HHS Public Access

Author manuscript

J Org Chem. Author manuscript; available in PMC 2019 July 20.

Published in final edited form as:

J Org Chem. 2018 July 20; 83(14): 7447–7452. doi:10.1021/acs.joc.8b00517.

Improved synthesis of 4-cyanotryptophan and other tryptophan analogs in aqueous solvent using variants of TrpB from *Thermotoga maritima*

Christina E. Boville[+], **David K. Romney**[+], **Patrick J. Almhjell**, **Michaela Sieben**, and **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

The use of enzymes has become increasingly widespread in synthesis as chemists strive to reduce their reliance on organic solvents in favor of more environmentally benign aqueous media. With this in mind, we previously endeavored to engineer the tryptophan synthase β-subunit (TrpB) for production of noncanonical amino acids that had previously been synthesized through multistep routes involving water-sensitive reagents. This enzymatic platform proved effective for the synthesis of analogs of the amino acid tryptophan (Trp), which are frequently used in pharmaceutical synthesis as well as chemical biology. However, certain valuable compounds, such as the blue fluorescent amino acid 4-cyanotryptophan (4-CN-Trp), could only be made in low yield, even at elevated temperature (75°C). Here, we describe the engineering of TrpB from *Thermotoga maritima* that improved synthesis of 4-CN-Trp from 24% to 78% yield. Remarkably, although the final enzyme maintains high thermostability (T_{50} = 93°C), its temperature profile is shifted, such that high reactivity is observed at ~37°C (76% yield), creating the possibility for *in vivo* 4-CN-Trp production. The improvements are not specific to 4-CN-Trp; a boost in activity at lower temperature is also demonstrated for other Trp analogs.

TOC image

[*] frances@cheme.caltech.edu, Phone: (626) 395-4162, Fax: (626) 568-8743.
[+]These authors contributed equally to this work.

Supporting Information

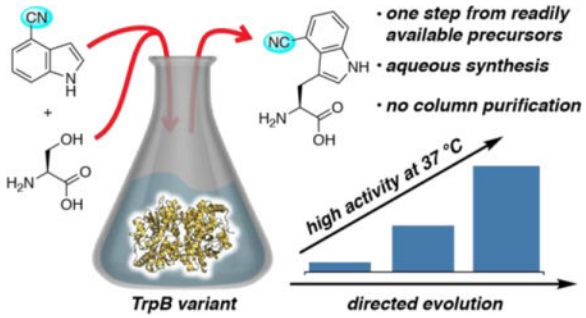
Results of site-saturation mutagenesis libraries. LCMS calibration curves for Chart 1. HPLC data for Figure 2. HPLC data for Chart 1 and indole. NMR spectrum of 4-cyanotryptophan from Scheme 2.

ORCID:

Christina E. Boville: 0000-0002-2577-9343
David K. Romney: 0000-0003-0498-7597
Patrick J. Almhjell: 0000-0003-0977-841X
Michaela Sieben: 0000-0002-4412-7148
Frances H. Arnold: 0000-0002-4027-364X

Notes

The authors declare the following competing financial interest(s): 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.



Introduction

Noncanonical α-amino acids (ncAAs) resemble the building blocks of natural proteins but are not themselves used in protein synthesis. Despite this, ncAAs are prevalent precursors for functional synthetic compounds, including over 12% of the 200 top-selling pharmaceuticals.[1] However, ncAAs are challenging synthetic targets, since they possess at minimum two reactive functional groups (the amine and carboxylic acid) and typically have at least one stereocenter. As a result, synthetic routes to ncAAs typically require multiple steps, most of which use organic solvents.[2,3] One of the most direct routes to ncAAs is to add a nucleophile to the β-position of a serine-derived lactone[4–6] or aziridine[7,8] (Figure 1a), but this approach has certain drawbacks, such as the need to pre-synthesize the water-sensitive electrophilic reactants.

Enzymes are widely applied to the synthesis of ncAAs since they circumvent many of the limitations of chemical methods. Not only do these catalysts function in aqueous media, but they also exhibit chemoselectivity that obviates the need for protecting groups, thereby trimming synthetic steps. In addition, the reactions are often highly stereoselective. Unfortunately, most enzymatic methods to synthesize ncAAs, such as those that rely on hydrolases or transaminases, require that the majority of the final product be synthesized in advance, usually by chemical means, with the enzyme only appearing at the end to set the stereochemistry. By contrast, enzymes like tryptophan synthase,[9–14] which uses the cofactor pyridoxal 5′-phosphate (PLP, Figure 1b), can form ncAAs by nucleophilic substitution at the β-position of readily available amino acids like serine. In this reaction scheme, the enzyme forms an active electrophilic species, the amino-acrylate (Figure 1c), directly in the active site, which is then intercepted by a nucleophilic substrate. These reactions can be run in aqueous conditions that would hydrolyze the serine-derived lactones or aziridines. Furthermore, the enzyme active site can bind the substrates to accelerate the reaction and control the regioselectivity of nucleophilic substitution.

The ncAA 4-cyanotryptophan (4-CN-Trp) was previously reported to exhibit blue fluorescence (λ_{max} ~ 405 nm) with a high quantum yield and long lifetime.[15] These properties, among others, make 4-CN-Trp an attractive small-molecule fluorophore for imaging studies *in vitro* and *in vivo*. However, the chemical synthesis requires multiple steps, including a low-yielding Pd-catalyzed cyanation reaction (Scheme 1a). We were excited to observe that an engineered variant of the β-subunit of tryptophan synthase (TrpB) from hyperthermophilic bacterium *Thermotoga maritima* could form 4-cyanotryptophan in

one step from readily available 4-cyanoindole and serine (Scheme 1b).[16] We therefore wished to engineer this variant further to improve 4-CN-Trp production.

Results

Increasing activity with 4-cyanoindole

As the starting enzyme, we chose a variant designated *Tm*2F3 (Table 1), which is derived from *T. maritima* TrpB and has seven mutations. We selected this variant because in previous studies it exhibited high activity with other 4-substituted indole substrates.[16] In addition, this variant, like its wild-type progenitor, tolerated high temperatures (up to 75°C), which accelerated the reaction. In the development of *Tm*TrpB-derived variants, we found that activating mutations were distributed throughout the protein sequence without any obvious patterns. One exception was the mutation I184F, which resides in the putative enzyme active-site. Although incorporation of the I184F mutation increased the production of 4-CN-Trp with *Tm*2F3,[16] it was not beneficial for other 4-substituted indoles. Moving forward, we therefore decided to exclude the I184F mutation and instead perform global random mutagenesis on the *Tm*2F3 gene, with the option to revisit I184 at a later stage.

We observed from test reactions that conversion of 4-cyanoindole to 4-CN-Trp was accompanied by an increase in absorption at 350 nm. This spectral shift allowed us to screen the enzyme library rapidly by running reactions in 96-well plates and then monitoring the change in absorption at 350 nm using a plate reader. After screening 1,760 clones, we identified a new variant *Tm*9D8 (E30G and G228S) that appeared to exhibit a 2.5-fold increase in the yield of 4-CN-Trp. Strangely, when we retested *Tm*9D8 in vials, we found that it was no more active than the parent, *Tm*2F3 (Figure 2). We hypothesized that although the plate and vial reactions were ostensibly conducted at the same temperature (75°C), the reaction mixtures in the plate may have actually been at lower temperature, due to the inherent difficulties in heating a 96-well plate uniformly. We therefore retested *Tm*9D8 and *Tm*2F3 at lower temperatures and found that *Tm*9D8 was almost 2-fold better than *Tm*2F3 at 50°C and almost 5-fold better at 37°C. Notably, *Tm*9D8 performed better at 37°C than *Tm*2F3 did at 75°C. This ability to function at lower temperature is not only advantageous for process development, but also creates the possibility of synthesizing 4-CN-Trp *in vivo*.

We previously found that introduction of the mutation I184F into *Tm*2F3 improved the production of 4-CN-Trp.[16] We therefore constructed an 8-variant recombination library in which positions 30, 184, and 228 could either be the wild-type or the mutated residues. Screening this set would reveal if E30G and G228S were both responsible for the first-round improvement and whether I184F was still beneficial in this new variant. We found that the best variant, *Tm*9D8*, indeed retained all three mutations, boosting production of 4-CN-Trp to ~76–78% at both 37 and 50°C (Figure 2). We also tested libraries in which positions 30, 184, and 228 were separately randomized to all twenty canonical amino acids; screening showed that glycine and serine were favored at positions 30 and 228, respectively (see Figures S1 and S2). At position 184, leucine also improved activity compared to the native isoleucine (see Figure S3), but rescreening showed this mutation was not as beneficial as phenylalanine. Thus, we adopted *Tm*9D8* for production of 4-CN-Trp.

Large-scale production of 4-CN-Trp

Although the final variant exhibits a low initial turnover frequency (0.95 ± 0.05 min^{-1} at 37°C) and requires a relatively high catalyst loading (0.1 mol %) to achieve the yields in Figure 2, its expression level is such that enzyme from a 1-L culture can synthesize almost 800 mg of 4-CN-Trp at 55°C (Scheme 2). Since the reaction is performed in aqueous media, most of the product precipitates directly from the reaction mixture and can be purified by simple wash steps. The new variant also retains excellent thermostability (T_{50} ~90°C) (Table 1), which allows it to be prepared as heat-treated lysate, facilitating removal of cell debris, and used in the presence of organic solvents, improving solubility of hydrophobic substrates.

Activity with other substrates

We tested the TrpB variants with other indole analogs to see how the mutations affected specificity (Chart 1). To highlight the improved activity at lower temperature, reactions with *Tm*2F3 and *Tm*9D8 were screened at 50°C, whereas reactions with *Tm*9D8* were screened at 37°C. Although *Tm*2F3 already exhibits good activity (3,250 turnovers at 50°C) with 4-bromoindole (**1**), the activity is improved in the later variants, with *Tm*9D8* performing a similar number of turnovers (3,750), but at lower temperature (37°C). All variants, however, exhibited negligible activity with 4-nitroindole (**2**), suggesting that the active site is highly sensitive to the geometry of substituents at the 4-position.

Previously, *T. maritima* TrpB variants had excelled in reactions with 5-substituted indoles.[16,17] However, 5-nitroindole (**3**) exhibited significantly inferior results with the later variants compared to *Tm*2F3. The later variants, however, exhibited significant improvements with 5,7-disubstituted substrates, providing almost quantitative conversion of **4** to product, even at 37°C. With substrate **5**, the activity is improved almost an order of magnitude from the starting variant.

Discussion

Effect of evolution on TrpB activity

4-Cyanoindole is an especially challenging substrate because its nucleophilicity is attenuated both electronically, due to the electron-withdrawing influence of the cyano group, and sterically, since substituents at the 4-position occlude the site of C–C bond formation. However, the new variant *Tm*9D8* exhibits improved activity with this substrate and even functions well at 37°C. The high expression level of the protein (~40 mg *Tm*9D8* per L culture), the availability of the starting materials, and the convenient reaction setup and product recovery make this an effective method for laboratory preparation of 4-CN-Trp.

Mutations discovered to enhance activity with 4-cyanoindole also improved activity at lower temperatures for other structurally and electronically distinct substrates, such as 4-bromoindole (**1**) and disubstituted indoles **4** and **5**. In all cases, the final variant *Tm*9D8* gave higher yield at 37°C than the starting variant *Tm*2F3 did at 50°C. This general boost in activity at lower temperature is valuable because it not only facilitates process development, but also enables future exploration of substrates that might be unstable in water at elevated temperature. The mutations, however, did not engender general tolerance for 4-substitution,

since the enzyme showed negligible activity with 4-nitroindole (**2**). Surprisingly, activity with 5-nitroindole (**3**) decreased dramatically, even though *Tm*2F3 had originally been evolved for activity with this substrate.[16] These data suggest that the mutations have significantly reconfigured the active site compared to *Tm*2F3, although activity with the native substrate indole remains high for all variants (see Supporting Information).

Role of mutations

The seven mutations in the parent protein *Tm*2F3 were all previously identified in a TrpB homolog from *Pyrococcus furiosus*, which had been evolved through global random mutagenesis and screening to accept 4-nitroindole as a nucleophilic substrate. Although the *P. furiosus* homolog has only 64% sequence identity to *T. maritima* TrpB,[17] we found that these seven mutations were activating in both protein scaffolds. Furthermore, the homologous variants had distinct substrate profiles, with the *T. maritima* variant performing better than *P. furiosus* with 4- or 5-substituted indoles like those shown in Chart 1.

To date, our efforts to solve crystal structures of *T. maritima* TrpB variants have been unsuccessful. We therefore constructed a homology model[18] based on a 1.65-Å crystal structure of *S. typhimurium* TrpB (PDB ID: 4hpx, 58% sequence identity)[19] with the PLP-bound amino-acrylate in the active site. From this model, it is apparent that of the ten mutations in *Tm*9D8*, only two reside in the active site (I184F and G228S), with the other eight scattered throughout the protein structure (Figure 3a). The precise effects of these eight mutations are uncertain, but previous studies suggested that they stabilize the closed state of the enzyme,[16,17] which is known to promote product formation.

The G228S mutation is striking not only because it is an active-site mutation, but also because it is predicted to occur at the beginning of a loop (**G**GGS) that binds the phosphate moiety of the PLP cofactor. To speculate on the role of this mutation, we modeled 4-cyanoindole in the putative binding pose necessary for C–C bond formation (Figure 3b). The sidechains of residues L162, I166, and V188 extend into the active site and thus are expected to influence the positioning of the indole substrate through hydrophobic interactions. The active site also contains E105, a universally conserved residue that interacts with the endocyclic N–H of the native substrate, indole. It is immediately evident that the 4-cyano substituent would point directly toward the phosphate-binding loop and G228 in particular. Thus, the G228S mutation might reorganize the cofactor-binding site to create space for substituents at the 4-position. A survey of 5,738 TrpB homologs revealed that this GGGS sequence is almost universally conserved. This variant therefore serves as an example of how mutations of universally conserved residues can benefit reactions with non-natural substrates.

Conclusions

By applying global random mutagenesis to TrpB from *T. maritima*, we have engineered a variant with improved activity for the production of 4-CN-Trp directly from 4-cyanoindole and serine. Whereas the parent enzyme struggled to form 4-CN-Trp at 75°C, this new variant exhibits considerable activity even at 37°C, enabling production of 4-CN-Trp under mild conditions. The TrpB-catalyzed reactions occur in aqueous media with readily available

starting materials and without the need for protecting groups. Thus, we believe that the TrpB platform will serve as a powerful tool to develop more efficient and direct routes to ncAAs that minimize use of organic solvents.

Experimental Procedures

General experimental methods

Chemicals and reagents were purchased from commercial sources and used without further purification. The proton NMR spectrum was recorded 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, δ 2.50 ppm). The NMR spectrum was 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 CH_3OH as the strong solvent and H_2O (0.1% HCl by weight) as the weak solvent. Liquid chromatography/mass spectrometry (LCMS) was performed on an Agilent 1290 UPLC-LCMS equipped with a C-18 silica column (1.8 μm, 2.1 × 50 mm) using CH_3CN/H_2O (0.1% acetic acid by volume): 5% to 95% CH_3CN over 4 min; 1 mL/min. The optical purity of the products was determined by derivatization with *N*-(5-fluoro-2,4-dinitrophenyl)alanamide (FDNP-alanamide)[20] as described below.

Cloning, expression, and purification of *Tm*TrpB variants

*Tm*TrpB (UNIPROT ID P50909) was previously cloned into pET22(b)+ between the *Nde*I and *Xho*I sites with a 6× C-terminal His-tag.[17] This study used the previously described variant *Tm*2F3[16] as the parent for subsequent evolution. All variants were expressed in BL21(DE3) E. cloni® Express cells. Cultures were started from single colonies in 5 mL Terrific Broth supplemented with 100 μg/mL ampicillin (TB_{amp}) and incubated overnight at 37°C and 230 rpm. For expression, 2.5 mL of overnight culture were used to inoculate 250 mL TB_{amp} in a 1-L flask, which was incubated at 37°C and 250 rpm for three hours to reach OD_{600} 0.6 to 0.8. Cultures were chilled on ice for 20 minutes and expression was induced with a final concentration of 1 mM isopropyl β-D-thiogalactopyranoside (IPTG). Expression proceeded overnight (approximately 20 hours) at 25°C and 250 rpm. Cells were harvested by centrifugation at 5,000*g* for five minutes at 4°C and stored at −20°C.

Thawed cell pellets were resuspended in 9 mL of lysis buffer containing 50 mM potassium phosphate buffer, pH 8.0 (KPi buffer) with 1 mg/mL hen egg white lysozyme (HEWL), 200 μM PLP, 2 mM $MgCl_2$, 0.02 mg/mL DNase I. Pellets were vortexed until completely resuspended, and then cells were lysed with BugBuster® according to manufacturer's recommendations. Lysates were then heat treated at 75°C for 10 minutes. The lysate was centrifuged for 15 minutes at 15,000*g* and 4°C and the supernatant collected. Purification was performed with an AKTA purifier FPLC system (GE Healthcare) and a 1-mL Ni-NTA column. Protein was eluted by applying a linear gradient of 100 mM to 500 mM imidazole in 25 mM KPi buffer and 100 mM NaCl. Fractions containing purified protein were dialyzed into 50 mM KPi buffer, flash frozen in liquid nitrogen, and stored at −80 C. Protein concentrations were determined using the Bio-Rad Quick Start™ Bradford Protein Assay.

Construction of random mutagenesis libraries

Random mutagenesis libraries were generated from the gene encoding *Tm*2F3 by adding 200 to 400 μM $MnCl_2$ to a Taq PCR reaction as reported previously.[16,21] PCR fragments were treated with *Dpn*I for two hours at 37°C and purified by gel extraction. The purified library was then cloned into an empty pET22(b)+ vector *via* Gibson assembly and transformed into BL21(DE3) E. cloni® Express cells.[22]

Forward primer (*Nde*I):
GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATG

Reverse primer (*Xho*I): GCCGGATCTCAGTGGTGGTGGTGGTGGTGCTCGAG

Construction of recombination libraries

Recombination libraries used primers with a degenerate codon to cause a 50/50 amplification of mutant and wild-type residues at a given site (E30G, I184F, G228S) (Table S4). PCR with Phusion® Polymerase (NEB) produced four fragments of the *Tm*2F3 gene (*NdeI* to E30, E30 to I184, I184 to G228, G228 to *XhoI*). Fragments were treated with *Dpn*I for two hours at 37°C and purified by gel extraction. The fragments were assembled by PCR with flanking primers that correspond to the *Nde*I and *Xho*I sites of the pET-22(b)+ vector. The assembled gene was then cloned into an empty pET22(b)+ vector *via* Gibson assembly and transformed into BL21(DE3) E. cloni® Express cells.[22]

Construction of site-saturation libraries

Site saturation libraries were generated using NEB Q5® site directed mutagenesis kit per manufacturer's instructions using *Tm*9D8 as the parent. Primers were designed using NEBaseChanger® software and incorporated the degenerate 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) at the residue of interest (Table S5). Primers were mixed as reported previously.[23] Following PCR, samples were treated with KLD Enzyme Mix for five minutes, and transformed into BL21(DE3) E. cloni® Express cells.

Library expression and screening

BL21(DE3) E. cloni® cells carrying variant plasmids were cultured in 96-well deep-well plates along with parent and negative controls as described previously.[16,21] Overnight cultures were grown by inoculating 300 μL TB_{amp} with a single colony followed by incubation at 37 °C and 250 rpm with 80% humidity. The following day 20 μL of the overnight culture were added to 630 μL TB_{amp} and incubated at 37 °C and 250 rpm with 80% humidity for 3 hours. Cells were then chilled on ice for 20 minutes and induced by addition of IPTG (final concentration 1 mM) followed by incubation at 25°C and 250 rpm overnight (approximately 20 hours). Cells were pelleted by centrifugation at 5,000*g* for 5 minutes, and then decanted and stored at −20°C. Cell plates were thawed and resuspended in 300 μL/well 50 mM KPi buffer with 1 mg/mL HEWL, 200 μM PLP, 2 mM $MgCl_2$, and 0.02 mg/mL DNase. Cells were lysed by a 30-minute incubation at 37°C and heat treatment in a

75°C water bath for 30 minutes (recombination and site saturation) to 180 minutes (random mutagenesis). Lysate was clarified by centrifugation at 5,000*g* for 10 minutes.

Random mutagenesis screen

Reactions were performed in a UV-transparent 96-well assay plate with a total volume of 200 μL/well comprised of 40 μL heat-treated lysate, 5 mM 4-cyanoindole, and 50 mM serine with 5% (v/v) DMSO in 50 mM KPi buffer. Reactions proceeded in a 75°C water bath for 24 hours. Plates were centrifuged briefly to collect condensation and assayed by measuring absorption at 350 nm.

Recombination and site saturation screen

Reactions were performed in 96-well deep-well plates with a total volume of 200 μL/well comprised of 40 μL heat-treated lysate, 5 mM 4-cyanoindole, and 50 mM serine with 5% (v/v) DMSO in 50 mM KPi buffer. Reactions were sealed with Teflon sealing mats and incubated in a 75°C water bath for 24 hours. Plates were briefly chilled on ice and centrifuged to collect condensation. Each well was charged with 500 μL 1-M aq. HCl and 500 μL ethyl acetate. The plate was sealed with a Teflon sealing mat followed by vigorous agitation to dissolve all precipitates and partition the product and substrate between the aqueous and organic phases, respectively. The plates were centrifuged for 2 minutes at 5,000*g* and then 200 μL of the aqueous phase was transferred to a 96-well UV-transparent assay plate. Activity was determined by measuring the absorption at 300 nm.

Calibration for measuring HPLC yield

Using an authentic standard, mixtures of corresponding indole and tryptophan analogs in varied ratios (9:1, 3:1, 1:1, 1:3, and 1:9) were prepared in 1:1 1-M aq. HCl/CH_3CN with a total concentration of 1 mM. Each mixture was prepared in duplicate, then analyzed by LCMS. The ratios of the product and substrate peaks at 254 nm and 280 nm (reference 360 nm, bandwidth 100 nm) were correlated to the actual ratios by a linear relationship (see Figure S4). The authentic standard for 4-cyanotryptophan was obtained from the gram-scale preparation described below. Authentic standards for 4-bromotryptophan, 5-nitrotryptophan, and 5-bromo-7-fluorotyptophan were synthesized as reported previously.[16]

Reactions for Figure 2 and Chart 1

A 2-mL glass HPLC vial was charged with the nucleophilic substrate as a solution in DMSO (10 μL, 400 mM). Next, serine (20 mM final concentration) and purified enzyme (either 4 μM or 20 μM final concentration) were added as a solution in 190 μL of 50 mM KPi buffer. Reactions were heated to 37°C, 50°C, or 75°C for 24 hours. The reaction was then diluted with 800 μL of 1:1 1-M aq. HCl/CH_3CN and vortexed thoroughly. Finally, the reaction mixture was centrifuged at >20,000*g* for 10 minutes, and the supernatant was analyzed by HPLC. The identity of the product was confirmed by comparison to an authentic standard. The yield was determined by comparing the integrations of the HPLC peaks corresponding to product and starting material (see Supporting Information for more details). Experiments were conducted at least in duplicate.

Approximation of initial turnover frequency

Reactions with 4-cyanoindole were set up according to the procedure described above for Figure 2. The reactions were worked up after 1 hour and analyzed by HPLC. The integrations of 254-nm absorption peaks corresponding to product and starting material were used to calculate product formation. The reactions were conducted in triplicate. See Table S2 for full data.

Gram-scale preparation of 4-cyanotryptophan

Heat-treated lysate was prepared following the protocol described above for preparing the enzymes for purification. In a 1-L Erlenmeyer flask, 4-cyanoindole (1.0 g, 7.0 mmol) and serine (810 mg, 7.7 mmol) were suspended in DMSO (17.5 mL) and 50 mM KPi buffer (250 mL). Heat-treated lysate from four 250-mL expression cultures was added, then the reaction mixture was heated in a water bath at 55°C. After 72 hours, the reaction mixture was cooled on ice for 90 minutes. The precipitate was collected by filtration, washed twice with ethyl acetate and twice with water, then dried *in vacuo* to afford 4-CN-Trp as an off-white solid (797 mg, 49% yield).

The ^{1}H NMR spectrum was taken in a mixture of DMSO-d_6 and 20% DCl/D_2O and referenced to the residual DMSO peak (2.50 ppm).[1] **H NMR** (400 MHz, DMSO-d_6) δ 7.57 (dd, J = 8.2, 1.0 Hz, 1H), 7.32 (s, 1H), 7.32–7.29 (m, 1H), 7.09–7.03 (m, 1H), 4.00 (dd, J = 5.8, 2.9 Hz, 1H), 3.30 (**AB**X, J_{AX} = 8.7 Hz, J_{BX} = 6.2 Hz, J_{AB} = 15.2 Hz, ν_{AB} = 85.2 Hz, 2H). The data were in concordance with the previous literature.[16]

Determination of T_{50} values

A mastermix of 1 μM purified enzyme was prepared in 50 mM KPi buffer and 95 μL added to 12 PCR tubes. Ten test samples were incubated in a thermocycler for 60 minutes with a temperature gradient from 79°C to 99°C, while the two control samples were incubated at room temperature. All tubes were centrifuged for three minutes to pellet precipitated enzyme, and then 75 μL of the supernatant was transferred from each tube to a UV-transparent 96-well assay plate. Enzyme activity was determined by adding an additional 75 μL of 50 mM KPi buffer containing 1 mM indole and 1 mM serine to each well. Reactions were incubated for 45 min at 50°C (*Tm*9D8 and *Tm*9D8*) or 75°C (*Tm*2F3) then briefly centrifuged to collect condensation. Activity was determined by measuring the absorption at 290 nm. Activity was correlated to incubation temperature, and the half-inactivation temperatures (T_{50}) were determined. Measurements were conducted in triplicate.

Determination of Optical Purity

The optical purity of products was estimated by derivatization with FDNP-alanamide. In a 2-mL vial, a 200 uL reaction was carried out as described above. After 24 hours of incubation at 37°C, 100 μL of 1 M aq. $NaHCO_3$ was added to the reaction, and 125 μL of the reaction mixture (up to 1.1 μmol product) was transferred into two 2-mL vials. FDNP-alanamide (33 μL of a 33-mM solution in acetone, 1.1 μmol) was added to each vial, followed by incubation at 37°C and 230 rpm. After two hours, the reaction mixture was cooled to room temperature and then diluted with 1:1 CH_3CN/1-M aq. HCl (600 μL). The resulting solution

was analyzed directly by LCMS. Each amino acid was derivatized with both racemic and enantiopure FDNP-alanamide for comparison. Absolute stereochemistry was inferred by analogy to L-tryptophan. All products were >99% ee.

Structural modeling

A structure of TrpS from *S. enterica* has been reported (PDB ID: 4HPX), in which the β-subunit (*Se*TrpB) is in the closed state and contains benzimidazole and the Ser-derived amino-acrylate in the active site.[19] This structure served as a template for a homology model of wild-type TrpB from *T. maritima* (*Tm*TrpB, 58% sequence identity), which was constructed using the Swiss-Model program.[18] The homology model of *Tm*TrpB was aligned with the authentic structure of *Se*TrpB using PyMOL, which allowed the amino-acrylate and benzimidazole to be mapped directly into the homology model. Finally, the benzimidazole was replaced with a simulated structure of 4-cyanoindole, such that the indole moiety of 4-cyanoindole mapped onto the structure of benzimidazole.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was funded by the Jacobs Institute for Molecular Engineering for Medicine (JIMEM) and the Rothenberg Innovation Initiative (RI^2) at Caltech. C.E.B. was supported by a postdoctoral fellowship from the Resnick Sustainability Institute, D.K.R. was supported by a Ruth Kirschstein NIH Postdoctoral Fellowship (F32GM117635), and M.S. was supported by a postdoctoral fellowship from the German Academic Exchange Service (DAAD).

References

1Smith DT, , Delost MD, , Qureshi H, , Njardarson JT. Top 200 Pharmaceutical Products by Retail Sales in 2016 http://njardarson.lab.arizona.edu/sites/njardarson.lab.arizona.edu/files/2016Top200PharmaceuticalRetailSalesPosterLowResV2.pdf.

2Ager DJ. Synthesis of Unnatural/Nonproteinogenic α-Amino Acids. In: Hughes AB, editorAmino Acids, Peptides and Proteins in Organic Chemistry: Origins and Synthesis of Amino Acid Vol. 1. Wiley Online Library; 2010 495526

3Brittain WDG, Cobb SL. Negishi Cross-Couplings in the Synthesis of Amino Acids. Org Biomol Chem. 2017; 16(1):10–20. [PubMed: 29199315]

4Arnold LD, Kalantar TH, Vederas JC. Conversion of Serine to Stereochemically Pure β-Substituted α-Amino Acids via β-Lactones. J Am Chem Soc. 1985; 107:7105–7109.

5Arnold LD, Drover J, Vederas JC. Conversion of Serine β-Lactones to Chiral α-Amino Acids by Copper-Containing Organolithium and Organomagnesium Reagents. J Am Chem Soc. 1987; 109(15):4649–4659.

6Arnold LD, May RG, Vederas JC. Synthesis of Optically Pure α-Amino Acids via Salts of α-Amino-β-Propiolactone. J Am Chem Soc. 1988; 110(7):2237–2241.

7Tanner D. Chiral Aziridines—Their Synthesis and Use in Stereoselective Transformations. Angew Chem Int Ed. 1994; 33(6):599–619.

8Ishikawa T. Aziridine-2-Carboxylates: Preparation, Nucleophilic Ring Opening, and Ring Expansion. Heterocycles. 2012; 85(12):2837–2877.

9Corr MJ, Smith D, Goss R. One-Pot Access to l-5, 6-Dihalotryptophans and l-Alknyltryptophans Using Tryptophan Synthase. Tetrahedron. 2016; 72(46):7306–7310.

10Smith DRM, Willemse T, Gkotsi DS, Schepens W, Maes BUW, Ballet S, Goss RJM. The First One-Pot Synthesis of l-7-Iodotryptophan from 7-Iodoindole and Serine, and an Improved Synthesis of Other l-7-Halotryptophans. Org Lett. 2014; 16(10):2622–2625. [PubMed: 24805161]

11Perni S, Hackett L, Goss RJ, Simmons MJ, Overton TW. Optimisation of Engineered Escherichia Coli Biofilms for Enzymatic Biosynthesis of l-Halotryptophans. AMB Express. 2013; 3(1):66. [PubMed: 24188712]

12Winn M, Roy AD, Grüschow S, Parameswaran RS, Goss RJM. A Convenient One-Step Synthesis of l-Aminotryptophans and Improved Synthesis of 5-Fluorotryptophan. Bioorg Med Chem Lett. 2008; 18(16):4508–4510. [PubMed: 18667314]

13Goss RJM, Newill PLA. A Convenient Enzymatic Synthesis of l-Halotryptophans. Chem Commun. 2006; 47:4924–4925.

14Lee M, Phillips RS. Enzymatic Synthesis of Chloro-l-Tryptophans. Bioorg Med Chem Lett. 1992; 2(12):1563–1564.

15Hilaire MR, Ahmed IA, Lin C-W, Jo H, DeGrado WF, Gai F. Blue Fluorescent Amino Acid for Biological Spectroscopy and Microscopy. Proc Nat Acad Sci. 2017; 114(23):6005–6009. [PubMed: 28533371]

16Romney DK, Murciano-Calles J, Wehrmüller JE, Arnold FH. Unlocking Reactivity of TrpB: A General Biocatalytic Platform for Synthesis of Tryptophan Analogs. J Am Chem Soc. 2017; 139(31):10769–10776. [PubMed: 28708383]

17Murciano-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(38):11577–11581.

18SWISS-MODEL available online at https://swissmodel.expasy.org.

19Niks D, Hilario E, Dierkers A, Ngo H, Borchardt D, Neubauer TJ, Fan L, Mueller LJ, Dunn MF. Allostery and Substrate Channeling in the Tryptophan Synthase Bienzyme Complex: Evidence for Two Subunit Conformations and Four Quaternary States. Biochemistry. 2013; 52(37):6396–6411. [PubMed: 23952479]

20Brückner H, Gah C. High-Performance Liquid Chromatographic Separation of DL-Amino Acids Derivatized with Chiral Variants of Sanger's Reagent. Journal of Chromatography A. 1991

21Buller 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(47):14599–14604. [PubMed: 26553994]

22Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO. Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases. Nature Methods. 2009; 6(5):343–345. [PubMed: 19363495]

23Kille S, Acevedo-Rocha CG, Parra LP, Zhang ZG, Opperman DJ, Reetz MT, Acevedo JP. Reducing Codon Redundancy and Screening Effort of Combinatorial Protein Libraries Created by Saturation Mutagenesis. ACS Synth Biol. 2013; 2(2):83–92. [PubMed: 23656371]

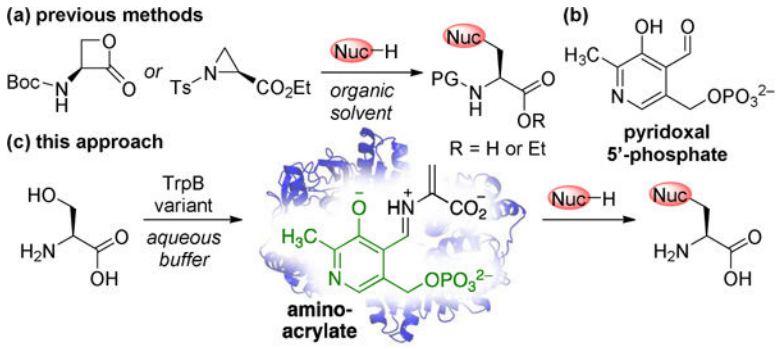


Figure 1.

Amino acid synthesis by nucleophilic substitution at the β-position. (a) Approach using preformed lactone or aziridine. (b) Cofactor used by TrpB enzymes. (c) Alternative approach in which an enzyme forms an amino-acrylate *in situ* from stable precursors like serine. Boc, *tert*-butoxycarbonyl; Ts, 4-toluenesulfonyl; PG, protecting group.

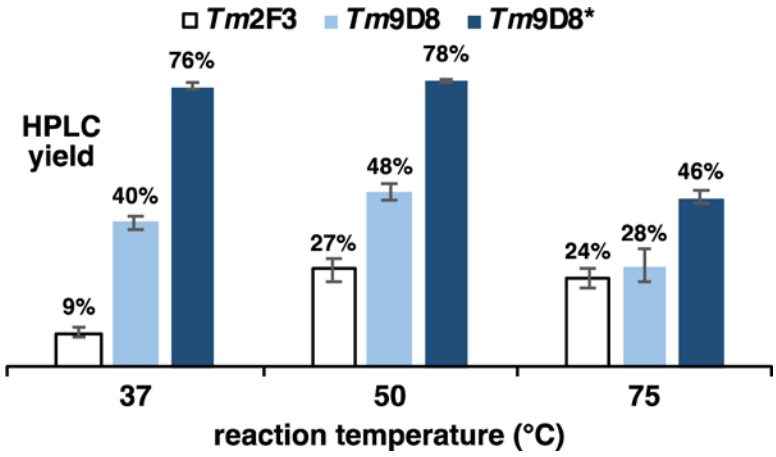


Figure 2.

Production of 4-CN-Trp at different temperatures from equimolar 4-cyanoindole and serine (maximum of 1000 turnovers). Yields are averages of two replicates. Full data are reported in Table S1.

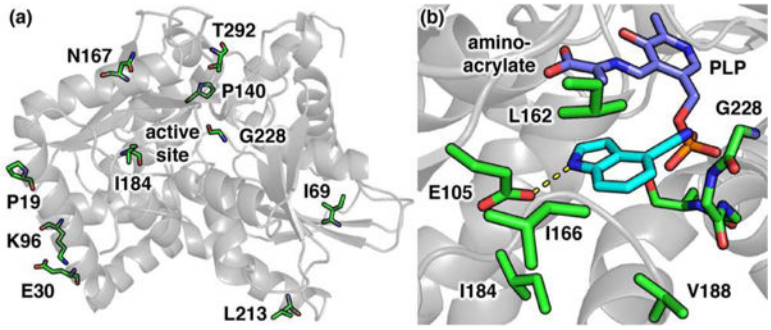
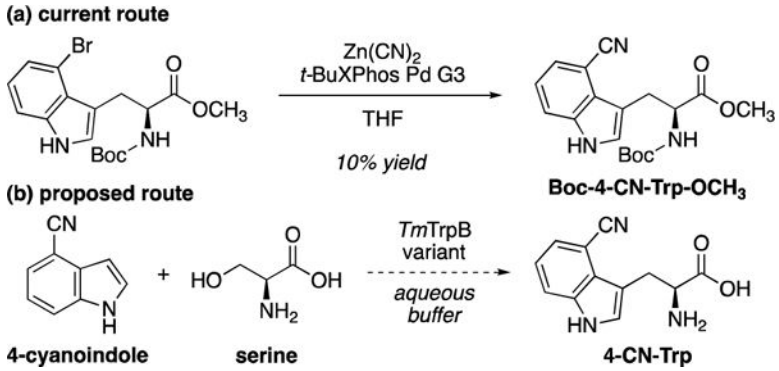
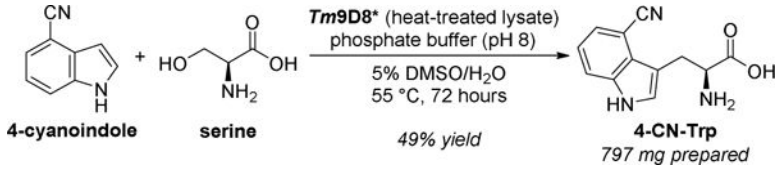


Figure 3.

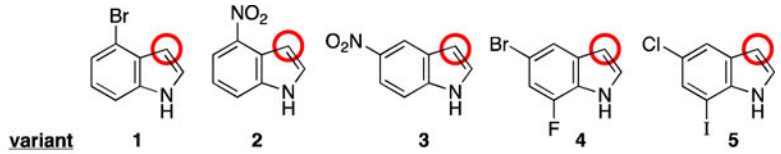
Homology model of the *T. maritima* TrpB showing (a) the whole protein structure with the mutated sites and (b) the active site with the PLP-bound amino-acrylate, 4-cyanoindole in a reactive binding pose, and residues predicted to interact with 4-cyanoindole highlighted.



Scheme 1.
Direct synthesis of 4-cyanotryptophan



Scheme 2.
Enzymatic preparation of 4-CN-Trp



variant	1	2	3	4	5
*Tm*2F3 :[b]	65%	2.6%	21%	34%	2.3%
*Tm*9D8 :[b]	71%	2.5%	5.3%	94%	20%
*Tm*9D8* :[c]	75%	ND	7.5%	92%	18%

Chart 1. HPLC yield of Trp analogs with TrpB variants[a]

[a]Reactions had 0.02 mol % catalyst loading (maximum 5000 turnovers) and 1 equiv serine relative to indole substrates.

[b]Reactions run at 50°C.

[c]Reactions run at 37°C. Red circles indicated site of C–C bond formation. Yields are averages of two replicates. Full data are reported in Table S3.

Table 1

Summary of *T. maritima* TrpB variants

Designation	Mutations	T_{50}(° C)[a]
*Tm*2F3	P19G, I69V, K96L, P140L, N167D, L213P, T292S	91.5 ± 0.8
*Tm*9D8	*Tm*2F3 + E30G, G228S	88.2 ± 0.7
*Tm*9D8*	*Tm*9D8 + I184F	92.7 ± 0.2

[a]Measurements conducted in triplicate.