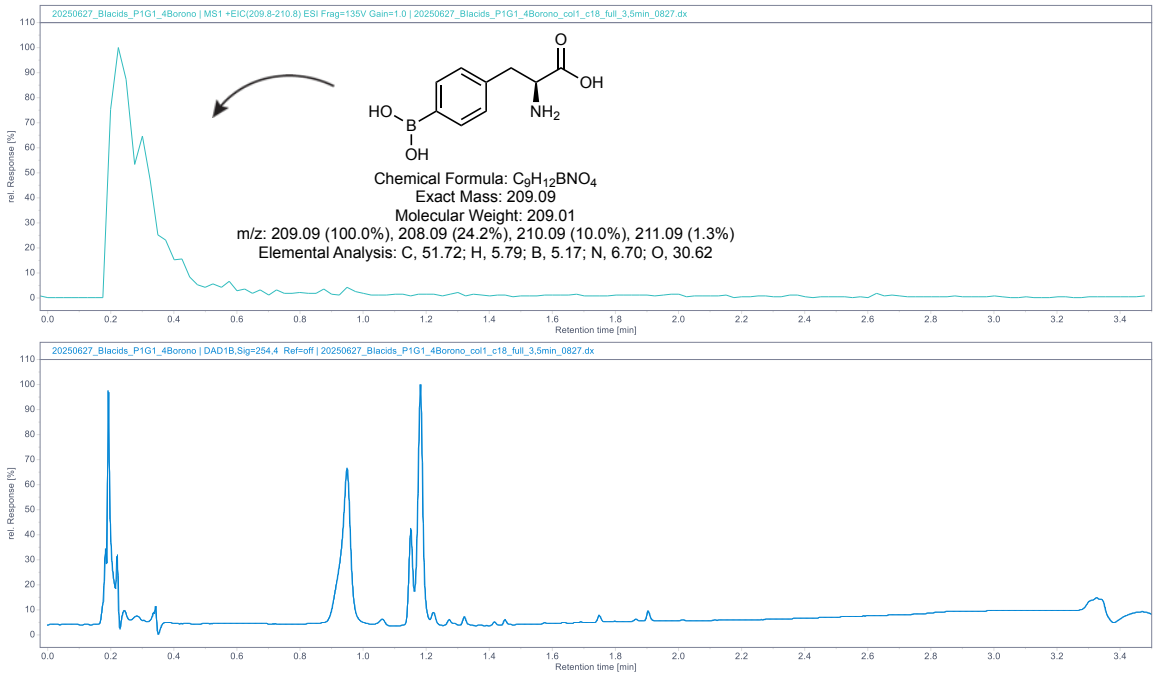
Supplementary Figure 31. In vivo biosynthesis of the nsAA 6e from carboxylic acid precursor (6a), as confirmed by MS. Mass spectra data were collected in $n = 1$.



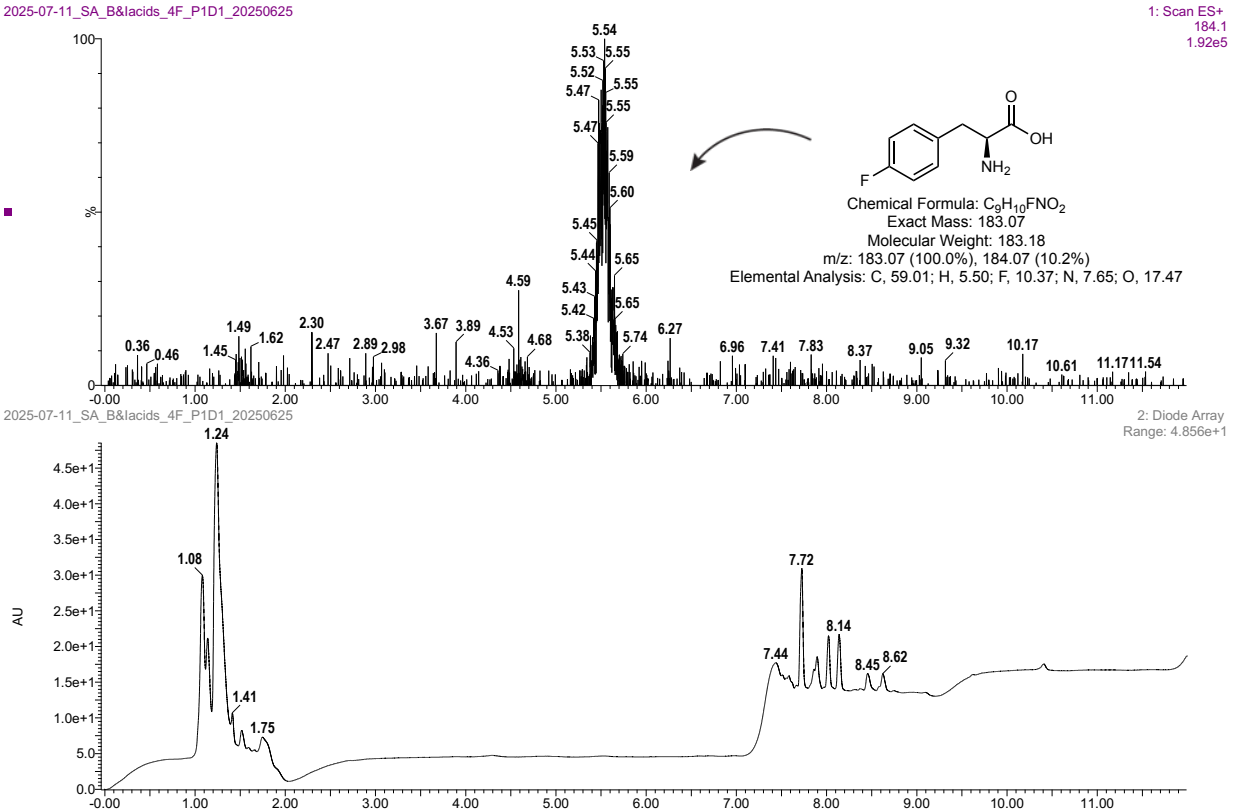
Supplementary Figure 32. In vivo biosynthesis of the nsAA 7e from carboxylic acid precursor (7a), as confirmed by MS. Mass spectra data were collected in $n = 1$.



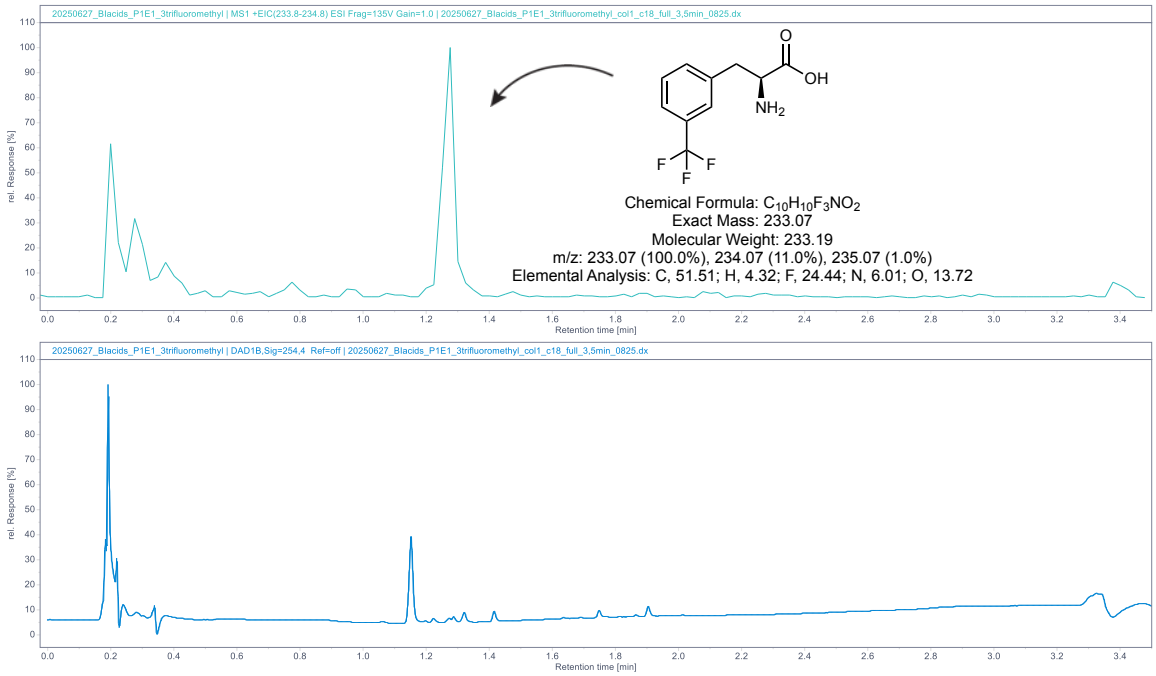
Supplementary Figure 33. In vivo biosynthesis of the nsAA 8e from carboxylic acid precursor (8a), as confirmed by MS. Mass spectra data were collected in $n = 1$.



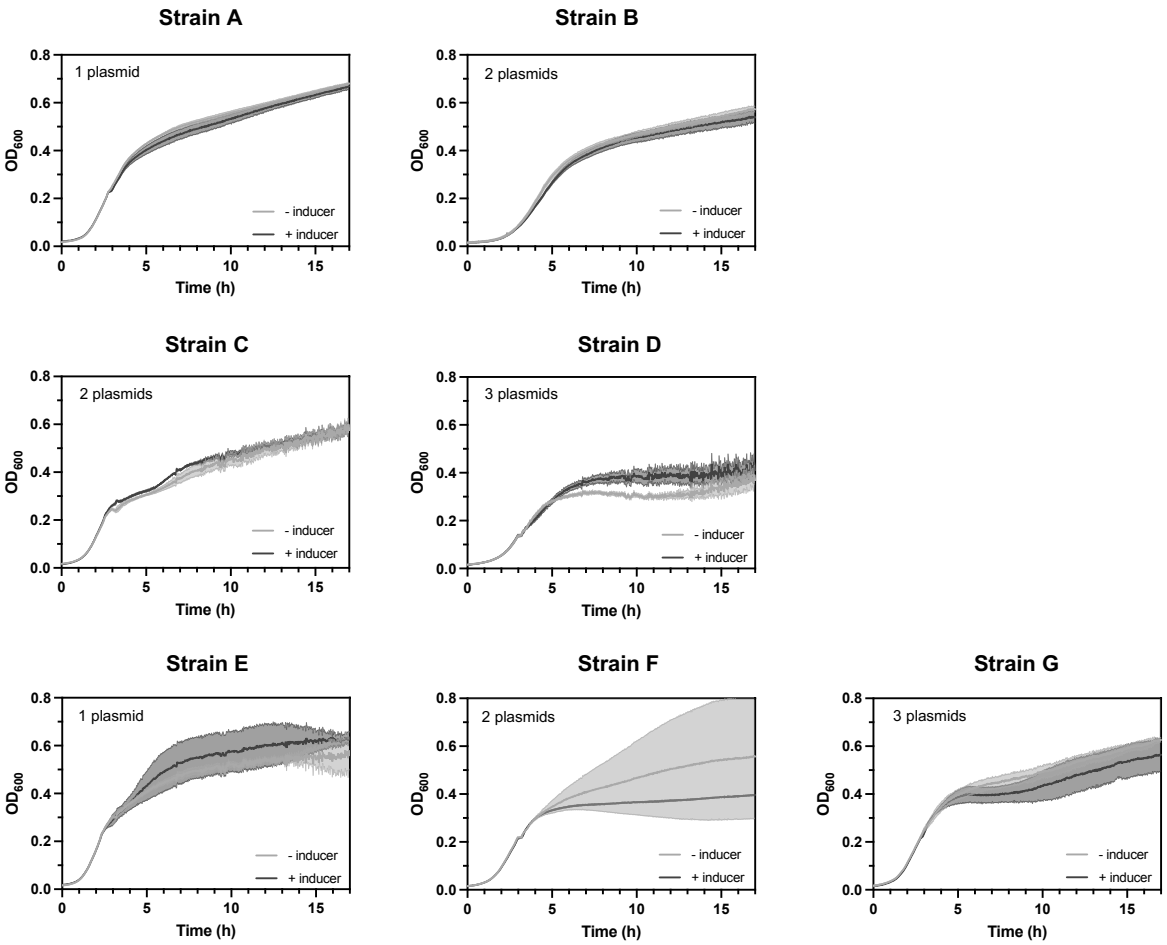
Supplementary Figure 34. In vivo biosynthesis of the nsAA 9e from carboxylic acid precursor (9a), as confirmed by MS. Mass spectra data were collected in $n = 1$.



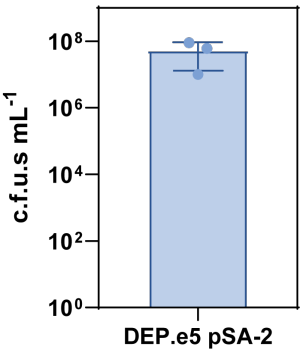
Supplementary Figure 35. In vivo biosynthesis of the nsAA 11e from carboxylic acid precursor (11a), as confirmed by MS. Mass spectra data were collected in $n = 1$.



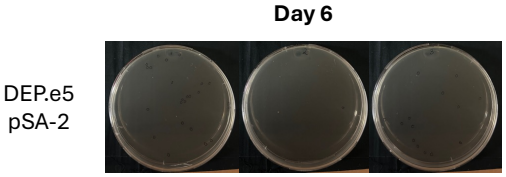
Supplementary Figure 36. In vivo biosynthesis of the nsAA 12e from carboxylic acid precursor (12a), as confirmed by MS. Mass spectra data were collected in $n = 1$.



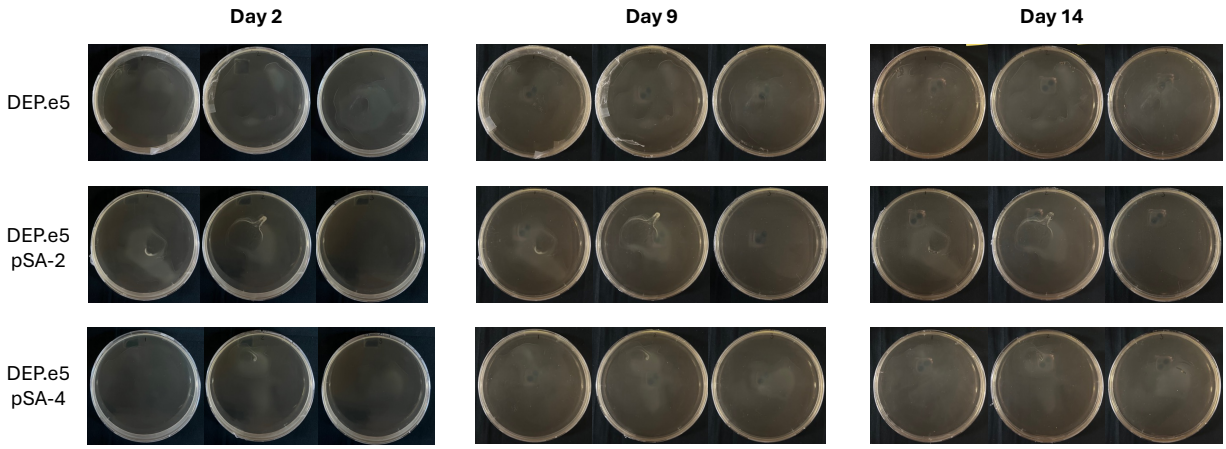
Supplementary Figure 37. Monitoring the growth of the engineered strains containing the genetic constructs for biosynthesis and incorporation of nsAAs. Strains A-G were grown in 200 µL cultures in a 96-well, flat bottom plate for continuous OD_{600} monitoring via a SpectraMax plate reader at 37 °C. At mid-exponential growth phase, inducers appropriate for each strain were added in triplicate conditions and monitored in comparison to conditions that were not provided inducer. The strains used in this study were as follows: A = S52, B = S36, C =S54, D = S55, E = S53, F = S44, G = S48 (**Table S1**). Sample sizes are $n = 3$ using biological replicates and data shown are mean ± s.d. Source data are provided as a Source Data file.



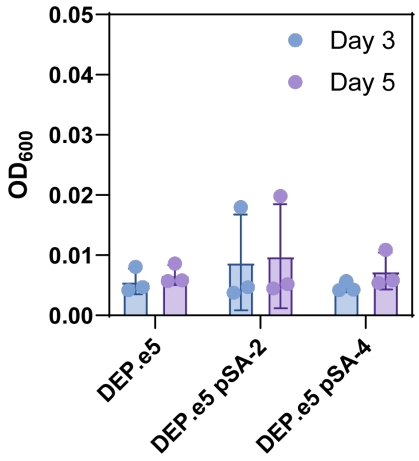
Supplementary Figure 38. C.f.u.s ml^{-1} values of DEP.e5 transformed with pSA-2. Calculated from 10-fold serial dilutions of the samples grown on solid permissive media for 24 h following incubation liquid media containing 500 μM **4b**. Sample sizes are $n = 3$ using biological replicates and data shown are mean ± s.d. Source data are provided as a Source Data file.



Supplementary Figure 39. Monitoring for escape of the synthetic auxotroph transformed with the biosynthetic pathway following provision with 4b. No colonies observed on solid non-permissive media for up to 6 days following incubation liquid media containing 500 μM **4b**. Sample sizes are $n = 3$ using biological replicates.



Supplementary Figure 40. Monitoring for escape of the synthetic auxotroph transformed with the biosynthetic pathway. Cultures (in triplicate) were washed in LB containing 100 mM L-Thr and the entirety of the 3 mL culture volumes were plated across three non-permissive agar plates. Growth was monitored for 14 days at 34 °C. No colonies were observed over the timeframe, and the film observed on some of the plates (sampled after 9 days) was shown to not contain viable cells (**Supplementary Figure 41**). Sample sizes are n = 3 using biological replicates.



Supplementary Figure 41. Monitoring the OD_{600} of the film for growth. No observable growth after inoculation of the film from the nonpermissive plates (shown in **Supplementary Figure 40**) in nonpermissive liquid media and incubation at 34 °C. OD_{600} was monitored after 3 and 5 days. Sample sizes are n = 3 using biological replicates and data shown are mean ± s.d. Source data are provided as a Source Data file.

III. Supplementary References

1. Kunjapur, A. M., Tarasova, Y. & Prather, K. L. J. Synthesis and accumulation of aromatic aldehydes in an engineered strain of Escherichia coli. *J. Am. Chem. Soc.* **136**, 11644–11654 (2014).
2. Dickey, R. M., Jones, M. A., Butler, N. D., Govil, I. & Kunjapur, A. M. Genome engineering allows selective conversions of terephthalaldehyde to multiple valorized products in bacterial cells. *AIChE J.* **69**, e18230 (2023).
3. Butler, N. D., Anderson, S. R., Dickey, R. M., Nain, P. & Kunjapur, A. M. Combinatorial gene inactivation of aldehyde dehydrogenases mitigates aldehyde oxidation catalyzed by *E. coli* resting cells. *Metab. Eng.* **77**, 294–305 (2023).
4. Kunjapur, A. M. *et al.* Synthetic auxotrophy remains stable after continuous evolution and in coculture with mammalian cells. *Sci. Adv.* **7**, eabf5851 (2021).
5. Anderson, S. R. *et al.* A One-Pot Biocatalytic Cascade to Access Diverse l-Phenylalanine Derivatives from Aldehydes or Carboxylic Acids. *ChemCatChem* **n/a**, e202401979.
6. Butler, N. D., Sen, S., Brown, L. B., Lin, M. & Kunjapur, A. M. A platform for distributed production of synthetic nitrated proteins in live bacteria. *Nat. Chem. Biol.* **19**, 911–920 (2023).
7. Jones, M. A. *et al.* Discovery of l-threonine transaldolases for enhanced biosynthesis of beta-hydroxylated amino acids. *Commun. Biol.* **6**, 929 (2023).
8. Gopal, M. R. *et al.* Reductive Enzyme Cascades for Valorization of Polyethylene Terephthalate Deconstruction Products. *ACS Catal.* **13**, 4778–4789 (2023).
9. Kunjapur, A. M. *et al.* Engineering posttranslational proofreading to discriminate nonstandard amino acids. *Proc. Natl. Acad. Sci.* **115**, 619–624 (2018).
10. Speight, L. C. *et al.* Efficient Synthesis and In Vivo Incorporation of Acridon-2-ylalanine, a Fluorescent Amino Acid for Lifetime and Förster Resonance Energy Transfer/Luminescence Resonance Energy Transfer Studies. *J. Am. Chem. Soc.* **135**, 18806–18814 (2013).
11. Chin, J. W., Martin, A. B., King, D. S., Wang, L. & Schultz, P. G. Addition of a photocrosslinking amino acid to the genetic code of Escherichia coli. *Proc. Natl. Acad. Sci.* **99**, 11020–11024 (2002).
12. Alfonta, L., Zhang, Z., Uryu, S., Loo, J. A. & Schultz, P. G. Site-Specific Incorporation of a Redox-Active Amino Acid into Proteins. *J. Am. Chem. Soc.* **125**, 14662–14663 (2003).
13. Seitchik, J. L. *et al.* Genetically Encoded Tetrazine Amino Acid Directs Rapid Site-Specific in Vivo Bioorthogonal Ligation with trans-Cyclooctenes. *J. Am. Chem. Soc.* **134**, 2898–2901 (2012).
14. Miyake-Stoner, S. J. *et al.* Generating Permissive Site-Specific Unnatural Aminoacyl-tRNA Synthetases. *Biochemistry* **49**, 1667–1677 (2010).
15. Cooley, R. B., Karplus, P. A. & Mehl, R. A. Gleaning Unexpected Fruits from Hard-Won Synthetases: Probing Principles of Permissivity in Non-canonical Amino Acid–tRNA Synthetases. *ChemBioChem* **15**, 1810–1819 (2014).
16. Amiram, M. *et al.* Evolution of translation machinery in recoded bacteria enables multi-site incorporation of nonstandard amino acids. *Nat. Biotechnol.* **33**, 1272–1279 (2015).
17. Chin, J. W. *et al.* Addition of *p* -Azido- L -phenylalanine to the Genetic Code of *Escherichia c oli*. *J. Am. Chem. Soc.* **124**, 9026–9027 (2002).
18. Xie, J., Liu, W. & Schultz, P. G. A genetically encoded bidentate, metal-binding amino acid. *Angew. Chem. Int. Ed Engl.* **46**, 9239–9242 (2007).
19. Minnihan, E. C., Young, D. D., Schultz, P. G. & Stubbe, J. Incorporation of Fluorotyrosines into Ribonucleotide Reductase Using an Evolved, Polyspecific Aminoacyl-tRNA Synthetase. *J. Am. Chem. Soc.* **133**, 15942–15945 (2011).
20. Young, D. D. *et al.* An Evolved Aminoacyl-tRNA Synthetase with Atypical Polysubstrate Specificity,. *Biochemistry* **50**, 1894–1900 (2011).
21. Blizzard, R. J. *et al.* Ideal Bioorthogonal Reactions Using A Site-Specifically Encoded Tetrazine Amino Acid. *J. Am. Chem. Soc.* **137**, 10044–10047 (2015).
22. Wang, J., Xie, J. & Schultz, P. G. A Genetically Encoded Fluorescent Amino Acid. *J. Am. Chem. Soc.* **128**, 8738–8739 (2006).
23. Wang, L., Brock, A. & Schultz, P. G. Adding l-3-(2-Naphthyl)alanine to the Genetic Code of E. coli. *J. Am. Chem. Soc.* **124**, 1836–1837 (2002).
24. Wang, L., Zhang, Z., Brock, A. & Schultz, P. G. Addition of the keto functional group to the genetic code of Escherichia coli. *Proc. Natl. Acad. Sci.* **100**, 56–61 (2003).

25. Mehl, R. A. *et al.* Generation of a Bacterium with a 21 Amino Acid Genetic Code. *J. Am. Chem. Soc.* **125**, 935–939 (2003).
26. Tsao, M.-L., Summerer, D., Ryu, Y. & Schultz, P. G. The Genetic Incorporation of a Distance Probe into Proteins in Escherichia coli. *J. Am. Chem. Soc.* **128**, 4572–4573 (2006).
27. Zimmerman, E. S. & Thanos, C. Non-natural amino acid tRNA synthetases for pyridyl tetrazine. (2019).
28. Wang, Y.-S., Fang, X., Wallace, A. L., Wu, B. & Liu, W. R. A Rationally Designed Pyrrolysyl-tRNA Synthetase Mutant Has a Broad Substrate Spectrum. *J. Am. Chem. Soc.* **134**, 2950–2953 (2012).
29. Beránek, V., Willis, J. C. W. & Chin, J. W. An Evolved Methanomethylophilus alvus Pyrrolysyl-tRNA Synthetase/tRNA Pair Is Highly Active and Orthogonal in Mammalian Cells. *Biochemistry* **58**, 387–390 (2019).
30. Ai, H., Shen, W., Sagi, A., Chen, P. R. & Schultz, P. G. Probing Protein–Protein Interactions with a Genetically Encoded Photo-crosslinking Amino Acid. *ChemBioChem* **12**, 1854–1857 (2011).