Published in final edited form as:

ACS Catal. 2019 September 6; 9(9): 8726–8730. doi:10.1021/acscatal.9b02089.

Nitroalkanes as Versatile Nucleophiles for Enzymatic Synthesis of Noncanonical Amino Acids

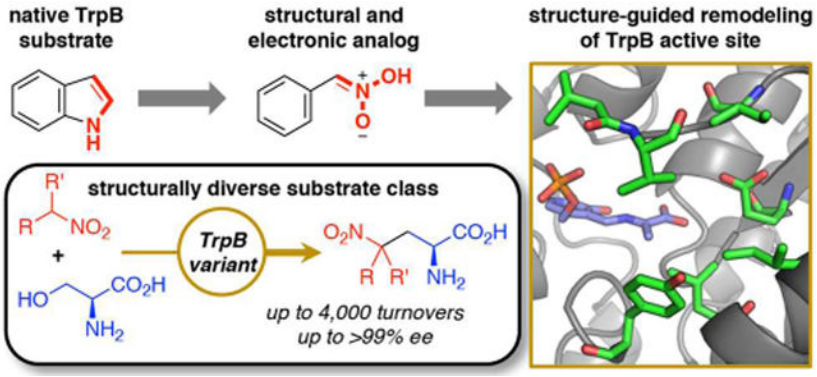
David K. Romney[*,†,‡], **Nicholas S. Sarai**[†], **Frances H. Arnold**[*,†]

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

Abstract

C–C bond-forming reactions often require nucleophilic carbon species rarely compatible with aqueous reaction media, thus restricting their appearance in biocatalysis. Here we report the use of nitroalkanes as a structurally versatile class of nucleophilic substrates for C–C bond formation catalyzed by variants of the β-subunit of tryptophan synthase (TrpB). The enzymes accept a wide range of nitroalkanes to form noncanonical amino acids, here the nitro group can serve as a handle for further modification. Using nitroalkane nucleophiles greatly expands the scope of compounds made by TrpB variants and establishes nitroalkanes as a valuable substrate class for biocatalytic C–C bond formation.

Graphical Abstract



Keywords

biocatalysis; C–C bond formation; tryptophan synthase; noncanonical amino acid; nitroalkane

[*]**Corresponding Author** dromney@aralezbio.com, frances@cheme.caltech.edu.
[‡]Current address: Aralez Bio, 101 West Ferndale Avenue, Sunnyvale, CA, 94085

Conflict of Interest

The authors declare the following competing financial interest: The contents of this paper are the subject of a patent application submitted by Caltech, and some authors are entitled to a royalty on revenues arising from that patent.

Supporting Information. The Supporting Information is available free of charge on the ACS Publications website. Additional figures, experimental procedures, and NMR spectra of products (PDF).

Enzymes are useful tools for organic synthesis because they can be produced sustainably and engineered to catalyze reactions with high rate acceleration and selectivity. Biocatalysts such as transaminases and ketoreductases are thus frequently deployed in synthetic processes.[1,2] However, the ability of biocatalysis to mediate formation of C–C bonds, one of the most important transformations for organic synthesis, is hindered by the fact that the nucleophilic carbon species ubiquitous in synthetic chemistry, such as organometallic compounds, tend to be unstable under the aqueous conditions required by most enzymes. Aldolases overcome this by forming a covalent bond with the nucleophilic substrate, which then forms the active nucleophilic species in the substrate-binding pocket.[3-5] Although these enzymes accept a broad range of *electrophiles*, the requirement for the nucleophilic substrate to form a covalent adduct with the enzyme inherently limits the *nucleophile* scope. Conversely, enzymes can activate *electrophilic* substrates, either covalently or through polar interactions such as hydrogen bonding, which can then be engaged by *nucleophiles* that are sufficiently reactive, but stable to aqueous conditions. One exemplar of the latter enzyme class is the β-subunit of tryptophan synthase (TrpB), which uses pyridoxal phosphate (PLP) as a cofactor to convert serine into a highly electrophilic amino-acrylate species (Figure 1a).[6] This in turn can react with a nucleophile to form an enantiopure amino acid. In nature, the nucleophilic substrate is indole, an electron-rich arene that leads to the amino acid tryptophan as the product. Due to the modular and convergent nature of this reaction, we envisioned TrpB as a general biocat-alytic platform capable of combining structurally diverse nucleophiles with serine (or an analogous substrate) to make enantioenriched noncanonical amino acids (ncAAs). However, while we have been able to use protein engineering to expand the substrate scope of TrpB to include dozens of indole analogs,[7-9] we struggled to find an active carbon nucleophile that did not require the indole substructure.

Nitroalkanes have been used as nucleophiles in organic synthesis dating back to the venerable Henry reaction.[10,11] Even in the absence of base, these compounds readily tautomerize to a structure that is nucleophilic at carbon (**1**,Figure 1b), and as such have been used as substrates for C–C bond-forming reactions catalyzed by electrophile-activating enzymes.[12-15] In principle, any nitroalkane with a proton at the α-carbon could react; this would make nitroalkanes a more versatile substrate class than other mild carbanions like malonates, which require at least two electron-withdrawing groups. Indeed, nitroalkanes have been shown to react with chemically produced amino-acrylates to form a wide range of amino acids,[16-20] although these methods were not enantioselective. These reactions also required strong base or harsh conditions (e.g, neat nitroalkane in a microwave reactor), suggesting that biocatalytic conditions might not support this reaction design. Nonetheless, we wondered whether TrpB variants could promote similar reactions under milder conditions.

From our previous efforts engineering activity for non-native substrates, we accrued a collection of TrpB variants with diverse properties (e.g., active-site configuration, conformational dynamics). Since we were unsure which combination of properties would provide the best starting point for such a fundamentally different substrate class as nitroalkanes, we began by testing a broad panel of TrpB variants for activity with serine and (nitromethyl)benzene (**2a**, Scheme 1), which was chosen due to its structural resemblance to

indole. These variants all possess excellent thermostability, and previous reactions with serine and indole analogs were run at 75 °C. At this temperature, however, we observed significant decomposition of **2a** over the 12-hour reaction time (Figure S1); test reactions were thus run at 50 °C. We were pleased to observe by HPLC that most tested variants did indeed form desired product **3a**. The best activity (1,200 turnovers, Figure 2) was obtained with a variant that had been engineered to synthesize β-methyltryptophan from L-threonine and indole.[21] This variant, hereafter called *Pf*(β-Me), is derived from TrpB of *Pyrococcus furiosus* and contains eight mutations (see Table 1).

Though initial results with substrate **2a** were encouraging, we wanted to develop a more general method for ncAA synthesis that would accept substrates bearing no structural resemblance to indole. We therefore tested *Pf*(β-Me) with nitrocyclohexane (**2b**), which lacks any aromatic moieties and has a more sterically congested nucleophilic carbon. To our surprise, this substrate was also accepted by the enzyme, forming product **3b** with about 500 turnovers (Figure 2).

Since none of the eight mutations in *Pf*(β-Me) are in the active site, the substrate-binding pocket is still optimized for the native substrates, indole and serine. We consulted a crystal structure of *Pf*(β-Me) with serine bound as the amino-acrylate intermediate and identified seven residues that comprised the indole-binding site (Figure 3).[22] We then constructed seven enzyme libraries in which each residue was subjected independently to site-saturation mutagenesis and tested each against substrates **2a** and **2b**, reasoning that these structurally distinct substrates would favor different active-site mutations.

Turnover for both substrates increased with mutation of L161 to smaller residues like valine, alanine, and serine, though the optimal mutations were different. Substrate **2a**, for example, favored serine, which almost doubled activity to 2,100 turnovers, while alanine provided a more modest improvement and valine was essentially equivalent to parent. Substrate **2b**, on the other hand, exhibited an almost three-fold increase in turnover in all three single-mutants. The other libraries either revealed no improvements over parent or gave results that were completely orthogonal between substrates. In particular, the variant with the I183P mutation, hereafter called *Pf*(NMB), showed twice the activity for substrate **2a** (2,700 turnovers), but no improvement with substrate **2b**, while the variant with the I165T mutation had almost quadruple the activity for **2b** (1,900 turnovers), but no improvement for substrate **2a**. Since the mutations at L161 and I183 do not introduce new polar or charged sidechains, their role is most likely to reshape the nucleophile-binding pocket to promote reactive binding poses of the nitroalkanes. The mutation of I165 to threonine, on the other hand, adds a polar functional group to what was previously an entirely aliphatic sidechain; interestingly, serine was also beneficial at this position, suggesting that the improvement particularly depends on the hydroxyl group. To react, however, the nitro group of substrate **2b** must face away from residue 165, so a direct interaction with the hydroxyl group and the substrate is unlikely. Thus, the precise effect of this mutation is as yet unclear.

Having identified multiple improved single-mutant variants, we next constructed libraries to search for beneficial combinations of these mutations. For both substrates, position 161 was allowed to be L (wild-type), V, A, or S. For substrate **2a**, this focused L161 library was

randomly recombined with I183P, whereas for substrate **2b**, this library was randomly recombined with I165T. In the screen with substrate **2a**, the top variants were all single mutants, suggesting that mutations at L161 and I183 were not beneficial in combination. In the screen with substrate **2b**, on the other hand, the best variant combined mutations I165T and L161V to give 2,700 turnovers, an improvement greater than five-fold relative to parent; this variant is hereafter called *Pf*(NC).

The activity of the enzyme can almost certainly be optimized further. For the purposes of this study, however, we elected to examine how the reaction conditions affect the activity with nitroalkanes. We were particularly interested to know if a more basic pH would increase the nucleophilicity of the nitroalkane substrates. To this end, reactions at different pH were run to low conversion, and the relative amounts of product formation were compared (see SI Section 2.8). This should approximate the relative rates of product formation as a function of pH. As expected, the reaction rates for both substrates were lowest at pH 6, but increased with more basic pH before remaining constant above pH 7–8 (Figure 4). This effect was much more pronounced for substrate **2b**, which reacted four times as quickly at pH > 8 compared to pH 6.

Although initial rate is an important reaction property, the utility of a preparative reaction ultimately depends on turnovers to product. Fortunately, total turnovers followed a similar trend (Figure S2), with product **3a** experiencing an additional boost in turnover at pH > 8.5; this boost appears to be correlated with a decrease in formation of a byproduct that results from hydrolysis of product **3a**. Thus, subsequent reactions were run at pH 9.

The actual concentration of the nitroalkane substrates in the reactions is difficult to know, since the reactions are heterogeneous and the substrate miscibility is likely dependent on the tautomeric equilibrium. These factors preclude precise determination of kinetic factors like k_{cat} and K_M. However, it can be observed that under these reaction conditions, the active-site mutations greatly increase the initial rates of the reactions with the nitroalkane substrates. Whereas *Pf*(β-Me) converted substrate **2a** with only 3 turnovers per minute, the I183P mutation increases this rate ten-fold (Table 2, entry 1). With substrate **2b**, which is consumed much more slowly, the improvement is even greater, with mutations L161V and I165T increasing the initial rate twenty-fold (Table 2, entry 2). If the initial rates are taken as an approximation for k_{cat}, then the optimized activity with the nitroalkanes is comparable to the activity of the natural enzyme with the native substrates (indole and serine).[23] Strangely, the active-site mutations also increased the rate of the native reaction (Table 2, entry 3).

Under standard reaction conditions (5,000 maximum turnovers, 12 hours), products **3a** and **3b** were formed with 4,000 and 2,700 turnovers, respectively (Chart 1). Although not as optimized, the catalysts also exhibited promising activity with a range of other nitroalkanes, including nitrocyclopentane, which was converted to product **3c** with 1,500 turnovers. The enzymes can also tolerate analogs of substrate **2a** with substituents at all positions on the aryl ring (e.g., products **3d**–**f**). The method-substituted product **3g** was not detected with any of the enzymes, though this may be due to product instability.

Although products like **3a** technically contain two stereocenters, the carbon adjacent to the nitro group is readily deprotonated, leading to an approximately 1:1 mixture of diastereomers. We therefore tested an analog of substrate **2a**, in which one of the acidic protons was replaced with a methyl group. We also tested substrate **2a** with threonine in place of serine, since previously we showed that TrpB enzymes could convert β-substituted serine analogs into products with defined stereocenters at both the α and β positions.[21,24] Although we observed new compounds matching the expected masses of products **3h** and **3i**, these were formed in low levels. With further engineering, it may be possible to use this methodology to synthesize ncAAs with multiple stereocenters.

A key feature of the nitro group is that it can be directly converted into other functional groups or removed entirely. Since Crossley and co-workers have already shown how compounds resembling product **3b** are precursors to other classes of amino acid,[25] we focused on exploring product **3a** as a precursor (Scheme 2). For example, the nitro group can be hydrolyzed to ketone **4**, which is an analog of kynurenin, a component of the antibiotic daptomycin. The nitro group can also be removed entirely to make homophenylalanine (**5**), a component of the anticancer drug carfilzomib. Thus, the enzymatic products can be used to prepare a much wider range of compounds.

In conclusion, nitroalkanes were demonstrated to be a versatile substrate class for variants of the enzyme TrpB from *Pyrococcus furiosus*. Coupling of nitroalkanes with serine delivers enantioenriched noncanonical amino acids with aryl and alkyl side-chains. Even for substrates with little structural similarity to indole, as few as two mutations were sufficient to increase the reaction rate to levels comparable to the native enzymatic reaction. Identification and exploitation of substrates that form water-compatible nucleophilic carbon species can greatly expand the products accessible to biocatalysis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENT

This work was funded by the Rothenberg Innovation Initiative at Caltech, and by a NIH R01 grant (GM125887-01A1). The authors thank Kai Chen for assistance with experiments, and Ella Watkins-Dulaney and Kelly Zhang for helpful discussions regarding the manuscript

REFERENCES

(1). Hughes DL Biocatalysis in Drug Development—Highlights of the Recent Patent Literature. Org. Proc. Res. Dev 2018, 22, 1063–1080.

(2). Choi J-M; Han S-S; Kim H-S Industrial Applications of Enzyme Biocatalysis: Current Status and Future Aspects. Biotechnol. Adv 2015, 33, 1443–1454. [PubMed: 25747291]

(3). Schmidt NG; Eger E; Kroutil W Building Bridges: Biocat-alytic C–C-Bond Formation toward Multifunctional Products. ACS Catal. 2016, 6,4286–1311. [PubMed: 27398261]

(4). Miao Y; Rahimi M; Geertsema EM; Poelarends GJ Recent Developments in Enzyme Promiscuity for Carbon-Carbon Bond-Forming Reactions. Curr. Opin. Chem. Biol 2015, 25, 115–123. [PubMed: 25598537]

(5). Müller M Recent Developments in Enzymatic Asymmetric C–C Bond Formation. Adv. Synth. Catal 2012, 354, 3161–3174.

(6). Dunn MF; Niks D; Ngo H; Barends TRM; Schlichting I Tryptophan Synthase: The Workings of a Channeling Nanomachine. Trends Biochem. Sci 2008, 33, 254–264. [PubMed: 18486479]

(7). Murciano-Calles J; Romney DK; Brinkmann-Chen S; Buller AR; Arnold FH A Panel of TrpB Biocatalysts Derived from Tryptophan Synthase through the Transfer of Mutations That Mimic Allosteric Activation. Angew. Chem. Int. Ed 2016, 55, 11577–11581.

(8). Romney DK; Murciano-Calles J; Wehrmüller JE; Arnold FH Unlocking Reactivity of TrpB: A General Biocatalytic Platform for Synthesis of Tryptophan Analogues. J. Am. Chem. Soc 2017, 139, 10769–10776. [PubMed: 28708383]

(9). Boville CE; Romney DK; Almhjell PJ; Sieben M; Arnold FH Improved Synthesis of 4-Cyanotryptophan and Other Tryptophan Analogues in Aqueous Solvent Using Variants of TrpB from Thermotoga Maritima. J. Org. Chem 2018, 83, 7447–7452. [PubMed: 29651849]

(10). Henry LCR Formation Synthétique d'alcools Nitrés. Acad. Sci. Ser. C 1895, 1265–1268.

(11). Henry FCR Formation Synthétique d'alcools Nitrés. Bull. Soc. Chim. Fr 1895, 13, 999–1002.

(12). Gruber-Khadjawi M; Purkarthofer T; Skranc W; Griengl H Hydroxynitrile Lyase-Catalyzed Enzymatic Nitroaldol (Henry) Reaction. Adv. Synth. Catal 2007, 349, 1445–1450.

(13). Milner SE; Moody TS; Maguire AR Biocatalytic Approaches to the Henry (Nitroaldol) Reaction. Eur. J. Org. Chem 2012, 3059–3067.

(14). Garrabou X; Macdonald DS; Hilvert D Chemoselective Henry Condensations Catalyzed by Artificial Carboligases. Chem. Eur.J2017, 23, 6001–6003. [PubMed: 28070900]

(15). Garrabou X; Verez R; Hilvert D Enantiocomplementary Synthesis of γ-Nitroketones Using Designed and Evolved Carboligases. J. Am. Chem. Soc 2017, 139, 103–106. [PubMed: 27992715]

(16). Wieland T; Ohnacker G; Ziegler W Aminosaure-Synthesen mit α-Acylamino-Acrylestern. Chem. Ber 1957, 90, 194–201.

(17). Bigge CF; Wu J-P; Drummond JT; Coughenour LL; Hanchin CM Excitatory Amino Acids: 6-Phosphonomethyltetrahy-dro-4-Pyrimidinecarboxylic Acids and Their Acyclic Analogues Are Competitive N-Methyl-D-Aspartic Acid Receptor Antagonists. Bioorg. Med. Chem. Lett 1992, 2, 207–212.

(18). Crossley M; Tansey C A Convenient Method for the Synthesis of β-Substituted α-Amino Acids. Diastereoselective Conjugate Addition of Nitronates to a Chiral Dehydroalanine Derivative. Aust. J. Chem 1992, 45, 479–181.

(19). Crossley MJ; Fung YM; Potter JJ; Stamford AW Convenient Route to γ-Nitro-α-Amino Acids: Conjugate Addition of Nitroalkanes to Dehydroalanine Derivatives. J. Chem. Soc., Perkin Trans 1 1998, 1113–1122.

(20). Ballini R; Balsamini C; Bartoccini F; Gianotti M; Martinelli C; Savoretti N Improved Synthesis of γ-Nitro-α-Amino Ester and Acid Derivatives. Synthesis 2005, 296–300.

(21). Herger M; van Roye P; Romney DK; Brinkmann-Chen S; Buller AR; Arnold FH Synthesis of β-Branched Tryptophan Analogues Using an Engineered Subunit of Tryptophan Synthase. J. Am. Chem. Soc 2016, 138, 8388–8391. [PubMed: 27355405]

(22). Buller AR; van Roye P; Cahn JKB; Scheele RA; Herger M; Arnold FH Directed Evolution Mimics Allosteric Activation by Stepwise Tuning of the Conformational Ensemble. J. Am. Chem. Soc 2018, 140, 7256–7266. [PubMed: 29712420]

(23). Buller AR; Brinkmann-Chen S; Romney DK; Herger M; Murciano-Calles J; Arnold FH Directed Evolution of the Tryptophan Synthase β-Subunit for Stand-Alone Function Recapitulates Allosteric Activation. Proc. Nat. Acad Sci 2015, 112, 14599–14604. [PubMed: 26553994]

(24). Boville CE; Scheele RA; Koch P; Brinkmann-Chen S; Buller AR; Arnold FH Engineered Biosynthesis of β-Alkyl Tryptophan Analogues. Angew. Chem. Int. Ed 2018, 57, 14764–14768.

(25). Crossley MJ; Fung YM; Kyriakopoulos E; Potter JJ New Synthetic Routes to α-Amino Acids and γ-Oxygenated α-Amino Acids. Reductive Denitration and Oxidative Transformations of γ-Nitro-α-Amino Acids. J. Chem. Soc., Perkin Trans 1 1998, 1123–1130.

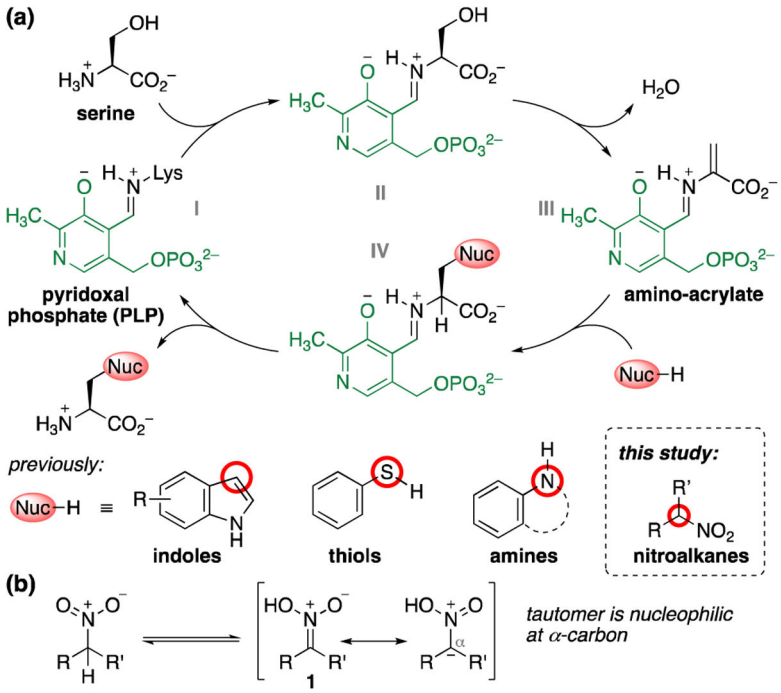


Figure 1.

Amino-acid synthesis using TrpB. (a) Mechanism of enzymatic reaction. (b) Tautomerism of nitroalkanes.

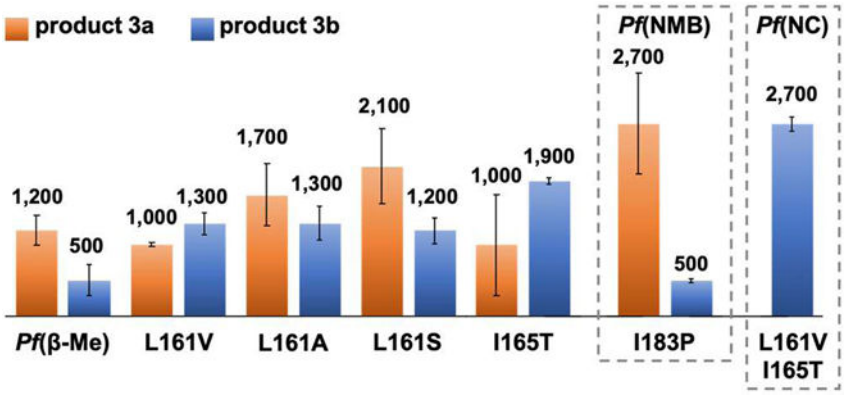


Figure 2.

Turnovers in 12 hours by parent enzyme *Pf*(β-Me) and variants with active-site mutations (see SI Section 2.6 for details).

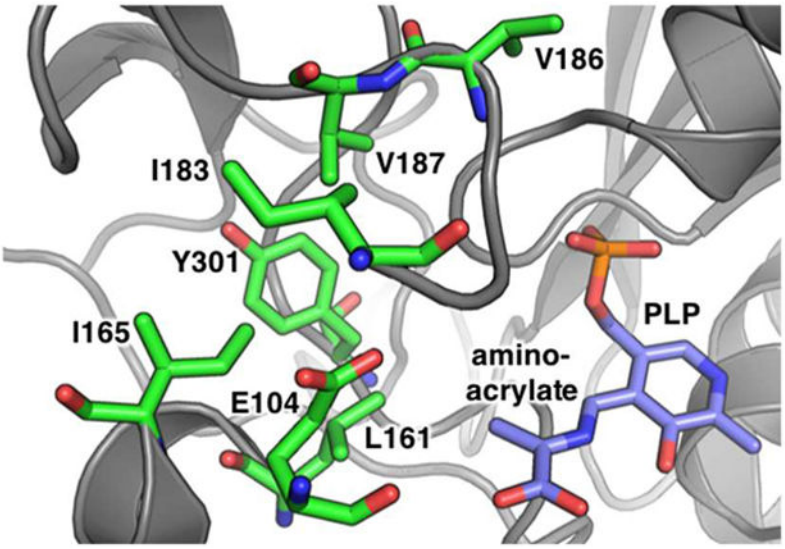


Figure 3.
Active site of *Pf*(β-Me) with the PLP-bound amino-acrylate (PDB ID: 5vm5).

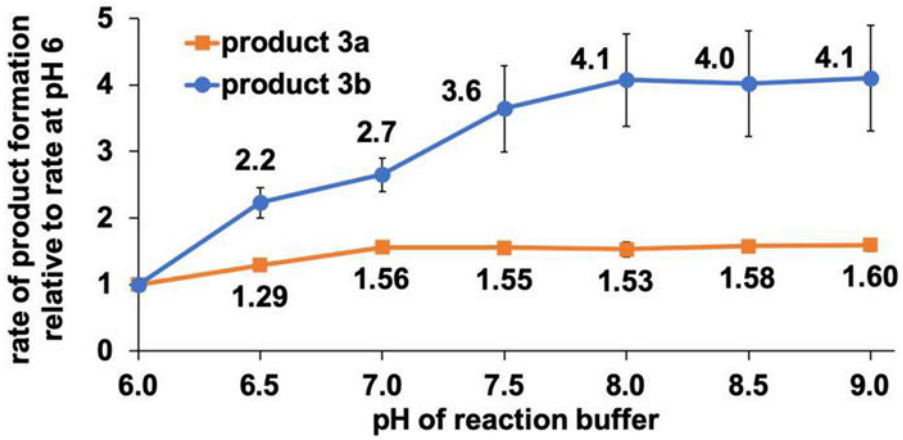
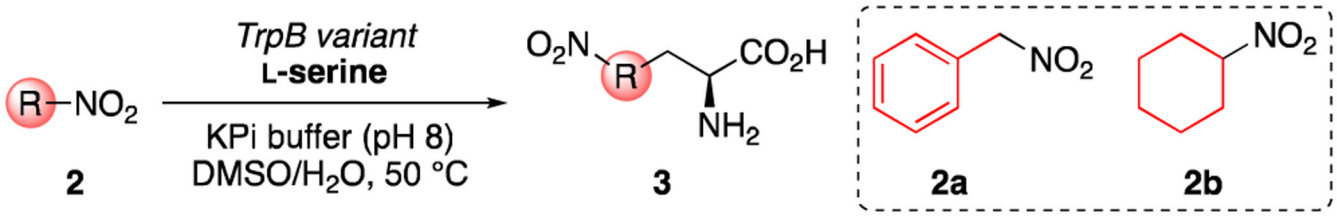
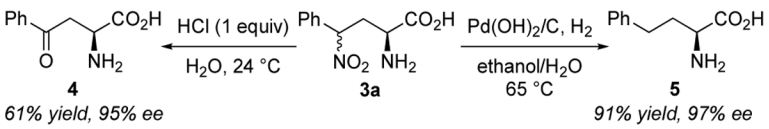


Figure 4.
Dependence of initial rate on pH of reaction buffer (see SI Section 2.8 for details).



Scheme 1.
Screening reactions for enzyme engineering



Scheme 2.

Product 3a as a precursor to other compounds

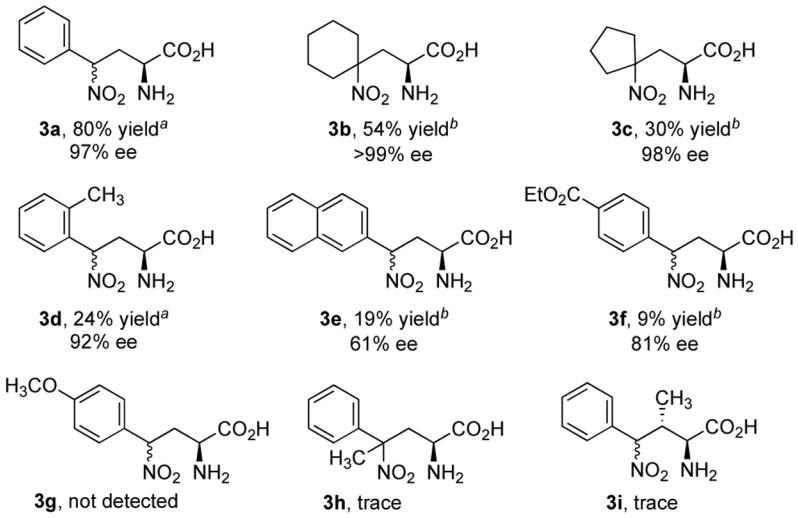


Chart 1. Scope of amino acid products

HPLC yield in 12 hours at 50 °C (50 mM of each substrate, 5,000 maximum turnovers). See SI Section 2.11 for details. [a]Reaction with *Pf*(NMB). [b]Reaction with *Pf*(NC).

Table 1.

Mutations in TrpB variants

variant	mutations
Pf(β-Me)	I16V, E17G, I68V, F95L, F274S, T292S, T321A, V384A
Pf(NMB)	*Pf*(β-Me) + I183P
Pf(NC)	*Pf*(β-Me) + L161V, I165T

Table 2.

Initial rates in reactions with serine[a]

		turnover frequency (min $^{-1}$)		
entry	**substrate**	***Pf*(β-Me)**	***Pf*(NMB)**	***Pf*(NC)**
1	**2a**	3.1 ± 0.4	29.7 ± 0.4	–
2	**2b**	0.43 ± 0.01	–	8.9 ± 0.2
3	indole	33.5 ± 0.5	82 ± 3	62 ± 3

[a]See SI Section 2.10 for details.