JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Communication
pubs.acs.org/JACS

Synthesis of β-Branched Tryptophan Analogues Using an Engineered Subunit of Tryptophan Synthase

Michael Herger,[†] Paul van Roye,[†] David K. Romney, Sabine Brinkmann-Chen, Andrew R. Buller,* 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

Supporting Information

ABSTRACT: We report that L-threonine may substitute for L-serine in the β-substitution reaction of an engineered subunit of tryptophan synthase from *Pyrococcus furiosus*, yielding (2*S*,3*S*)-β-methyltryptophan (β-MeTrp) in a single step. The trace activity of the wild-type β-subunit on this substrate was enhanced more than 1000-fold by directed evolution. Structural and spectroscopic data indicate that this increase is correlated with stabilization of the electrophilic aminoacrylate intermediate. The engineered biocatalyst also reacts with a variety of indole analogues and thiophenol for diastereoselective C−C, C−N, and C−S bond-forming reactions. This new activity circumvents the 3-enzyme pathway that produces β-MeTrp in nature and offers a simple and expandable route to preparing derivatives of this valuable building block.

Amino acids possess a rich array of functionality that has made them ideal building blocks for complex molecular structures. In addition to the small set of amino acids that form the basis of naturally occurring proteins, nature has produced a myriad of noncanonical amino acids (ncAAs) that are non-proteinogenic but are nonetheless prevalent in other classes of natural products.[1] Chemists have applied both natural and artificial ncAAs in the development of chemical probes and asymmetric catalysts as well as industrial products such as agrochemicals and small-molecule pharmaceuticals.[2] However, the efficient synthesis of ncAAs is notoriously challenging because it requires methods that can form the products in high enantioenrichment while tolerating their dense functionality.[3] The difficulty is compounded with β-branched amino acids, which necessitate both enantio- and diastereoselective synthetic methods.[4] As a result, access to these compounds often requires lengthy synthetic routes with the use of numerous protecting groups. Biocatalytic routes to ncAAs present an attractive alternative because they often feature exquisite stereoselectivity and do not require protecting groups. As such, there is continued demand for enzymes that can enable production of diverse ncAAs.

A striking example of an enzyme with a broad substrate scope is tryptophan synthase (TrpS), a pyridoxal phosphate (PLP)-dependent enzyme that catalyzes the condensation of indole (**1**) and L-serine (Ser, **2**) to form L-tryptophan (Trp, **3**, Figure 1a).[5] It is well-precedented that TrpS can accept nucleophiles other than

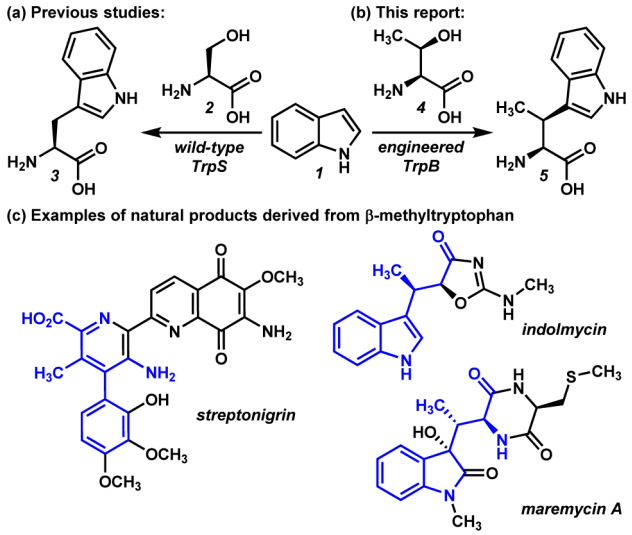


Figure 1. (a) Condensation of **1** and **2** using wild-type TrpS from different species. (b) Proposed reaction with **4** using engineered variants of TrpB. (c) Examples of natural products with β-MeTrp as a biosynthetic intermediate.

1,[6] but the only electrophiles that are known to replace **2** are amino acids such as L-cysteine and β-chloro-L-alanine, which lead to the same products as **2**, albeit with diminished efficiency.[7] This desirable reaction is catalyzed by the β-subunit (TrpB) of the αββα heteromeric enzyme complex, but further expansion of the substrate scope through protein engineering is complicated by the presence of the α-subunit (TrpA), without which the catalytic efficiency of TrpB is significantly attenuated.[5] Recently, we used directed evolution of TrpB from the hyperthermophile *Pyrococcus furiosus* to create a variant, *Pf*TrpB^{0B2}, that maintains high catalytic efficiency outside of the native heteroenzyme complex.[8] We explored the substrate scope of this stand-alone enzyme with the hypothesis that any promising activity could be optimized through further evolution for efficient production of ncAAs.

We tested analogues of **2** with *Pf*TrpB^{0B2}, with the goal of expanding the scope of substrates that can serve as electrophiles in this reaction manifold (Figure S1). In the course of this exploration, we were pleased to discover that L-threonine (Thr,

Received: May 10, 2016
Published: June 29, 2016

DOI: 10.1021/jacs.6b04836
J. Am. Chem. Soc. 2016, 138, 8388−8391

4) could substitute for Ser to yield a single product that we determined to be (2*S*,3*S*)-*β*-methyltryptophan (*β*-MeTrp, **5**, Figure 1b), corresponding to substitution with retention of configuration (Figure 2).[9]

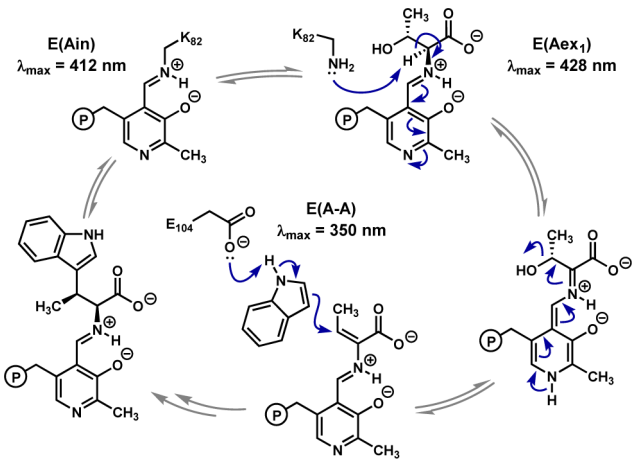


Figure 2. Proposed catalytic cycle showing the expected UV−vis absorbance peaks for key intermediates. Detailed discussion of the TrpS mechanism is given in ref 5.

To our knowledge, this reactivity is unprecedented for TrpB. While an analogous reaction has been reported for TrpS from *Salmonella typhimurium*, this required a substantially stronger nucleophile (benzyl mercaptan) to obtain even 1% of the activity relative to Ser and the stereochemistry of the product was not determined.[10] Compound **5** is valuable because it is an intermediate in the biosynthetic pathways to several natural products such as maremycin and the antitumor agent streptonigrin (Figure 1c).[9,11] The chemical synthesis of **5** has excellent selectivity but requires many chemical manipulations.[12] Similarly, the biosynthesis of **5**, which uses three enzymes in a four-step pathway from **3** using (*S*)-adenosylmethionine as the methyl source,[9,11] is unsuitable for large-scale production of **5** or its derivatives. The reaction reported here represents a significantly more efficient route to this valuable intermediate. We therefore sought to optimize the enzyme for this novel reaction using directed evolution.

To identify the best parent enzyme for subsequent engineering, we first measured production of **5** using purified enzymes from each generation in the evolution of *Pf*TrpB^{0B2}. *Pf*TrpB^{WT} was an inefficient catalyst, performing just 66 turnovers in 24 h (Table S1). This activity was enhanced approximately 6-fold with the single amino acid substitution T292S and further to ∼660 TTN with the enzyme *Pf*TrpB^{4D11}, which incorporates four additional mutations (E17G, I68V, F274S, and T321A). *Pf*TrpB^{0B2} exhibited a decrease in activity with Thr (Table S1) compared to *Pf*TrpB^{4D11}, despite having higher activity with Ser.[8] We therefore selected *Pf*TrpB^{4D11} as the parent for directed evolution.

Previous studies have shown that the catalytic activity of TrpB is governed by open−close transitions of the communication (COMM) domain,[5,8] and mutations that alter this dynamic in the reactivated subunits are distributed throughout the protein structure.[8] As such, we elected to apply random mutagenesis for the first round of evolution for activity with Thr. To determine an optimal mutational load, we measured the retention of function of 352 clones from libraries with different mutation rates (Figure S2). In this process, we found six missense mutations in five clones with at least a 2-fold increase in V_{max}, obviating the need for screening a more expansive library. The most active clone from this library, *Pf*TrpB^{4G1}, had ∼3.8-fold increased activity and contained the single mutation F95L, which abuts the COMM domain.

Notably, another activated variant contained two mutations, I16V and Q89L, of which the former is adjacent to the E17G mutation already present in the parent enzyme. These six mutations were randomly recombined and the resulting variants screened for activity, which resulted in the final variant *Pf*TrpB^{2B9}. Three mutations, F95L, I16V, and V384A, were retained in this enzyme, which is eight mutations away from wild type and has at least a 6000-fold boost in productivity (Figure 3). This increase is in part due to an elevated expression level of soluble protein, from ∼80 mg *Pf*TrpB^{WT} per L medium to ∼350 mg of *Pf*TrpB^{2B9}.

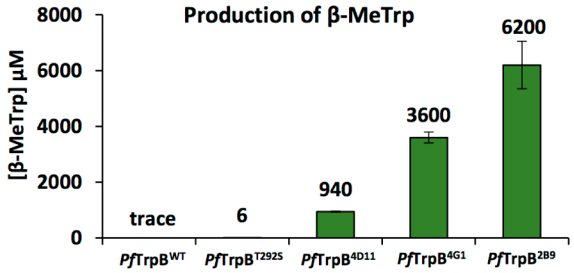


Figure 3. Productivity of **5** with heat-treated lysates of *Escherichia coli* expressing *Pf*TrpB enzymes. Productivity is defined as the amount of **5** formed after 90 min at 75 °C from a starting reaction mixture containing 20 mM of **1** and **4**.

To assess the basis of this new activity, we used UV−vis spectroscopy, which reports on the steady-state distribution of the PLP-bound intermediates in the catalytic cycle (Figure 2). Addition of Ser to *Pf*TrpB^{WT} shifts λ_{max} from 412 to 428 nm, corresponding to transimination of the lysine-bound internal aldimine, E(Ain), to a Ser-bound external aldimine, E(Aex$_1$).[8] Unexpectedly, addition of 20 mM Thr to *Pf*TrpB^{WT} did not result in any spectral shift indicative of Thr binding. However, this experiment cannot distinguish between Thr simply not binding or Thr binding noncovalently but not favoring formation of E(Aex$_1$). We soaked Thr into crystals of *Pf*TrpB^{WT}, and diffraction yielded a 1.54 Å structure that shows Thr bound noncovalently (Figure 4a, Table S2). The Thr hydroxyl forms a 2.7 Å hydrogen bond with the side chain of Asp300, an

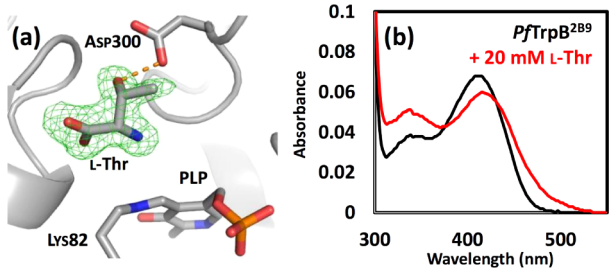


Figure 4. (a) Cocrystal structure of *Pf*TrpB^{WT} with **4** shows that Thr binds noncovalently. Mesh corresponds to F_o-F_c omit map displayed at 2.5σ. (b) UV−vis spectrum of *Pf*TrpB^{2B9} (black) shows a peak at 412 nm, corresponding to E(Ain). Addition of 20 mM **4** (red) has a mixed population of E(Aex$_1$) at 428 nm and E(A-A) at ∼350 nm, indicating covalent substrate binding.

DOI: 10.1021/jacs.6b04836
J. Am. Chem. Soc. 2016, 138, 8388−8391

interaction that is also observed in the Ser E(Aex_1). Modeling of a hypothetical Thr E(Aex_1) that maintains the Asp300 hydrogen bond reveals a strong steric clash between the Thr methyl group and the backbone carbonyl of Gly298 (Figure S3), accounting for the lack of E(Aex_1) observed by UV−vis spectroscopy.

UV−vis spectroscopy was used to probe the basis for enhanced activity through directed evolution. Addition of 20 mM Thr to the engineered proteins from each generation resulted in a clear trend (Figures 4b and S4): the 412 nm peak decreased while a new absorbance band at 350 nm with a broad shoulder out to 550 nm appeared, consistent with accumulation of the electrophilic aminoacrylate species E(A-A). A similar phenomenon was observed during directed evolution for independent TrpB function;[8] however, comparison of the Ser- and Thr-bound *Pf*TrpB^{2B9} spectra indicates that E(A-A) with Thr is much less stable (Figure S5).

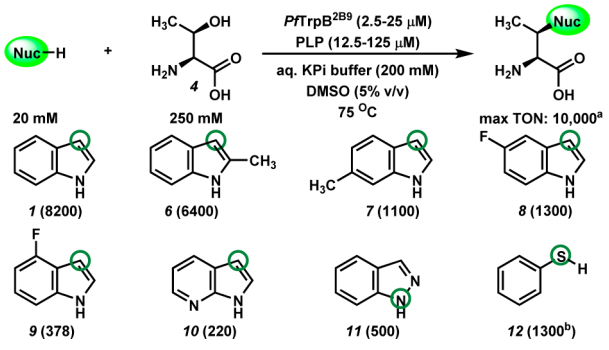
The engineered *Pf*TrpB^{2B9} enzyme has several desirable features as a catalyst for β-MeTrp production. It is robustly expressed in *E. coli* and can be prepared in a moderately pure form as a heat-treated lysate (Figure S6), and its thermal stability permits high reaction temperatures, routinely up to 75 °C, which greatly increases the solubility of hydrophobic substrates. However, in reactions with 1 equiv each of indole and Thr, we observed only 44% conversion to product, corresponding to 2220 turnovers (Figure S7). A critical clue explaining this low conversion came from UV−vis spectroscopy, which revealed that addition of Thr to *Pf*TrpB^{2B9} results in a time-dependent increase in absorbance at 320 nm, while the remainder of the spectrum remains constant (Figure S8). We attribute this to α-ketobutyrate production from the well-described deamination reaction that results when a nucleophile does not add into C-β, and E(Ain) is reformed through transimination.[7] The precise mechanism of deamination is unknown, but the net effect is an abortive reaction wherein the aminoacrylate hydrolyzes to form α-ketobutyrate and ammonium. We therefore added additional equivalents of Thr to the β-substitution reaction with indole (Figure S7), which enabled >99% conversion of indole to β-MeTrp and up to 8200 total turnovers to the desired product (Chart 1) with >99% ee and de.

With these reaction conditions, we next characterized the substrate scope of the reaction with Thr. Previously, we screened a small panel of indole-like nucleophiles for reaction with *Pf*TrpB^{0B2} and Ser.[8] We expanded on this panel of nucleophiles and tested for activity with *Pf*TrpB^{2B9} using a 10-fold molar excess of Thr to nucleophile (Chart 1). Reactions were run to intermediate yield to determine the TTN with a given nucleophile. The identity of the products was established with a separate preparative-scale reaction using 100 mM nucleophile, 1.0 M Thr, and 0.02−0.13 mol % of *Pf*TrpB^{2B9} catalyst. We observed good reactivity with the 2-methyl and 6-methylindole substrates, but decreased TTNs compared to indole demonstrate that the active site is sensitive to steric perturbations. To probe the role of electronic effects on the C−C bond-forming step, we tested activity with 4-fluoro- and 5-fluoroindole, which are more closely isosteric with indole but have decreased electron density in the π-system. We observed product formation with each substrate and 3.4-fold lower TTN with the 4-fluoro substrate, which is more electron-deficient at C-3 than the 5-fluoro substituent. Increased steric constraints of the active site were again clear, as we did not observe any activity with the 5-chloro-, 5-bromo-, and 6-hydroxyindoles, which do undergo reaction with Ser.[8]

A productive reaction with 220 turnovers was observed with 7-azaindole, which is a substantially weaker nucleophile than indole. Interestingly, a second product was detected in the reaction with 7-azaindole that is consistent with N-alkylation. This regioselectivity is well-known with indazole, which we found reacts exclusively to form an N-alkylated product. Surprisingly, we did not observe product formation with indoline, which is a stronger nucleophile than both indazole and indole and reacts significantly faster in the reaction with Ser. Future studies will be needed to ascertain the origin of this effect. Lastly, we tested for S-alkylation using thiophenol and measured 1300 turnovers.

In conclusion, we have discovered a new, non-natural enzymatic route for the production of (2*S*,3*S*)-β-MeTrp. This activity lies in the β-subunit of TrpS, which was subsequently engineered for increased activity with Thr as the amino acid donor. Development of the resultant catalyst, *Pf*TrpB^{2B9}, was greatly facilitated by previous efforts to replace the native two-enzyme system with a single stand-alone catalyst.[8] The engineered enzymes have high thermal stability and expression in *E. coli*, enhancing their utility as practical catalysts. This enzymatic route to β-MeTrp is dramatically shorter than previous synthetic routes and also the native 3-enzyme pathway to this natural metabolite.[9,11] This study highlights the ability of engineered biocatalysts to efficiently produce complicated molecules from simple precursors and offers a simple and expandable route for the production of β-methyl ncAA analogues.

Chart 1. Production of β-MeTrp Analogues



[a]Total turnover numbers are given in parentheses. Reactions conducted in triplicate. Circles indicate the site of alkylation. [b]Single reaction conducted at 25 °C with 12.5 mM DTT. See Supporting Information for further detail.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b04836.

Additional figures, experimental details, and characterization (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*abuller@caltech.edu

*frances@cheme.caltech.edu

Author Contributions

†M.H. and P.vR. contributed equally.

Notes

The authors declare no competing financial interest.

DOI: 10.1021/jacs.6b04836
J. Am. Chem. Soc. 2016, 138, 8388−8391

■ ACKNOWLEDGMENTS

The authors thank Hans Renata, Javier Murciano-Calles, and Chris Prier for helpful discussion and comments on the manuscript. We thank Dr. Jens Kaiser of the Caltech Molecular Observatory, which is supported by the Gordon and Betty Moore Foundation, the Beckman Institute, and the Sanofi-Aventis Bioengineering Research Program at Caltech. We thank Dr. Scott Virgil and the Center for Catalysis and Chemical Synthesis, and Dr. Mona Shahgholi and Naseem Torian from the Caltech Mass Spectrometry Laboratory. This work was funded through the Jacobs Institute for Molecular Engineering for Medicine and Ruth Kirschstein NIH Postdoctoral Fellowships F32G110851 (to A.R.B.) and F32GM117635 (to D.K.R.).

■ REFERENCES

(1) Alkhalaf, L. M.; Ryan, K. S. *Chem. Biol.* **2015**, *22* (3), 317−328.
(2) (a) Wennemers, H. *Chem. Commun.* **2011**, *47* (44), 12036−12041. (b) Davie, E. A. C.; Mennen, S. M.; Xu, Y. J.; Miller, S. J. *Chem. Rev.* **2007**, *107* (12), 5759−5812. (c) Kiick, K. L.; Saxon, E.; Tirrell, D. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (1), 19−24. (d) Fowler, V. G.; Boucher, H. W.; Corey, G. R.; Abrutyn, E.; Karchmer, A. W.; Rupp, M. E.; Levine, D. P.; Chambers, H. F.; Tally, F. P.; Vigliani, G. A.; Cabell, C. H.; Link, A. S.; DeMeyer, I.; Filler, S. G.; Zervos, M.; Cook, P.; Parsonnet, J.; Bernstein, J. M.; Price, C. S.; Forrest, G. N.; Fatkenheuer, G.; Gareca, M.; Rehm, S. J.; Brodt, H. R.; Tice, A.; Cosgrove, S. E. *N. Engl. J. Med.* **2006**, *355* (7), 653−665.
(3) Kieffer, M. E.; Repka, L. M.; Reisman, S. E. *J. Am. Chem. Soc.* **2012**, *134* (11), 5131−5137.
(4) Sawai, Y.; Mizuno, M.; Ito, T.; Kawakami, J.; Yamano, M. *Tetrahedron* **2009**, *65* (34), 7122−7128.
(5) Dunn, M. F. *Arch. Biochem. Biophys.* **2012**, *519* (2), 154−166.
(6) (a) Phillips, R. S. *Tetrahedron: Asymmetry* **2004**, *15* (18), 2787−2792. (b) Smith, D. R. M.; Willemse, T.; Gkotsi, D. S.; Schepens, W.; Maes, B. U. W.; Ballet, S.; Goss, R. J. M. *Org. Lett.* **2014**, *16* (10), 2622−2625.
(7) (a) Crawford, I. P.; Ito, J. *Proc. Natl. Acad. Sci. U. S. A.* **1964**, *51* (3), 390−397. (b) Ahmed, A. S.; Miles, E. D. *J. Biol. Chem.* **1994**, *269* (23), 16486−16492.
(8) Buller, A. R.; Brinkmann-Chen, S.; Romney, D. K.; Herger, M.; Murciano-Calles, J.; Arnold, F. H. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (47), 14599−14604.
(9) Zou, Y.; Fang, Q.; Yin, H.; Liang, Z.; Kong, D.; Bai, L.; Deng, Z.; Lin, S. *Angew. Chem., Int. Ed.* **2013**, *52* (49), 12951−12955.
(10) Esaki, N.; Tanaka, H.; Miles, E. W.; Soda, K. *Agric. Biol. Chem.* **1983**, *47* (12), 2861−2864.
(11) Kong, D.; Zou, Y.; Zhang, Z.; Xu, F.; Brock, N. L.; Zhang, L.; Deng, Z.; Lin, S. *Sci. Rep.* **2016**, *6*, 20273−20280.
(12) (a) Liu, Y.; Zhang, L.; Jia, Y. *Tetrahedron Lett.* **2012**, *53* (6), 684−687. (b) Han, G.; Lewis, A.; Hruby, V. J. *Tetrahedron Lett.* **2001**, *42*, 4601−4603.

DOI: 10.1021/jacs.6b04836
J. Am. Chem. Soc. 2016, 138, 8388−8391