Supporting Information

Nitroalkanes as Versatile Nucleophiles for Enzymatic Synthesis of Noncanonical Amino Acids

David K. Romney,*,†,‡ Nicholas S. Sarai,† Frances H. Arnold*,†

†Division of Chemistry and Chemical Engineering 210-41, California Institute of Technology, 1200 East California Boulevard, Pasadena, California 91125, United States

‡Current address: Aralez Bio, 101 West Ferndale Avenue, Sunnyvale, CA, 94085

Corresponding authors

*dromney@aralezbio.com

*frances@cheme.caltech.edu

Table of Contents

1. Additional figures

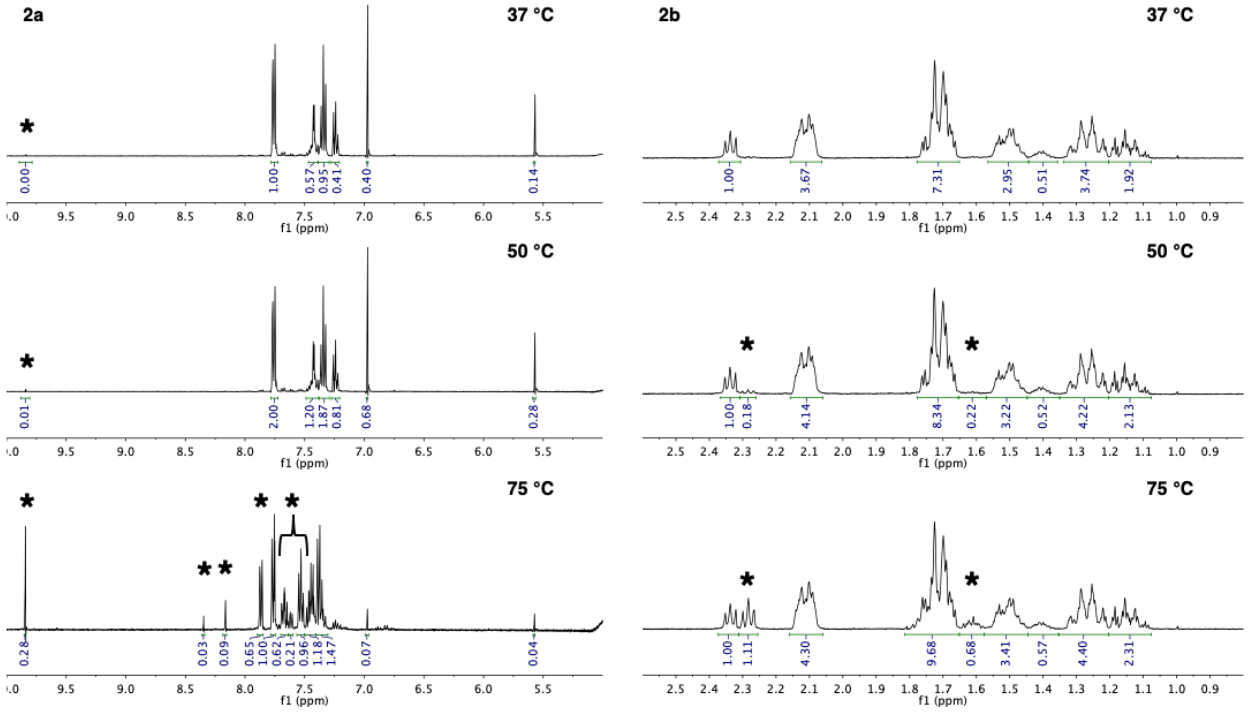


Figure S1. ^{1}H NMR spectra of compounds **2a** and **2b** in aqueous potassium phosphate buffer (pH 8) after 12 hours at the indicated temperature. *Signals associated with decomposition products.

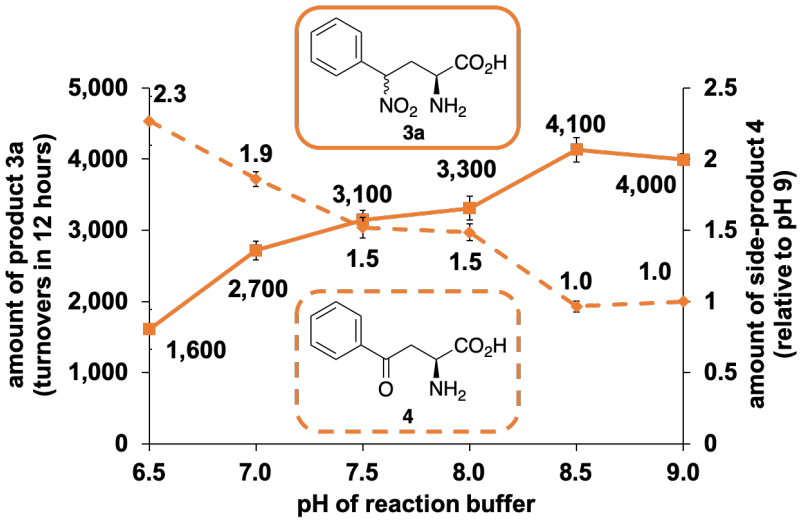


Figure S2. Formation of product **3a** at different pH (solid line) and accompanying formation of hydrolysis byproduct **4** (dashed line). See Section 2.9 for experimental details.

2. Procedures for enzyme expression and characterization

2.1 Cloning, expression, and purification of TrpB variants

The gene encoding *Pf*(β-Me) was previously cloned into pET22(b)+ with a *C*-terminal His-tag.[1] Protein expression of the variants was carried out in *Escherichia coli* BL21 *E. cloni* Express cells (Lucigen) by inoculating 5 mL of Lysogeny Broth containing 100 μg/mL carbenicillin (LB_{carb}) with a single colony and incubating this pre-culture overnight at 37 °C and 230 rpm. The overnight cultures were used to inoculate 500 mL of Terrific Broth containing 100 μg/mL carbenicillin (TB_{carb}). The expression cultures were shaken at 37 °C and 230 rpm for ~3 h, at which point the OD_{600} was 0.6–0.8. The cultures were chilled on ice for >30 min and then induced by the addition of 1 M aq. isopropyl β-D-thiogalactopyranoside (IPTG, 500 μL, final concentration of 1 mM). Expression of the variants took place at 230 rpm and 20 °C for another 20 h. The cultures were centrifuged at 5,000×*g* and 4 °C for 5 minutes. The cell pellets were decanted, then frozen and stored at –20 °C until further use.

For protein purification, cells were thawed, then re-suspended in potassium phosphate buffer (25 mM, pH 8) that contained 20 mM imidazole, 100 mM NaCl, 400 μM PLP, 1 mg/mL of hen egg white lysozyme (HEWL, Sigma Aldrich), and 0.1 mg/mL of bovine pancreas DNase I. BugBuster (Novagen) was added, then the mixture was vortexed to suspend the pellet. The suspension was shaken at 37 °C and 230 rpm for 15 min, then centrifuged at 5000×*g* and 4 °C for 10 minutes. Without decanting, the cell lysate was immersed in a water bath at 75 °C. After 30 minutes, the suspension was centrifuged (15,000×*g* and 4 °C for 15 minutes). The supernatant was purified using a 1-mL HisTrap HP column with an AKTA purifier FPLC system (GE Healthcare) and a linear gradient from buffer A (25 mM potassium phosphate, 20 mM imidazole, 100 mM NaCl, pH 8) to buffer B (25 mM potassium phosphate, 500 mM imidazole, 100 mM NaCl, pH 8) over 10 volumes. Proteins eluted at approximately 140 mM imidazole. Purified proteins were dialyzed into potassium phosphate buffer (50 mM, pH 8), then flash-frozen in liquid N_2 and stored at –80 °C until further use. Protein concentrations were determined *via* the Bradford assay (Bio-Rad).

2.2 Construction of site-saturation mutagenesis libraries

PCR was conducted using Phusion polymerase (New England Biolabs) according to the standard protocol. For the given site of mutagenesis, three primers were designed containing codons NDT (encoding for Ile, Asn, Ser, Gly, Asp, Val, Arg, His, Leu, Phe, Tyr, and Cys), VHG (encoding for Met, Thr, Lys, Glu, Ala, Val, Gln, Pro, and Leu), and TGG (Trp), respectively, thereby including all 20 natural amino acids. These three primers were mixed in a ratio 12:9:1 according to the previously described protocol.[2] Then, the plasmid was constructed by site-directed mutagenesis by overlap extension (SOE) PCR using a plasmid that contained the parent gene in the pET22(b)+ vector as template. The linear plasmid was digested with *Dpn*I (1 μL per 50 μL PCR reaction, incubated at 37 °C for 2 hours), purified by preparative agarose gel, then cyclized *via* the Gibson method.[3]

2.3 Construction of random recombination libraries

These libraries were constructed in an analogous manner to the site-saturation libraries, using primers that coded for both the native residue and the mutation. The mutant genes were first constructed as fragments, which were purified by preparative agarose gel, then assembled into a contiguous gene using primers that corresponded to the

*Nde*I and *Xho*I sites of the pET22(b)+ vector (GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATG and GCCGGATCTCAGTGGTGGTGGTGGTGGTGCTCGAG, respectively). After a final purification by agarose gel, the assembled gene was cloned into an empty pET22(b)+ vector between restriction sites *Nde*I and *Xho*I using the Gibson method.[3]

2.4 Transformation of BL21 *E. coli* cells

In preparation, SOC medium, 50 µL aliquots of electrocompetent BL21 *E. coli* cells, and electroporation cuvettes were chilled in ice. The plasmid (1 µL) was added to the cells, which were then transferred to a sterile electroporation cuvette. An electric potential was applied with a Gene Pulser Xcell (2.5 kV, 25 µF, 200 Ω). Then, SOC medium (750 µL) was immediately added and the cuvette was shaken at 37 °C and 230 rpm. After 45 min, aliquots of cell suspension were plated onto LB_{amp} agar plates. The plates were incubated overnight at 37 °C, then stored at 4 °C until further use.

2.5 Library expression and screening

BL21 *E. cloni* Express cells carrying parent and variant plasmids were grown in 96-well deep-well plates (300 µL/well TB_{amp}) at 37 °C and 80% humidity. After shaking at 250 rpm overnight, 20 µL of the overnight cultures were transferred to new deep-well plates containing 630 µL/well TB_{amp}, which were allowed to grow at 37 °C and 80% humidity. After shaking at 250 rpm for 3 h, the plates were chilled on ice for 30 min, then induced by the addition of IPTG in TB_{amp} (1 mM final concentration). The cultures were shaken at 250 rpm and 20 °C. After 20 hours, the cultures were centrifuged at 4,000×*g* for 10 min. The cell pellets were frozen at –30 °C for a minimum of 2 hours. For screening, cells were thawed at room temperature and then lysed by the addition of 400 µL/well of potassium phosphate buffer (50 mM, pH 8.0), with 1 mg/mL HEWL, 0.1 mg/mL DNase I, 40 µM PLP, and 2 mM $MgCl_2$. The plates were incubated for 1 h at 37 °C, then transferred to a water bath equilibrated to 75 °C. After 30 min, the plates were chilled in ice, then centrifuged at 5,000×*g* and 4°C for 20 min.

The reactions were performed in 96-well deep-well plates. In general, each well was charged with the nitroalkane substrate as a solution in DMSO (10 µL/well). Then, a solution of serine in potassium phosphate buffer (50 mM, pH 8.0, 140 µL) was added. Finally, the enzymes were added as heat-treated lysate (50 µL), such that the total volume in the wells was 200 µL. All libraries were screened with 10 µmol of nitroalkane and 10 µmol of serine.

The plates were sealed with Teflon sealing mats, then immersed in a water bath equilibrated to 50 °C. After ~12 hours, the plates were chilled in ice and briefly centrifuged (5,000×*g*, 5 min) to settle the reaction contents to the bottom of the wells. The reactions were then processed depending on the assay being used.

For analysis using a UV-visible light plate reader, each well was charged with 500 µL of ethyl acetate and 900 µL of aq. 1 M HCl. The plates were again sealed with Teflon sealing mats, then shaken vigorously to dissolve all precipitates and partition the product and substrate between the aqueous and organic phases, respectively. The plates were again centrifuged (5,000×*g*, 2 min), then 200 µL of the aqueous phase were transferred to 96-well UV-vis assay plates. The activity of each well was determined by measuring the absorption at 260 nm.

For analysis by UHPLC/MS, the reactions were diluted with 800 μL of 1:1 CH_3CN/H_2O that contained 1-naphthylamine (0.1 mg/mL). The plates were centrifuged at 5,000×*g* and 20 °C for 15 minutes, then the supernatants were transferred to a fresh assay plate. Each well was analyzed with a C-18 silica column (4.6 × 50 mm) using acetonitrile/water (0.1% acetic acid by volume): 5% to 95% acetonitrile over 2 minutes, 95% for 1 min; 1 mL/min. The activity was approximated by dividing the integration of product ion with that of 1-naphthylamine.

2.6 Small-scale reactions with heat-treated lysate (data in Figure 2)

The enzyme was expressed as a 5-mL culture in TB_{amp} according to the procedure in Section 2.1. The cell pellet was suspended in 400 μL potassium phosphate buffer (50 mM, pH 8) that contained 1 mg/mL HEWL, 0.1 mg/mL DNase I, and 200 μM PLP. BugBuster was added, then the suspension was shaken at 37 °C and 230 rpm. After 15 minutes, the suspension was chilled in ice, then centrifuged at 5,000×*g* and 4 °C (10 minutes). The supernatant was transferred to a 1.5-mL Eppendorf tube, then heated to 75 °C. After 30 minutes, the suspensions were chilled in ice, then centrifuged at 20,000×*g* and 4 °C (10 minutes). The concentration of enzyme in the lysate was estimated using the Bradford assay, and then adjusted to 100 μM.

A 2-mL HPLC vial was charged with nitroalkane as a 1 M solution in DMSO (10 μL, 10 μmol nitroalkane). Next, Ser (10 μmol) was added as a solution in 170 μL of potassium phosphate buffer (50 mM). Finally, 20 μL of heat-treated lysate was added (2×10^{-3} μmol), then the reaction was heated to 50 °C. After 12 hours, the reaction was chilled in ice, then diluted with 800 μL of 1:1 CH_3CN/H_2O. The reaction mixture was centrifuged at 20,000×*g* and 4 °C (10 minutes), then the supernatant was analyzed by HPLC with a C-18 silica column (1.8 μm, 2.1 × 50 mm) using acetonitrile/water (0.1% acetic acid by volume): 5% to 95% acetonitrile over 4 min; 1 mL/min. The amount of product was calculated from the integration of the product signal in the diode array according to a calibration curve that used chromatographically purified material. Turnover is the theoretical maximum turnovers (5,000) multiplied by the fraction of product formed relative to theoretical maximum (10 μmol). At least three replicates were performed to validate each result.

2.7 Small-scale reactions with purified protein

A 2-mL HPLC vial was charged with nitroalkane as a 1 M solution in DMSO (10 μL, 10 μmol nitroalkane). Next, Ser (10 μmol) was added as a solution in 174 μL of potassium phosphate buffer (50 mM). Finally, 15 μL of purified protein solution was added (concentration was adjusted depending on the desired catalyst loading). The reaction was then heated to 50 °C in a water bath. After 12 hours, the reaction was diluted with 1:1 CH_3CN/H_2O (800 μL) and then centrifuged at 20,000×*g* (10 minutes, 4 °C). Finally, the supernatant was analyzed by UHPLC/MS according to the procedure in Section 2.6.

2.8 Relative initial reaction rates at different pH

Reaction buffers at different pH were prepared by mixing different amounts of dibasic and monobasic potassium phosphate buffers (50 mM). Reactions were run and analyzed using purified protein according to the procedure in Section 2.7 (final enzyme concentration 10 μM), except with a reaction time of ~15 minutes. The rates were

expressed relative to the rate at pH 6 by dividing the integration of the product signal in the diode array by the analogous signal for reactions at pH 6. Three replicates were performed for each reaction condition.

Table S1. Integration of signal corresponding to product 3a at different pH (short reaction times)

	254 nm			230 nm			280 nm		
pH	**#1**	**#2**	**#3**	**#1**	**#2**	**#3**	**#1**	**#2**	**#3**
6.0	165	162	80.0	379	368	174	93.6	91.0	42.2
6.5	227	205	99.1	515	459	213	127	116	52.9
7.0	257	240	129	601	546	279	148	135	68.7
7.5	245	248	131	556	569	283	139	140	69.8
8.0	231	244	134	531	552	288	133	138	71.1
8.5	249	248	135	577	557	286	143	140	71.6
9.0	254	244	138	585	560	294	144	139	74.2

Table S2. Data from Table S1 normalized to corresponding value at pH 6.0

	254 nm			230 nm			280 nm		
pH	**#1**	**#2**	**#3**	**#1**	**#2**	**#3**	**#1**	**#2**	**#3**
6.0	1	1	1	1	1	1	1	1	1
6.5	1.38	1.27	1.24	1.36	1.25	1.23	1.36	1.27	1.25
7.0	1.56	1.48	1.61	1.59	1.48	1.61	1.58	1.48	1.63
7.5	1.49	1.53	1.63	1.47	1.55	1.63	1.49	1.54	1.65
8.0	1.41	1.51	1.67	1.40	1.50	1.66	1.42	1.52	1.68
8.5	1.51	1.53	1.69	1.52	1.51	1.65	1.53	1.54	1.70
9.0	1.55	1.50	1.73	1.54	1.52	1.69	1.54	1.52	1.76

Table S3. Integration of signal corresponding to product 3b at different pH (short reaction times)

	absorption at 230 nm		
pH	**#1**	**#2**	**#3**
6.0	5.0	4.8	6.2
6.5	13	10	12
7.0	15	12	15
7.5	21	20	17
8.0	23	22	19
8.5	21	23	18
9.0	24	22	19

	normalized		
pH	**#1**	**#2**	**#3**
6.0	1	1	1
6.5	2.5	2.1	2.0
7.0	3.0	2.5	2.5
7.5	4.1	4.1	2.7
8.0	4.5	4.6	3.1
8.5	4.3	4.9	3.0
9.0	4.7	4.6	3.0

Table S4. Average relative rates for formation of products 3a and 3b at different pH

Product 3a

pH	**average**	**standard deviation**
6.0	1	–
6.5	1.29	0.06
7.0	1.56	0.06
7.5	1.55	0.07
8.0	1.53	0.11
8.5	1.58	0.07
9.0	1.60	0.09

Product 3b

pH	**average**	**standard deviation**
6.0	1.0	–
6.5	2.2	0.2
7.0	2.7	0.2
7.5	3.6	0.6
8.0	4.1	0.7
8.5	4.0	0.8
9.0	4.1	0.8

2.9 Production of products 3a and 3b at different pH (12 hours)

Reaction buffers at different pH were prepared by mixing dibasic and monobasic potassium phosphate buffers (50 mM). Reactions were run according to the procedure described in Section 2.7 (final enzyme concentration 10 μM). Three replicates were performed for each reaction condition.

Table S5. Turnovers to product in 12 hours at different pH

Product 3a

pH	#1	#2	#3	average	std. dev.
6.5	1213	1758	1855	1609	282
7.0	2537	2846	2768	2717	131
7.5	3098	3012	3335	3148	137
8.0	3101	3326	3509	3312	167
8.5	3894	4276	4223	4131	169
9.0	3885	4035	4063	3994	78

Product 3b

pH	#1	#2	#3	average	std. dev.
6.0	1354	1254	1211	1273	60
6.5	1985	1766	1746	1832	108
7.0	2552	2278	2300	2377	124
7.5	2800	2675	2649	2708	66
8.0	2857	2559	2564	2660	139
8.5	2821	2821	2843	2828	11
9.0	3366	2780	2731	2959	288

Table S6. Integration of signal corresponding to hydrolysis byproduct 4

	254 nm			280 nm		
pH	#1	#2	#3	#1	#2	#3
6.0	1784	1678	1462	217	203	178
6.5	1223	1234	1684	146	148	203
7.0	1082	1050	1241	129	126	150
7.5	861	917	981	103	109	118
8.0	879	853	964	104	102	116
8.5	554	550	599	70	69	76
9.0	554	560	658	70	70	83

Table S7. Data from Table S6 normalized to results at pH 9

	254 nm			280 nm				standard
pH	#1	#2	#3	#1	#2	#3	average	deviation
6.0	3.22	2.99	2.22	3.09	2.92	2.15	2.76	0.42
6.5	2.21	2.20	2.56	2.08	2.12	2.44	2.27	0.17
7.0	1.95	1.87	1.89	1.83	1.81	1.80	1.86	0.05
7.5	1.55	1.64	1.49	1.46	1.56	1.42	1.52	0.07
8.0	1.59	1.52	1.47	1.48	1.46	1.39	1.49	0.06
8.5	1.00	0.98	0.91	1.00	0.99	0.91	0.96	0.04
9.0	1	1	1	1	1	1	1	–

2.10 Turnover frequency with substrates 2a, 2b, and indole

Reactions were performed in pH 9.0 potassium phosphate buffer according to the procedure detailed in Section 2.7, except with reaction times of 15 minutes using substrate **2a** or indole and 45 minutes using substrate **2b**. The rates were expressed as turnovers per minute by dividing the number of turnovers by the reaction time. Three replicates were performed for each reaction condition.

Table S8. Turnover frequency (min^{-1}) to products 3a and 3b (short reaction time)

Product 3a

variant	#1	#2	#3	avg.	std. dev.
Pf(β-Me)	3.56	3.12	2.53	3.07	0.42
Pf(NMB)	30.16	29.70	29.09	29.65	0.44

Product 3b

variant	#1	#2	#3	avg.	std. dev.
Pf(β-Me)	0.42	0.42	0.44	0.43	0.01
Pf(NC)	8.72	8.95	9.09	8.92	0.15

Table S9. Turnover frequency (min^{-1}) to tryptophan (short reaction time)

Tryptophan

variant	#1	#2	#3	avg.	std. dev.
Pf(β-Me)	33	34	34	33.5	0.5
Pf(NMB)	83	79	85	82	3
Pf(NC)	57	63	64	62	3

2.11 Activity with other nitroalkane substrates (Chart 1)

Reactions with other nitroalkane substrates were performed in pH 9.0 potassium phosphate buffer according to the procedure detailed in Section 2.7 (final enzyme concentration 10 μM). For each substrate, individual reactions with *Pf*(β-Me), *Pf*(NMB), and *Pf*(NC) were run to determine which catalyst provided the highest activity. Then, reactions with the best catalyst were run in triplicate, and the product formation was determined by HPLC (diode array detector). For products **3c**–**e**, purified product was used to calibrate HPLC signal to product formation. For products **3f**–**h**, the product yield was estimated by comparing the integration of the product signal to that of the starting material signal. For product **3i**, the HPLC signal was calibrated using purified product **3a**.

Table S10. Turnovers to other amino acid products in 12 hours

product	variant	#1	#2	#3	avg.	std. dev.
3c	*Pf*(NC)	1541	1414	1608	1521	80
3d	*Pf*(NMB)	1232	1222	1199	1217	14
3e	*Pf*(NC)	966	905	972	947	30
3f	*Pf*(NC)	453	455	478	462	11

3. Synthesis and characterization of noncanonical amino acid products

3.1 General information for biocatalytic reactions

Proton and carbon NMR spectra were recorded either on a Bruker 400 MHz (100 MHz) spectrometer equipped with a cryogenic probe. Proton chemical shifts are reported in ppm (δ) relative to tetramethylsilane and calibrated using the residual solvent resonance (DMSO-d_6, δ 2.50 ppm or D_2O, δ 4.79 ppm). Data are reported as follows: chemical shift (multiplicity [singlet (s), doublet (d), doublet of doublets (dd), doublet of doublets of doublets (ddd), triplet (t), multiplet (m)], coupling constants [Hz], integration). Carbon NMR spectra were recorded with complete proton decoupling. Carbon chemical shifts are reported in ppm relative to tetramethylsilane and calibrated using the residual solvent ^{13}C resonance (DMSO-d_6, δ 39.52); for spectra in D_2O, the residual solvent proton resonance was used as an absolute reference. All NMR spectra were recorded at ambient temperature (about 25 °C). Preparative reversed-phase chromatography was performed on a Biotage Isolera One purification system, using C-18 silica as the stationary phase, with methanol as the strong solvent and water (0.1% HCl by weight) as the weak solvent. The

gradient of the eluent (∇) is given as % strong solvent/column volume (CV). High-resolution mass spectrometry (HRMS) was conducted with an Agilent 6200 TOF, with samples ionized by electrospray ionization (ESI). All starting materials were purchased from commercial sources and used without further purification. Liquid chromatography/mass spectrometry (LCMS) was performed on an Agilent 1290 UPLC-LCMS equipped with a 2.1 × 50 mm C-18 silica column, using acetonitrile as the strong solvent and 0.1% (v/v) acetic acid/water as the weak solvent. The optical purity of the products was determined by derivatization with *N*-(5-fluoro-2,4-dinitrophenyl)alaninamide (FDNP-alaninamide) as described below.

The substrate (nitromethyl)benzene (**2a**) was purchased from Sigma Aldrich and purified by flash chromatography (silica, 1–5% ethyl acetate/hexanes). The substrates nitrocyclohexane (**2b**) and nitrocyclopentane (**2c**) were purchased from Sigma Aldrich and used as received. Substrates **2d**–**h** were prepared according to the procedure of Kozlowski and co-workers.[4] Substrate **2i** was prepared according to the procedure of Tomaselli and co-workers.[5] An authentic standard of (*S*)-**5** was purchased from Chem-Impex.

3.2 Preparative reactions for product characterization

Reactions were conducted using up to 1 mmol of substrate. A 20-mL reaction vial with stir bar was charged with nitroalkane, followed by a solution of serine in potassium phosphate buffer (50 mM, pH 9), such that the nitroalkane and serine both had a final concentration of 50 mM (for substrates **2d-f**, 5% v/v DMSO was used as a co-solvent). The corresponding enzyme (see Chart 1) was prepared as heat-treated lysate and added to achieve a final enzyme concentration of 0.010 mM. The reaction vessel was sealed, and then heated to 50 °C with gentle stirring (300 rpm). After 12 hours, the reaction was purified directly by C-18 silica (see below).

For products **3b-c**, the reaction mixture was injected directly onto C-18 silica, eluting with methanol/water (typical gradient: 1 column volume of 1% methanol/water, then 1–100% methanol/water over 10 column volumes). After lyophilization, the products were obtained as the free base.

The other products can exist as the nitro or *aci*-nitro tautomers (see Figure 1b). To isolate the compounds in the nitro form, the reaction mixtures were first acidified to pH < 3 with 0.1% w/w aqueous HCl. This mixture was injected directly onto C-18 silica, and then the product was eluted as the nitro tautomer using a gradient of 1–100% methanol/water (0.1% w/w HCl) over 10 column volumes. Following lyophilization, the product was obtained as the hydrochloride salt.

3.3 Determination of optical purity

FDNP-alaninamide was used as a solution in acetone (33 mM). In a 2-mL vial, the amino acid (0.50 μmol) was dissolved in 1 M aq. $NaHCO_3$ (100 μL). FDNP-alaninamide (10 μL, 0.33 μmol) was added, then the vial was placed in an incubator at 37 °C and shaken at 230 rpm. After 2 h, the reaction mixture was allowed to cool to room temperature, then diluted with 1:1 CH_3CN/H_2O (600 μL). The resulting solution was analyzed directly by LCMS (5% to 95% acetonitrile, Figure S3). Each amino acid was derivatized with both racemic and enantiopure FDNP-alaninamide. The epimerizable compounds **3a** and **3d–g** could not be analyzed directly, so compound **3a** was first converted to compound **5** (see Section 3.6), while compounds **3d–g** were hydrolyzed to the ketone derivatives.

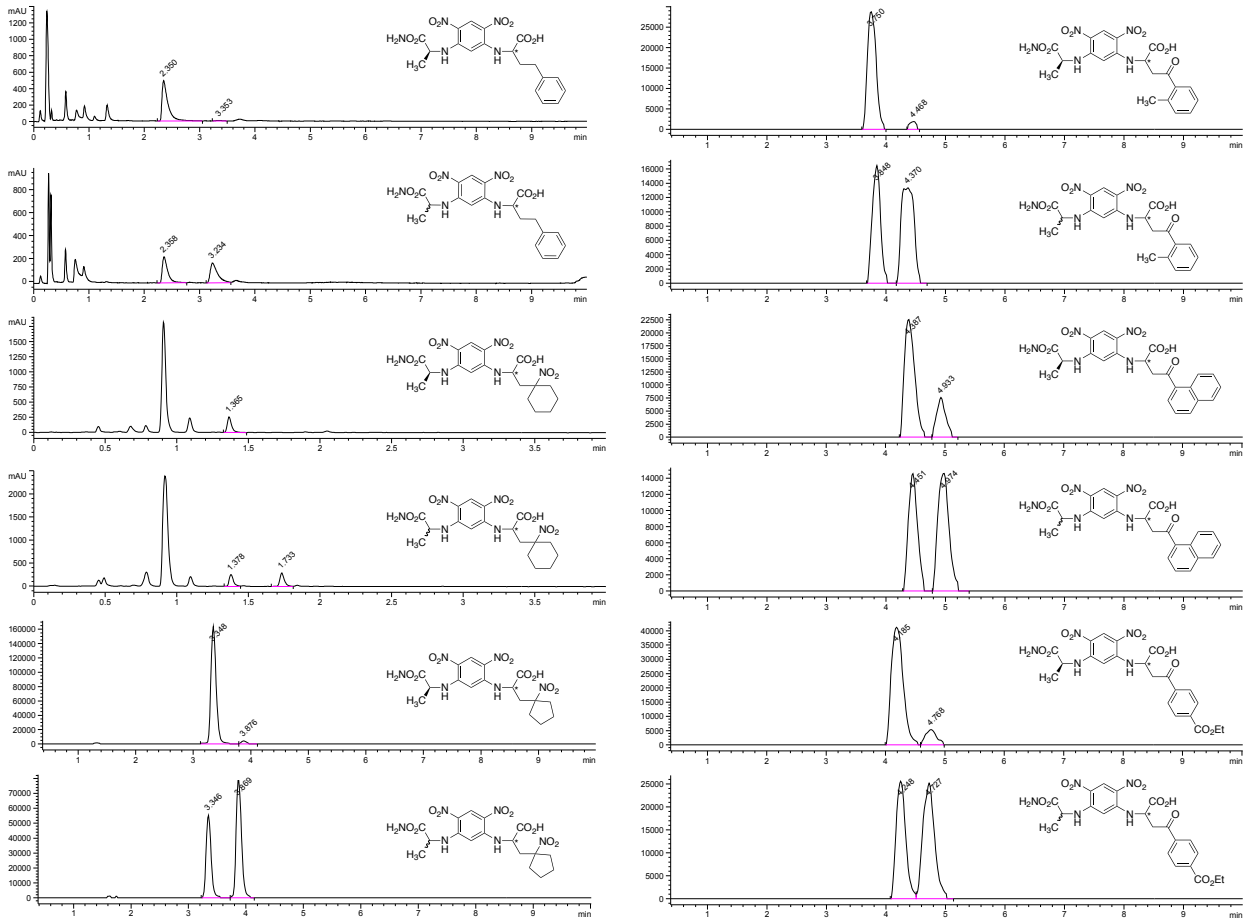


Figure S3. HPLC traces of products derivatized with FDNP-L-alaninamide and FDNP-DL-alaninamde.

3.4 Determination of absolute stereochemistry

Purified product **3a** was converted to compound **5** according to the procedure in Section 3.6. This compound was then derivatized with FDNP-alaninamide (see Section 3.3). Separately, an authentic sample of (*S*)-**5** was derivatized with FDNP-alaninamide by an identical procedure. Comparison of the derivatized compounds by LCMS confirmed that the synthetic compound **5** and, by extension, compounds **3a** and **4** all had the *S* configuration in the major enantiomer (Figure S3). In addition, the specific rotation of synthetic compound **5**•HCl in 3 M aq. HCl ($[\alpha]_D$ = +35.516 at 24.8 °C) corresponds closely to the reported value.[6] The major enantiomers of the other products were inferred to have the *S* configuration by analogy. Notably, the same absolute configuration is favored in the native reaction of the wild-type enzyme. Furthermore, for all products derivatized with FDNP-alaninamide, the major enantiomer eluted first; although there is no strict correlation between elution order on HPLC and absolute configuration, it was observed that L-amino acids derivatized with FDNP-alaninamide tend to elute earlier than their D enantiomers.[7]

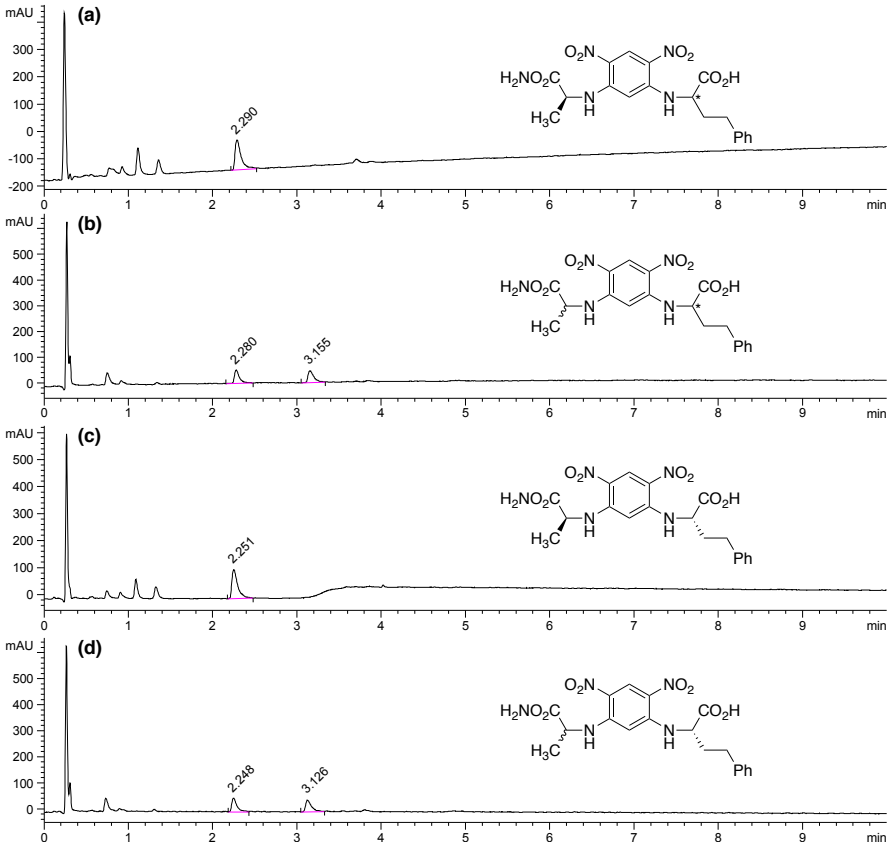


Figure S4. HPLC traces (absorption at 254 nm) of (a) synthetic product **5** tagged with FDNP-L-alaninamide, (b) synthetic product **5** tagged with FDNP-DL-alaninamide, (c) authentic (*S*)-**5** tagged with FDNP-L-alaninamide, and (d) authentic (*S*)-**5** tagged with FDNP-DL-alaninamide.

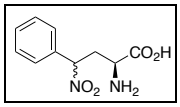
3.5 Hydrolysis of product 3a to ketone 4

In a 5-mL reaction vial with a stir bar, product **3a** (26 mg, 100 μmol) was suspended in water (1000 μL) and stirred vigorously at 24 °C. After 24 h, the product was purified directly by C-18 silica, eluting with methanol and 0.1% w/w aqueous HCl (1 CV of 1% methanol, then 1–100% methanol over 10 CV). The hydrochloride salt of the product was obtained as a white solid (14 mg, 61% yield).

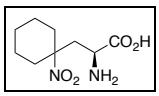
3.6 Denitration of product 3a to product 5

In a 5-mL reaction vial with stir bar, product **3a** (20.7 mg, 92 μmol) was dissolved in 1:1 ethanol/water (1 mL). $Pd(OH)_2/C$ (3.2 mg, 20% w/w, 4.6 μmol) was added. The vessel head-space was flushed with a hydrogen balloon, and then put under static hydrogen pressure (1 bar). The reaction was heated to 65 °C with vigorous stirring. After 12 hours, the reaction was allowed to cool to ambient temperature. The ethanol was removed *in vacuo*, and then the residue was dissolved in 1 M aq. HCl and purified directly by C-18 silica (10 column volumes of 1% methanol/water, then 1–100% methanol/water over 10 column volumes). Fractions containing the product were concentrated *in vacuo*, and then the residue was dissolved in 1 M aq. HCl and transferred to a vial. After lyophilization, the product was obtained as a white solid (18.3 mg, 91% yield of the hydrochloride salt).

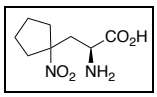
3.7 Characterization of isolated amino acids



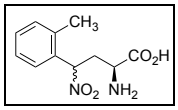
Compound 3a (HCl salt). The product was obtained as a 3:2 mixture of diastereomers. Major diastereomer: **^{1}H NMR** (400 MHz, DMSO-d_6) δ 8.69 (s, 3H), 7.60–7.44 (m, 5H), 6.19 (dd, J = 8.4, 5.7 Hz, 1H), 3.74 (t, J = 7.1 Hz, 1H), 3.10 (dt, J = 14.7, 8.0 Hz, 1H), 2.57 (dt, J = 14.8, 6.3 Hz, 1H). **^{13}C NMR** (100 MHz) δ 169.8, 134.6, 130.0, 129.2, 127.7, 86.4, 49.6, 33.2. Minor diastereomer: **^{1}H NMR** (400 MHz, DMSO-d_6) δ 8.69 (s, 3H), 7.60–7.44 (m, 5H), 6.27 (t, J = 7.3 Hz, 1H), 3.61 (t, J = 6.6 Hz, 1H), 3.00 (dt, J = 14.8, 7.3 Hz, 1H), 2.65 (dt, J = 15.0, 6.8 Hz, 1H). **^{13}C NMR** (100 MHz) δ 169.7, 134.1, 130.2, 129.3, 127.9, 85.9, 49.2, 33.5. **HRMS** (ESI) (m/z) for $[M+H]^+$ ($C_{10}H_{13}N_2O_4$) calculated 225.0870, observed 225.0866. **ee** 97%.



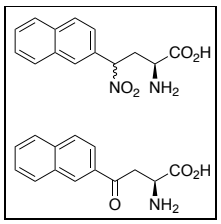
Compound 3b (free base). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 7.45 (s, 2H), 3.08 (dd, J = 7.2, 4.4 Hz, 1H), 2.27 (**AB**X, J_{AX} = 7.3 Hz, J_{BX} = 4.4 Hz, J_{AB} = 15.2 Hz, ν_{AB} = 223.4 Hz, 2H), 2.34–2.25 (m, 2H), 1.86–1.66 (m, 2H), 1.63–1.53 (m, 2H), 1.53–1.45 (m, 1H), 1.35–1.15 (m, 3H). **^{13}C NMR** (100 MHz) δ 169.5, 90.8, 49.9, 33.3, 32.7, 23.9, 21.9, 21.9. **HRMS** (ESI) (m/z) for $[M+H]^+$ ($C_9H_{17}N_2O_4$) calculated 217.1110, observed 217.1182. **ee** >99%.



Compound 3c (free base). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 7.40 (s, 2H), 3.07 (t, J = 6.2 Hz, 1H), 2.47–2.35 (m, 2H), 2.42 (**AB**X, J_{AX} = 7.0 Hz, J_{BX} = 5.4 Hz, J_{AB} = 15.1 Hz, ν_{AB} = 181.2 Hz, 2H), 2.02–1.87 (m, 2H), 1.76–1.57 (m, 4H). **^{13}C NMR** (100 MHz) δ 169.8, 98.3, 51.1, 37.0, 36.6, 23.8, 23.6. **HRMS** (ESI) (m/z) for $[M+H]^+$ ($C_8H_{15}N_2O_4$) calculated 203.1026, observed 203.1019. **ee** 98%.

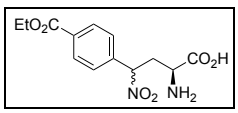


Compound 3d (HCl salt). The product was obtained as a 2:1 mixture of diastereomers. Major diastereomer: **^{1}H NMR** (400 MHz, DMSO-d_6) δ 8.58 (s, 3H), 7.46–7.25 (m, 4H), 6.31 (dd, J = 8.5, 6.0 Hz, 1H), 3.76 (t, J = 6.9 Hz, 1H), 3.14–3.05 (m, 1H), 2.61–2.49 (m, 1H), 2.48 (s, 3H). **^{13}C NMR** (100 MHz) δ 169.8, 137.4, 132.9, 131.1, 129.8, 126.9, 126.0, 82.5, 49.6, 32.8, 19.0. Minor diastereomer: **^{1}H NMR** (400 MHz, DMSO-d_6) δ 8.58 (s, 3H), 7.39–7.25 (m, 4H), 6.34 (dd, J = 9.0, 4.8 Hz, 1H), 3.89 (dd, J = 7.6, 4.9 Hz, 1H), 3.07–2.99 (m, 1H), 2.61–2.49 (m, 1H), 2.44 (s, 3H). **^{13}C NMR** (100 MHz) δ 169.8, 137.0, 132.9, 131.1, 129.8, 127.0, 125.9, 82.6, 49.2, 32.9, 18.9. **HRMS** (ESI) (m/z) for $[M+H]^+$ ($C_{11}H_{15}N_2O_4$) calculated 239.1026, observed 239.1031. **ee** 92%.

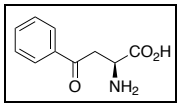


Compound 3e. **HRMS** (ESI) (m/z) for $[M+H]^+$ ($C_{14}H_{15}N_2O_4$) calculated 275.1026, observed 275.1028. **ee** 61%.

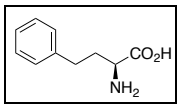
For characterization by NMR, the compound was converted to the ketone (HCl salt). **^{1}H NMR** (400 MHz, D_2O) δ 8.60 (d, J = 1.7 Hz, 1H), 8.06 (d, J = 8.1 Hz, 1H), 8.02–7.93 (m, 3H), 7.70 (ddd, J = 8.3, 6.9, 1.4 Hz, 1H), 7.64 (ddd, J = 8.2, 6.9, 1.3 Hz, 1H), 4.55 (t, J = 5.1 Hz, 1H), 4.01 (d, J = 5.1 Hz, 2H). **^{13}C NMR** (100 MHz, D_2O) δ 199.0, 171.8, 135.9, 132.3, 132.1, 131.1, 129.8, 129.5, 128.7, 127.8, 127.4, 123.0, 49.1, 38.2.



Compound 3f (HCl salt). The product was obtained as a 3:2 mixture of diastereomers.Major diastereomer: **^{1}H NMR** (400 MHz, DMSO-d_6) δ 8.69 (s, 3H), 8.03 (d, J = 8.3 Hz, 2H), 7.68 (d, J = 8.4 Hz, 2H), 6.32 (dd, J = 8.4, 5.5 Hz, 1H), 4.33 (q, J = 7.1 Hz, 2H), 3.84 (t, J = 7.0 Hz, 1H), 3.14 (dt, J = 15.3, 7.9 Hz, 1H), 2.58 (dt, J = 14.7, 6.1 Hz, 1H), 1.32 (t, J = 7.1 Hz, 3H). **^{13}C NMR** (100 MHz, DMSO-d_6) δ 169.7, 165.1, 139.1, 131.2, 129.9, 128.3, 85.7, 61.1, 49.5, 33.1, 14.2. Minor diastereomer: **^{1}H NMR** (400 MHz, DMSO-d_6) δ 8.69 (s, 3H), 8.05 (d, J = 8.3 Hz, 2H), 7.73 (d, J = 8.4 Hz, 2H), 6.38 (t, J = 7.1 Hz, 1H), 4.33 (q, J = 7.1 Hz, 2H), 3.69 (t, J = 6.4 Hz, 1H), 3.04 (dt, J = 14.6, 7.2 Hz, 1H), 2.69 (dt, J = 14.9, 6.6 Hz, 1H), 1.32 (t, J = 7.1 Hz, 3H). **^{13}C NMR** (100 MHz, DMSO-d_6) δ 169.6, 165.1, 138.7, 131.4, 130.0, 128.5, 85.3, 61.1, 49.0, 33.4, 14.2. **HRMS** (ESI) (m/z) for $[M+H]^+$ ($C_{13}H_{17}N_2O_6$) calculated 297.1081, observed 297.1087. **ee** 81%.



Compound 4 (HCl salt). **^{1}H NMR** (400 MHz, D_2O) δ 8.07–7.99 (m, 2H), 7.74 (ddt, J = 8.7, 7.0, 1.3 Hz, 1H), 7.64–7.54 (m, 2H), 4.48 (t, J = 5.2 Hz, 1H), 3.90 (d, J = 5.2 Hz, 2H). **^{13}C NMR** (100 MHz, D_2O) δ 199.5, 172.1, 135.0, 134.8, 129.0, 128.3, 49.2, 38.3. **HRMS** (ESI) (m/z) for $[M+H]^+$ ($C_{10}H_{12}NO_3$) calculated 194.0812, observed 194.0803. **ee** 95%.



Compound 5. **^{1}H NMR** (400 MHz, D_2O/NaOD) δ 7.10–7.03 (m, 2H), 7.03–6.93 (m, 3H), 2.94 (dd, *J* = 7.0, 5.7 Hz, 1H), 2.33 (t, *J* = 8.2 Hz, 1H), 1.64–1.56 (m, 1H), 1.56–1.45 (m, 1H). **^{13}C NMR** (100 MHz, D_2O) δ 183.2, 142.1, 128.6, 128.4, 128.2, 55.6, 36.7, 31.3. **HRMS** (ESI) (*m*/*z*) for $[M+H]^+$ ($C_{10}H_{14}NO_2$) calculated 180.1019, observed 180.1019. **[α]$_D$** of **5**•HCl in 3 M aq. HCl at 24.8 °C, +35.516°. **ee** 97%.

4. References

(1) Herger, M.; van Roye, P.; Romney, D. K.; Brinkmann-Chen, S.; Buller, A. R.; Arnold, F. H. Synthesis of β-Branched Tryptophan Analogues Using an Engineered Subunit of Tryptophan Synthase. *J. Am. Chem. Soc.* **2016**, *138* (27), 8388–8391.

(2) Kille, S.; Acevedo-Rocha, C. G.; Parra, L. P.; Zhang, Z.-G.; Opperman, D. J.; Reetz, M. T.; Acevedo, J. P. Reducing Codon Redundancy and Screening Effort of Combinatorial Protein Libraries Created by Saturation Mutagenesis. *ACS Synth. Biol.* **2013**, *2* (2), 83–92.

(3) Gibson, D. G.; Young, L.; Chuang, R.-Y.; Venter, J. C.; Hutchison, C. A.; Smith, H. O. Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases. *Nat. Methods* **2009**, *6* (5), 343–345.

(4) Walvoord, R. R.; Berritt, S.; Kozlowski, M. C. Palladium-Catalyzed Nitromethylation of Aryl Halides: An Orthogonal Formylation Equivalent. *Org. Lett.* **2012**, *14* (16), 4086–4089.

(5) Ballistreri, F. P.; Barbuzzi, E.; Tomaselli, G. A.; Toscano, R. M. Useful Oxidation Procedure of Oximes to Nitro Compounds with Benz-Mo in Acetonitrile. *Synlett* **1996**, *1996* (11), 1093–1094.

(6) L-Homophenylalanine hydrochloride at https://www.sigmaaldrich.com/catalog/product/aldrich/544051.

(7) Marfey, P. Determination of D-Amino Acids. II. Use of a Bifunctional Reagent, 1,5-Difluoro-2,4-Dinitrobenzene. *Carlsberg Res. Commun.* **1984**, *49* (6), 591–596.