Directed Evolution of an Allosteric Tryptophan Synthase to Create a Platform for Synthesis of Noncanonical Amino Acids

1

Javier Murciano-Calles, Andrew R. Buller, and Frances H. Arnold

Abstract

Tryptophan and its derivatives are important natural products and have many biochemical and synthetic applications. However, the more elaborate these derivatives are, the more complex the synthesis becomes. In this chapter, we summarize the development of an engineered enzymatic platform for synthesis of diverse tryptophan analogs. This endeavor utilizes the tryptophan synthase (TrpS) enzyme, an $\alpha_2\beta_2$ heterodimeric protein complex that catalyzes the last two steps in the biosynthetic pathway of tryptophan. Although the synthetically useful reaction (indole + Ser = Trp) takes place in the β-subunit (TrpB), the exquisite allosteric regulation of this enzyme impedes the use of isolated TrpB due to its dramatically decreased activity in the absence of the α-subunit (TrpA). This chapter discusses our efforts to engineer TrpB to serve as a general platform for the synthesis of noncanonical amino acids. We used directed evolution to enhance the activity of TrpB from *Pyrococcus furiosus* (*Pf*TrpB), so that it can act as a stand-alone biocatalyst. Remarkably, we found that mutational activation mimics the allosteric activation induced by binding of TrpA. Toward our goal of expanding the substrate scope of this reaction, we activated other homologs with the same mutations discovered for *Pf*TrpB. We found improved catalysts for the synthesis of 5-substituted tryptophans, an important biological motif. Finally, we performed directed evolution of TrpB for synthesis of β-branched amino acids, a group of products whose chemical syntheses are particularly challenging.

J. Murciano-Calles • A.R. Buller • F.H. Arnold (✉)
Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA, USA
e-mail: frances@cheme.caltech.edu

M. Alcalde (ed.), *Directed Enzyme Evolution: Advances and Applications*,
DOI 10.1007/978-3-319-50413-1_1

1.1 Introduction

In addition to being one of the standard 20 proteinogenic α-amino acids, tryptophan (Trp) is the immediate precursor of important biomolecules such as the neurotransmitter serotonin [1], vitamin B3 [2], and auxin phytohormones [3, 4]. In biosynthetic pathways of more complex natural products, modified tryptophan is frequently the core of the final biomolecule [5–9]. Tryptophan derivatives have also been used in chemical biology for a variety of applications [10]. These noncanonical amino acids (ncAAs) also serve as intermediates in the production of pharmaceuticals by chemical synthesis [11]. It is therefore important to develop efficient routes to preparing these compounds.

Tryptophan synthase (TrpS) has been used to make a wide variety of Trp analogs. TrpS catalyzes the last steps in the de novo pathway for Trp in all three domains of life. This enzyme synthesizes Trp from 3-indole-D-glycerol phosphate (IGP) and L-serine (Ser) in two steps that take place in two separate subunits of the heterodimeric enzyme [12]. TrpS has been used for the synthesis of many modified tryptophans by reaction of Ser and an appropriate nucleophile, typically a substituted indole [13–20]. However, these reactions frequently proceed with low yields (below 50%). Researchers have tried to expand the substrate scope to nitrogen nucleophiles through protein engineering, but those efforts concomitantly boosted an abortive side deamination reaction that prevented efficient use of the catalyst [21].

We believed that TrpS would be a good starting point to develop a platform for synthesizing a wider variety of Trp analogs than had been reported. In particular, we sought to generate catalysts for C-C, C-N, and C-S bond-forming reactions to make ncAAs from inexpensive starting materials using low catalysts loadings. Our approach was to first simplify the heterodimeric enzyme complex and engineer TrpB to function as a single, stand-alone enzyme. To help the reader understand this strategy, we describe the sophisticated mechanism of allosteric regulation that governs native TrpS activity.

1.1.1 TrpS

In the first steps of the native catalytic cycle, IGP binds in the α-subunit (TrpA), and Ser binds in the β-subunit (TrpB), where it forms a covalent Schiff base linkage to the cofactor, pyridoxal 5'-phosphate (PLP). Once bound, Ser undergoes dehydration to form an electrophilic amino acrylate intermediate. TrpA then catalyzes the retro-aldol cleavage of IGP, releasing indole, which diffuses through a tunnel connecting the two subunits and enters the TrpB active site. There it reacts with the amino acrylate to yield L-tryptophan (Fig. 1.1).

The efficient functioning of TrpS requires that all of the mechanistic steps be carefully synchronized [22]. In essence, both subunits have an open conformation that permits the entry of substrates, but exhibits a slow rate of catalysis, and a closed conformation that accelerates intermediate chemical steps, but cannot bind substrate or release product. Each subunit's equilibrium between open and closed states

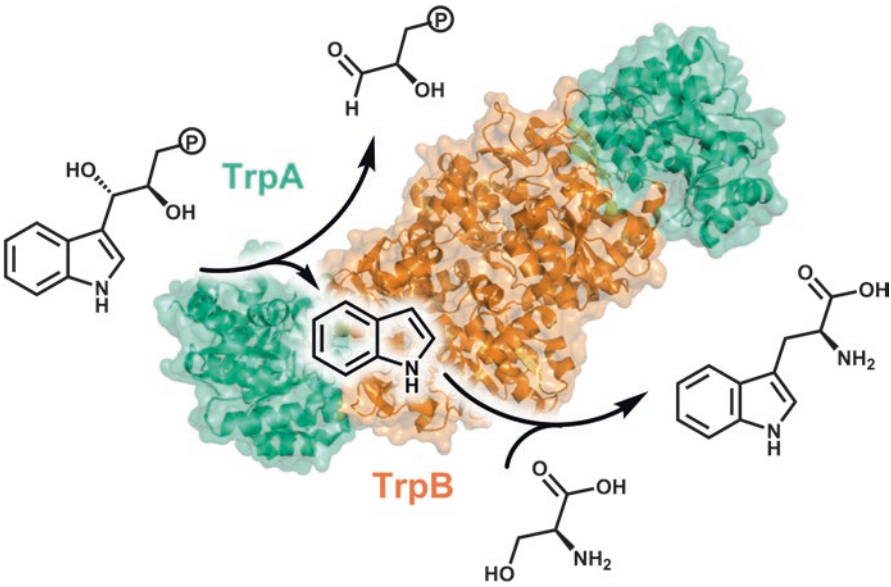


Fig. 1.1 Overall reaction catalyzed by TrpS. The α-subunit (*green*) cleaves IGP in a retro-aldol reaction that releases indole, which diffuses through a molecular tunnel into the β-subunit (*orange*). There, indole reacts with Ser in a PLP-dependent β-substitution reaction to yield Trp

is allosterically modulated by the other subunit, thus ensuring that intermediates are not released prematurely. In particular, the indole must be available to TrpB immediately upon formation of the amino acrylate intermediate, which could otherwise decompose through hydrolysis. However, if indole were released before the subunits are in a fully closed state, it would leak into the cellular medium, whereupon it would diffuse through the membrane and be lost. To accomplish the necessary synchronicity, each subunit acts as an allosteric effector to the other [12, 22].

The molecular basis for this synchronization comprises structural transitions in both subunits [23–25]. In TrpA, the majority of residues in the αL6 loop switch from disordered to well ordered, forming a closed state (Fig. 1.2). In TrpB, the structural change is more significant; around 20% of the residues change conformation. The so-called COMM domain, which refers to the region of TrpB that mediates the communication between subunits, undergoes a rigid-body motion and rotates between open, partially closed, and fully closed conformations (Fig. 1.2). Closure in the COMM domain impedes access to the active site from solution and stabilizes the closed conformation in TrpA. Concurrently with the COMM domain motion, the conformation of TrpB in the interface between the two subunits is also altered. These conformational changes combine to form a 25-Å tunnel between both subunits, allowing indole to diffuse from TrpA to the TrpB active site.

TrpS thus exhibits a sophisticated allosteric control mechanism in which both subunits play a fundamental role. However, only the β-subunit performs catalysis in the β-substitution reaction that is useful to make ncAAs. Unfortunately, the

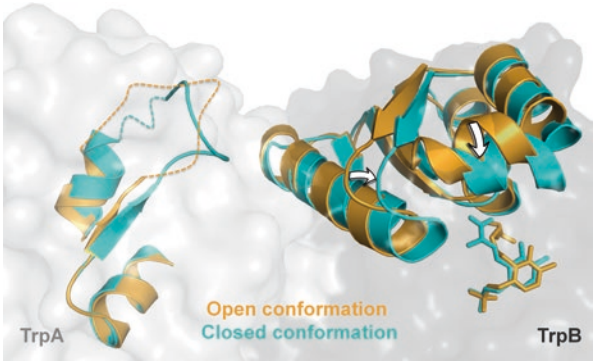


Fig. 1.2 Structural transitions in the conformational switch from open to closed states in *Salmonella typhimurium* TrpS. The open conformation is in orange (PDB ID: 1KFK), and the closed conformation is in *blue* (PDB ID: 2J9X). In TrpA (*light gray*), many residues in loop αL6 are disordered in the open conformation, but most become well ordered upon the switch to the closed conformation. In TrpB (*dark gray*), the *arrows* indicate the direction of the motion of the COMM domain from the open to the closed state. The PLP is represented in sticks

allosteric regulation has hindered the use of isolated TrpB, whose activity in isolation is seriously diminished compared to the full TrpS complex. Consequently, we sought to engineer a stand-alone TrpB for efficient catalysis in the absence of TrpA. We hypothesized that such a simplified biocatalyst could serve as an efficacious starting point for further engineering to expand reactivity.

1.2 Activation of TrpB from *Pyrococcus furiosus* by Directed Evolution

The first task in engineering a stand-alone TrpB catalyst was to identify a suitable parent for directed evolution. Most studies of TrpS have been done with the homolog from *Salmonella typhimurium* (*St*TrpS), a mesophilic organism. We decided to use TrpS from *Pyrococcus furiosus*, a thermophilic archaeon that survives at 100 °C. The ability of this enzyme to function at high temperature (75 °C) enables solubilization of highly hydrophobic substrates such as indole without addition of cosolvent. Another significant advantage of engineering a protein from a thermophilic organism is the possibility of accumulating more mutations that boost activity but may be destabilizing [26].

We wanted to evolve TrpB to catalyze its native reaction, the condensation of indole and Ser to give Trp, more efficiently as a stand-alone enzyme. For this we developed a high-throughput assay that measures the change in absorbance at 290 nm as indole is converted to Trp [27]. The use of a thermostable protein is highly advantageous for screening at this wavelength, because the background absorption from *E. coli* proteins can be reduced through heat treatment at 75 °C, which yields moderately pure *Pf*TrpB enzymes.

We constructed a random mutagenesis library of *Pf*TrpB by error-prone PCR and measured the activity of 528 clones. From this library, we identified 20 clones (3.8% of all variants) with at least a 35% increase in V_{max} relative to the wild-type enzyme [28]. The most active of these contained a single Thr→Ser mutation that increased the catalytic efficiency of *Pf*TrpB on indole by 20-fold, which is even greater than the change induced by TrpA binding. Twelve activating mutations were recombined, and screening yielded a new variant, *Pf*TrpB^{4D11}, that retained the T292S mutation and incorporated four more (E17G, I68V, F274S, T321A). This enzyme, which has a further 2.6-fold increase in catalytic efficiency, was used as the parent for a final round of random mutagenesis, from which we identified *Pf*TrpB^{0B2} harboring the single additional P12L mutation that increased the catalytic efficiency to 3.3×10^5 M^{-1} s^{-1} with indole, 83-fold higher than *Pf*TrpB and threefold higher than the allosterically activated *Pf*TrpS complex [28].

We next investigated whether the increased activity of this stand-alone TrpB was achieved through the same mechanism as the binding of TrpA. Several lines of evidence suggested this was the case. We performed our screening under saturating conditions of each substrate and nevertheless observed a coupled decrease in the K_M for each substrate. *Pf*TrpA binding not only causes a ~threefold increase in k_{cat} but also a decrease in K_M values for Ser and indole, by twofold and fourfold, respectively, suggesting a similar mechanism of activation at work during both effector binding and mutational activation.

To further understand how *Pf*TrpA regulates *Pf*TrpB, we used X-ray crystallography and UV-vis spectroscopy to probe conformational changes and the steady-state distribution of intermediates in the active site. Importantly, our data showed that the structural and spectroscopic properties of *Pf*TrpB and *Pf*TrpS are very similar to those reported in the distantly-related but well-studied enzyme from *Salmonella typhimurium* [25]. The UV-vis spectrum of the internal aldimine (i.e., when the PLP is bound to the catalytic lysine) has a λ_{max} of ~412 nm (Fig. 1.3a). Ser binding to *Pf*TrpB is associated with the large-scale conformational rearrangement of the COMM domain into a partially closed state and accumulation of an external aldimine intermediate, which shifts λ_{max} to 428 nm (Fig. 1.3a). When *Pf*TrpA is bound, the Ser-bound spectrum shifts to a characteristic λ_{max} at 350 nm, consistent with stabilization of the amino acrylate intermediate. Structural studies with *St*TrpS show that this species is stabilized by a fully closed state of the COMM domain (Fig. 1.2), which also has an increased affinity for indole [12]. As can be seen in Fig. 1.3b, the spectrum after addition of Ser to *Pf*TrpB^{0B2} corresponds to amino acrylate intermediate, further supporting the hypothesis that mutations increase activity through the same mechanism as allosteric effector binding, i.e., stabilization of the closed conformational state.

Each of these experiments probed changes in *Pf*TrpB during its native catalytic cycle with Ser and indole. We found that *Pf*TrpA binding also increases the relative rate of product formation upward of 200-fold with different indole analogs. Therefore, it was of interest to know whether mutations were also activating for ncAA synthesis. We found that our stand-alone *Pf*TrpB^{0B2} catalyst was also broadly activated for eight different indole analogs in C-C and C-N bond-forming reactions

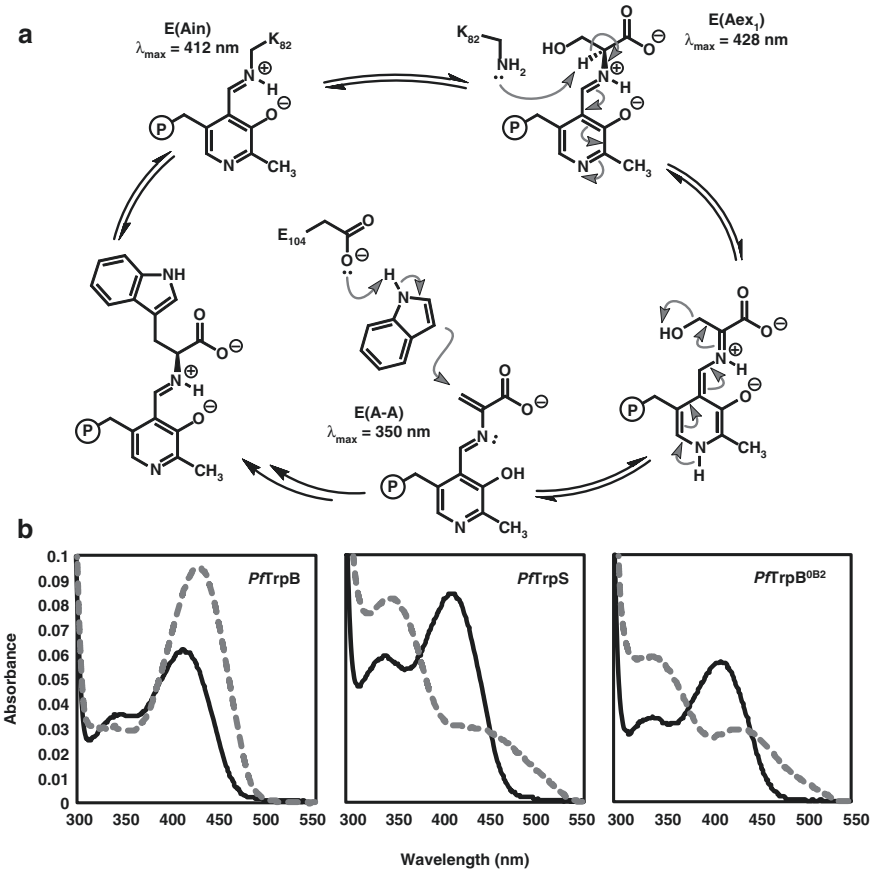


Fig. 1.3 Spectroscopic signatures of the TrpB catalytic cycle. (**a**) Different intermediates in the catalytic cycle of TrpB are shown with their corresponding λ_{max}. Protonation states are assigned according to Caulkins et al. [29] from study of *St*TrpS. Spectroscopic data are broadly similar for the two enzymes, supporting this assignment. (**b**) UV-vis spectra of *Pf*TrpB in isolation, in the native complex, and in the stand-alone engineered protein *Pf*TrpB^{0B2}. Spectra recorded with 20 μM of protein (*black lines*) and after addition of 20 mM Ser (*gray dashes*). The shifts in λ_{max} indicate that *Pf*TrpB accumulates E(Aex$_1$) at steady state, whereas *Pf*TrpS and *Pf*TrpB^{0B2} accumulate the E(A-A) intermediate

[28]. This trend was similar to the rate enhancement induced by *Pf*TrpA binding, with some differences emerging between the two enzyme systems. *Pf*TrpB^{0B2} was moderately faster at C-N bond-forming reactions with indole and indazole, whereas *Pf*TrpS was moderately faster with 5-bromoindole. Hence, mutations that increased activity with indole were broadly activating with other substrates. Again, *Pf*TrpB^{0B2} resembles the catalytic features described for *Pf*TrpS.

Given the considerable interest in understanding allosteric phenomena [30–34], a compelling question arose: are activating mutations confined to a distinct allosteric pathway or interface? The first round of evolution identified 27

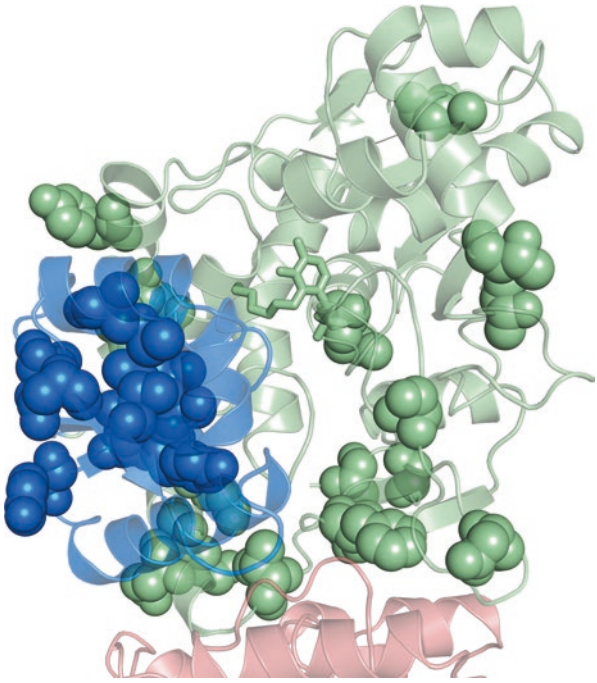


Fig. 1.4 Sites where mutation of *Pf*TrpB is associated with increased activity. In the structure of *Pf*TrpS (PDB ID: 5E0K), TrpA is shown in *pink*, and TrpB is shown in *green* and *blue*. PLP is depicted in sticks. Residues where mutations were found that gave validated increases in V_{max} during high-throughput screening are shown as spheres. There is an abundance of sites in the COMM domain (colored in *blue*) and at the subunit interface that give rise to increases in activity

mutations distributed across 20 improved clones. Although the effect of each mutation has not been measured individually, the data provide a broad picture of *Pf*TrpB activation (Fig. 1.4). It was described that some residues undergo switch-like behavior upon effector binding in *St*TrpS, but just one of the 27 possible activating mutations (N166D) was found along the homologous route for *Pf*TrpB [12, 35, 36]. Taking a more generous view of what an allosteric "pathway" might look like, 17 of the 27 activating mutations (63%) were found at either the protein-protein interface or within the COMM domain. These regions comprise 31% of the total *Pf*TrpB sequence, indicating modest enrichment within the areas previously thought to control allosteric signaling in TrpB. While the mutations have a positive effect on *Pf*TrpB catalysis, it is not clear how or why they influence the rate of the reaction. Indeed, almost 40% of activating mutations are outside of any region that one might *a priori* think has an impact on allosteric signaling or catalysis.

From the study of *Pf*TrpB activation, it is clear that mutations can reproduce complex conformational changes induced by effector binding. Similar effects were previously shown in a handful of other examples from the literature. Shi and Kay identified mutations for the activation of the bacterial HslV protease [37]. The proteolytic activity of HslV is increased ~200-fold upon binding of its partner protein HslU, which is essential for its role in cellular protein degradation [38]. A sensitive NMR analysis was used to monitor chemical shift perturbations in Ile, Leu, Met, Val, and Thr residues in HslV upon HslU binding. From these data, a pair of helices

that undergo conformational changes at the HslU binding site was identified. A small panel of conservative mutations at positions within these helices was constructed, and, strikingly, six of the mutations increased catalytic efficiency, the largest by ~20-fold. NMR analysis showed that increases in activity were correlated with substantial chemical shift perturbations similar to the effects of native effector binding.

Another example is the activation of LovD, an acyl transferase that transfers an α-methylbutyrate group that is covalently tethered to an acyl carrier protein, LovF [39, 40]. With substrate surrogates that are not bound to LovF, the acyltransferase activity of LovD is greatly reduced, indicating that the carrier protein also serves as an allosteric activator. Tang and collaborators employed nine rounds of directed evolution to increase activity on a nonnatural substrate in the absence of the protein effector as well as increase the thermal stability and tolerance to organic solvent [40]. They assessed the aggregate effect of these mutations on LovD conformational dynamics with molecular dynamics (MD) simulations. Their results suggested that the activity of LovD is altered upon LovF binding through the stabilization of a closed and catalytically active conformation. Interestingly, simulations suggested that engineered LovD enzymes experience comparable conformation changes in the absence of their effector. Testing this hypothesis experimentally would have been exceptionally difficult without further modification of LovD, as there is no chromophore (like PLP for TrpB) that enables one to probe the steady state of the catalytic cycle directly.

1.3 Activation of TrpB from Other Species by Transfer of Mutations

With our stand-alone *Pf*TrpB in hand, we wished to test whether the mutational activation of TrpB could be generalized to TrpBs from other organisms. All known TrpBs are subject to allosteric regulation by TrpA [12, 41, 42], and it is plausible that the allosteric mechanisms are conserved across TrpS homologs. Furthermore, our interest in expanding the substrate scope of TrpB might be helped by assessing other TrpB homologs, since enzyme homologs often exhibit different substrate scopes [43, 44]. For example, native *St*TrpS is a poor catalyst for N-alkylation, whereas *Pf*TrpS is moderately proficient. However, we did not wish to repeat the effort required to activate *Pf*TrpB (screening ~3100 clones) for the other homologs. Instead we tested whether activating mutations in *Pf*TrpB^{0B2} have the same effects when transferred to TrpBs from other species.

Successful transfer of beneficial mutations among homologous proteins has been reported numerous times, although not for allosteric properties, as far as we know. For instance, multiple sequence alignments of homologs allow the identification of consensus residues that tend to be thermostabilizing within the protein family [45]. Also, mutations that alter nicotinamide cofactor specificity can be transferred to homologous enzymes [46]; in this case, structural and sequence analyses of an entire protein family provided specialized “recipes” to

change specificity from NADPH to NADH. However, engineering allostery may be significantly more complex than transferring mutations that enhance thermostability, where the mutational effects are largely additive, or specificity, where the effects are usually more localized. Allosteric regulation occurs through a dynamic mechanism that transfers information between the allosteric binding site and the enzyme active site and involves many, if not all, residues in the protein. Often, experimental evidence establishes that a residue that participates in transmitting this information is not conserved across different homologs. This holds even when the allosteric mechanisms are superficially similar, and evolution frequently causes homologous proteins to develop different allosteric mechanisms [47].

To test the transferability of the allostery-mimicking mutations, we selected diverse TrpB homologs with differing sequence identities and well separated in the phylogenetic tree [48]. The closest homolog to *Pf*TrpB tested was from *Archaeoglobus fulgidus*, *Af*TrpB (72% sequence identity), which is a thermophilic archaeon. We also selected the TrpB from *Thermotoga maritima* (*Tm*TrpB, 64% sequence identity), a thermophile that belongs to the bacterial domain of life. Lastly, we chose the TrpB from *Escherichia coli* (*Ec*TrpB, 57% sequence identity), a mesophile. The sequence alignment of the homologs showed some differences at the positions of the activating mutations in *Pf*TrpB^{0B2}. Importantly, two mutations in *Pf*TrpB^{0B2} were already present as the wild-type residues in two of the homologs, A321 in *Tm*TrpB (mutation T321A in *Pf*TrpB^{0B2}) and S297 in *Ec*TrpB (mutation T292S in *Pf*TrpB^{0B2}).

Making the 0B2 mutations in the homologs led to varied levels of catalytic efficiency. *Af*TrpB^{0B2} was indeed activated, with a ~20-fold increase in catalytic efficiency with respect to wild type. Similar to *P. furiosus* enzymes, the UV-vis spectrum after addition of Ser to *Af*TrpB showed a λ_{max} at 428 nm, reflecting accumulation of the external aldimine intermediate. After addition of Ser to the 0B2 mutant, the spectrum was similar to *Af*TrpS, with a λ_{max} at 350 nm [49] corresponding to the amino acrylate intermediate. Hence, these mutations, which mimic the binding of the allosteric protein partner in *Pf*TrpB, appear to have the same effect in *Af*TrpB. However, this was not the case for the other two mutant homologs, where the catalytic efficiencies of *Tm*TrpB^{0B2} and *Ec*TrpB^{0B2} decreased by more than 40% relative to the wild-type enzymes.

To determine whether a subset of the mutations could activate the other two homologs, we made and screened recombination libraries of the 0B2 mutations for *Tm*TrpB and *Ec*TrpB. With this strategy, we found three activating mutations (P19G, I69V, and T292S) for *Tm*TrpB. T292S was the most activating single mutation, conferring a sixfold increase in catalytic efficiency. All the variants containing this mutation showed a UV-vis spectrum after addition of Ser similar to the native *Tm*TrpS complex [49]. The most activated *Tm*TrpB variant harbored the three mutations and exhibited a tenfold increase in catalytic efficiency.

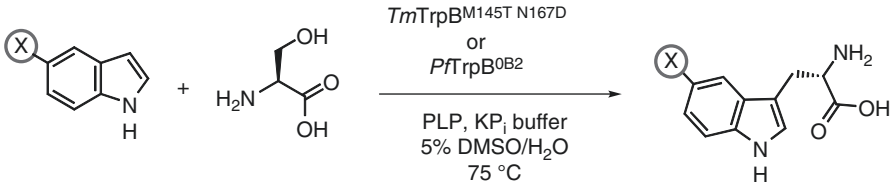
Activation of the last, *E. coli* homolog was more challenging: the recombination library of the 0B2 mutations in *Ec*TrpB produced no increased activity. The T292S mutation had a prominent effect in the other TrpB homologs, but Ser is the native

residue in the equivalent position of *Ec*TrpB. We made a saturation mutagenesis library at this site but again did not find any activated variants. In a final attempt to activate this homolog, we returned to the initial random mutagenesis performed on *Pf*TrpB, where 27 activating mutations were found [28]. We chose to test the mutations of the double mutant *Pf*TrpB$^{M144T\ N166T}$ because these residues are highly conserved across all known TrpBs and are located in the COMM domain. These mutations activated *Ec*TrpB, giving more than a twofold increase in catalytic efficiency. We tested the effects of these two mutations in the other homologs and found that all were activated, with a two- to fivefold increase in catalytic efficiency. The UV-vis spectra after addition of Ser to *Pf*TrpB$^{M144T\ N166D}$ or *Ec*TrpB$^{M149T\ N171D}$ did not have a λ_{max} at 350 nm and instead showed an accumulation of the external aldimine species. However, after addition of Ser to the corresponding *Af*TrpB and *Tm*TrpB mutants, the spectra showed a λ_{max} at 350 nm, characteristic of the amino acrylate species. These results suggest that this set of mutations also mimics the allosteric activation exerted by TrpA binding, although in some of the homologs, this activation does not completely reach TrpS-like behavior [49]. Similarly, in the initial evolution of *Pf*TrpB, the activating T292S mutation alone was not sufficient to shift the spectrum to the amino acrylate species, which required four more mutations.

Once we accumulated this panel of stand-alone TrpB enzymes, we sought to compare their substrate profiles with substituted indoles. We were particularly interested in accessing Trp derivatives with substituents in the 5 position because this site is often substituted in biologically relevant Trp-based compounds. For instance, the halogenase PyrH chlorinates tryptophan in the 5 position during the biosynthesis of the antifungal antibiotic pyrroindomycin B [8]. In the biosynthesis of the neurotransmitter serotonin and the hormone melatonin, a tryptophan hydroxylase mono-oxygenates tryptophan in the 5 position [1]. Halogenation or borylation in this position would generate analogs that could serve as intermediates for other biologically active compounds through cross-coupling reactions [50–52]. Different TrpS homologs have shown a substantial decrease in activity with 5-substituted indoles bearing a substituent bulkier than fluorine [17, 18].

We tested our panel of activated stand-alone biocatalysts with 5-bromoindole. Remarkably, one of the new activated TrpB, *Tm*TrpB$^{M145T\ N167D}$, was ~sixfold faster than *Pf*TrpB^{0B2}, the first enzyme we engineered (Table 1.1). We tested eight more 5-substituted indoles with the two enzymes and saw that *Tm*TrpB$^{M145T\ N167D}$ was always faster than *Pf*TrpB^{0B2}, from 1.4- to 7.5-fold (Table 1.1). As shown in Table 1.1, the yields for 5-chlorotryptophan and 5-bromotryptophan are 93% and 88%, respectively. Previous use of *St*TrpS for these two reactions reported yields of 61% for 5-chlorotryptophan and only 33% for 5-bromotryptophan [17]. Moreover, the chemical syntheses of these compounds described to date involve multiple steps, generate racemic products, and have final yields below 50% [6, 53–55]. Hence, our panel of stand-alone TrpBs is a useful set of biocatalysts for the synthesis of ncAAs and should be a fertile starting point for the further development of biocatalysts to make more synthetically challenging ncAAs.

Table 1.1 Yields and total turnovers of the reactions catalyzed by *Tm*TrpB$^{M145T\ N167D}$ to obtain 5-substituted tryptophan derivatives. The rates relative to *Pf*TrpB^{0B2} are also included



		Reaction catalyzed by *Tm*TrpB$^{M145T\ N167D}$	
X	Rate relative to *Pf*TrpB^{0B2}	Yield (%)	Total turnovers
Cl	3.0	93	9300
Br	5.6	88	4400
NO_2	7.5	25	1250
$B(OH)_2$	1.8	38	1900
CHO	1.9	32	1600
CN	4.5	49	2450
OH	1.4	93	9300
CH_3	1.4	91	9100
OCH_3	1.5	76	7600

1.4 Directed Evolution of *Pf*TrpB for the Synthesis of β-Branched Amino Acids

β-Branched amino acids are particularly desirable because the additional substituent at C_β alters the conformational properties of peptides and small molecules that bear it. For example, the clinically important antibiotic daptomycin features a β-methylglutamate residue, and the activity of the drug is greatly diminished without this modification [56, 57]. Unfortunately, chemical synthesis of β-branched amino acids is particularly challenging, owing to the need for both enantio- and diastereoselective transformations.

One strategy to access β-branched ncAAs would be to substitute Ser in the reaction with Thr. While the nucleophile scope of TrpB has been well explored and our panel of stand-alone variants can catalyze an array of substitutions, synthesis of Trp analogs by replacing Ser had not been reported. We screened our stand-alone *Pf*TrpB catalysts for activity with a variety of Ser analogs and found that the natural amino acid L-threonine (Thr) can replace Ser, yielding (2*S*,3*S*)-β-methyltryptophan (β-MeTrp) in a single step (Fig. 1.5). Previous attempts to perform reactions with Thr required the use of the strong nucleophile benzyl mercaptan, which proceeded with greatly diminished activity compared to Ser, and the stereochemistry was unknown [20].

β-MeTrp is an intermediate in the natural biosynthetic pathways to maremycin and streptonigrin [7, 9]. Studies have shown that this intermediate is produced in

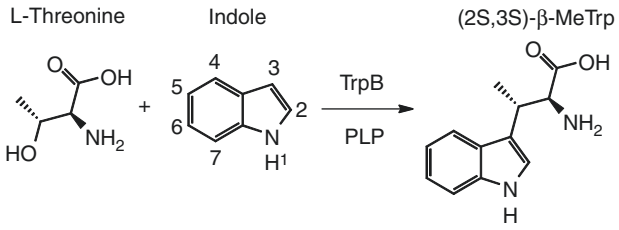


Fig. 1.5 β-Substitution of Thr with indole yields β-MeTrp. This reaction is catalyzed at trace levels by *Pf*TrpB, but the level increases with directed evolution. Activity was also observed with a variety of indole analogs

four steps from Trp, using three different enzymes with *S*-adenosylmethionine as the methyl source. Chemical synthetic routes to this ncAA have relatively good selectivity but require multiple steps, making it challenging to apply for the production of diverse β-MeTrp analogs. Therefore, we sought to evolve our platform for the synthesis of these challenging β-branched amino acids using Thr as a nonnative substrate.

We first explored the *Pf*TrpB lineage of stand-alone enzymes for activity with Thr. Wild-type *Pf*TrpB yielded just 66 turnovers in 24 h. The single mutant *Pf*TrpB^{T292S} showed ~sixfold enhanced activity, and reactions with variant *Pf*TrpB^{4D11} (containing T292S, E17G, I68V, F274S, and T321A mutations) yielded 660 turnovers. Interestingly, *Pf*TrpB^{0B2} gave slightly slower rates than *Pf*TrpB^{4D11}, despite having higher activity with Ser [28]. Therefore, we selected *Pf*TrpB^{4D11} for directed evolution to increase activity with Thr.

We screened 352 clones from a random mutagenesis library of *Pf*TrpB^{4D11} and identified six missense mutations in five clones with increased V_{max}. The most active variant, *Pf*TrpB^{4G1}, has a single additional mutation, F95L. Recombination of all of the mutations and screening resulted in variant *Pf*TrpB^{2B9}, which has mutations I16V, F95L, and V384A and significantly enhanced activity with Thr. Interestingly, the I16V mutation is adjacent to E17G present in the parent, and the F95L mutation is adjacent to the COMM domain of TrpB. Combined, these mutations provide at least a 6000-fold boost in productivity compared to the wild-type *Pf*TrpB enzyme.

This engineered enzyme has several positive features as a biocatalyst. As we screened for V_{max}, some of the increases in activity in cell lysates were due to boosts in the expression of soluble enzyme. Hence, the protein can now be prepared at ~350 mg/L of culture, facilitating larger-scale reactions. The enzyme also retains high activity at elevated temperatures, allowing for high substrate loading. However, reactions with a single equivalent of each substrate give only 44% conversion to products, which corresponds to 2220 total turnovers (TTN). UV-vis spectroscopy showed that an abortive deamination reaction is occurring when Thr is incubated with *Pf*TrpB^{2B9}. With the addition of more equivalents of Thr (up to ten), we observed >99% conversion based on indole and up to 8200 TTN. Additionally, this

reaction proceeds with >99% enantiomeric and diastereomeric excess, highlighting the exquisite selectivity of the enzyme.

Using these reaction conditions, we explored the nucleophile scope of this new β-substitution reaction. Indoles with methylation at the 2 and 6 positions (numbering of indole is shown in Fig. 1.5) were well tolerated by the enzyme, yielding 6400 and 1100 turnovers, respectively. We also observed product formation with 4- and 5-fluoroindole, with ~3.4-fold lower TTN for the 4-fluoroindole, which is more electron deficient at C-3 than the 5-fluoroindole analog. However, we did not observe product formation using 5-chloro-, 5-bromo-, or 6-hydroxyindoles, demonstrating a reduced substrate scope in the new reaction.

We observed that 7-azaindole reacted with 220 TTN and also found a secondary product corresponding to N-alkylation. This regioselectivity is identical to that of indazole, which we found reacts exclusively in an N-alkylation reaction that proceeds with 500 TTN. Lastly, we observed a productive reaction with thiophenol in 1300 TTN, demonstrating that the enzyme can catalyze stereoselective C-C, C-N, and C-S bond-forming reactions.

The development of this enzymatic platform for the synthesis of β-branched ncAAs was greatly facilitated by previous efforts to engineer a stand-alone enzyme. At all stages, only a single enzyme was necessary to catalyze the reaction, and this simplified system supported expression of the catalyst at high levels, ~350 mg of *Pf*TrpB^{2B9} per L culture, which will facilitate production of these important synthons [58].

1.5 Summary and Outlook

We have engineered the heterodimeric enzyme TrpS into a single enzyme platform for the synthesis of ncAAs from simple starting materials. This was accomplished by identifying mutations in TrpB that recapitulate the effects that are induced by its allosteric binding partner, TrpA. We identified dozens of mutations that are activating in *Pf*TrpB and showed that some of them could be transferred to homologs to produce an array of catalysts with varied properties. These new stand-alone TrpB enzymes can be evolved readily for altered function, as we demonstrated for the β-substitution of Thr. All of this was accomplished with mutations that were identified by screening random mutant libraries, and the mutations are distributed throughout the protein. Hence, each of the catalysts has a unique active site that can be engineered to increase activity with a particular nucleophile or electrophile. We believe this approach will continue to yield superior catalysts for the biocatalytic production of desirable ncAAs.

Acknowledgments We gratefully thank Sabine Brinkmann-Chen for a critical reading of this chapter. We also thank David K. Romney for his helpful discussions during the elaboration of the chapter. Javier Murciano-Calles acknowledges financial support from the Alfonso Martín Escudero Foundation. This work was funded through the Jacobs Institute for Molecular Engineering for Medicine and Ruth Kirschstein NIH Postdoctoral Fellowship F32GM110851 (to Andrew R. Buller).

References

1. Zhang J, Wu C, Sheng J, Feng X (2016) Molecular basis of 5-hydroxytryptophan synthesis in *Saccharomyces cerevisiae*. Mol Biosyst 12(5):1432–1435
2. Ikeda M et al (1965) Studies on the biosynthesis of nicotinamide adenine dinucleotide: II. A role of picolinic carboxylase in the biosynthesis of nicotinamide adenine dinucleotide from tryptophan in mammals. J Biol Chem 240(3):1395–1401
3. Stepanova AN et al (2008) TAA1-Mediated auxin biosynthesis is essential for hormone crosstalk and plant development. Cell 133(1):177–191
4. Tao Y et al (2008) Rapid synthesis of auxin via a new tryptophan-dependent pathway is required for shade avoidance in plants. Cell 133(1):164–176
5. Barry SM et al (2012) Cytochrome P450-catalyzed L-tryptophan nitration in thaxtomin phytotoxin biosynthesis. Nat Chem Biol 8(10):814–816
6. Kieffer ME, Repka LM, Reisman SE (2012) Enantioselective synthesis of tryptophan derivatives by a tandem Friedel-Crafts conjugate addition/asymmetric protonation reaction. J Am Chem Soc 134(11):5131–5137
7. Kong D et al (2016) Identification of (2S,3S)-β-methyltryptophan as the real biosynthetic intermediate of antitumor agent streptonigrin. Sci Rep 6:20273
8. Zehner S et al (2005) A regioselective tryptophan 5-halogenase is involved in pyrroindomycin biosynthesis in *Streptomyces rugosporus* LL-42D005. Chem Biol 12(4):445–452
9. Zou Y et al (2013) Stereospecific biosynthesis of β-methyltryptophan from L-tryptophan features a stereochemical switch. Angew Chem Int Ed 52(49):12951–12955
10. Lang K, Chin JW (2014) Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins. Chem Rev 114(9):4764–4806
11. Patel R (2013) Biocatalytic synthesis of chiral alcohols and amino acids for development of pharmaceuticals. Biomolecules 3(4):741
12. Dunn MF (2012) Allosteric regulation of substrate channeling and catalysis in the tryptophan synthase bienzyme complex. Arch Biochem Biophys 519(2):154–166
13. Corr MJ, Smith DRM, Goss RJM (2016) One-pot access to L-5,6-dihalotryptophans and L-alknyltryptophans using tryptophan synthase. *Tetrahedron* 72(45):7306–7310
14. Ferrari D, Niks D, Yang L-H, Miles EW, Dunn MF (2003) Allosteric communication in the tryptophan synthase bienzyme complex: roles of the β-subunit aspartate 305−arginine 141 salt bridge. Biochemistry 42(25):7807–7818
15. Goss RJM, Newill PLA (2006) A convenient enzymatic synthesis of L-halotryptophans. Chem Commun 47:4924–4925
16. Perni S, Hackett L, Goss RJ, Simmons MJ, Overton TW (2013) Optimisation of engineered *Escherichia coli* biofilms for enzymatic biosynthesis of L-halotryptophans. AMB Express 3(1):1–10
17. Smith DRM et al (2014) 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 16(10):2622–2625
18. Tsoligkas AN et al (2011) Engineering biofilms for biocatalysis. Chembiochem 12(9):1391–1395
19. Winn M, Roy AD, Grüschow S, Parameswaran RS, Goss RJM (2008) A convenient one-step synthesis of L-aminotryptophans and improved synthesis of 5-fluorotryptophan. Bioorg Med Chem Lett 18(16):4508–4510
20. Esaki N, Tanaka H, Miles EW, Soda K (1983) Enzymatic synthesis of S-substituted L-cysteines with tryptophan synthase of *Escherichia coli*. Agric Biol Chem 47(12):2861–2864
21. Ferrari D, Yang LH, Miles EW, Dunn MF (2001) β-D305A Mutant of tryptophan synthase shows strongly perturbed allosteric regulation and substrate specificity. Biochemistry 40(25):7421–7432
22. Niks D et al (2013) Allostery and substrate channeling in the tryptophan synthase bienzyme complex: evidence for two subunit conformations and four quaternary states. Biochemistry 52(37):6396–6411

23. Barends TRM et al (2008) Structure and mechanistic implications of a tryptophan synthase quinonoid intermediate. Chembiochem 9(7):1024–1028
24. Lai J et al (2011) X-ray and NMR Crystallography in an enzyme active site: the indoline quinonoid intermediate in tryptophan synthase. J Am Chem Soc 133(1):4–7
25. Ngo H et al (2007) Allosteric regulation of substrate channeling in tryptophan synthase: modulation of the L-serine reaction in stage I of the β-reaction by α-site ligands. Biochemistry 46(26):7740–7753
26. Bloom JD, Labthavikul ST, Otey CR, Arnold FH (2006) Protein stability promotes evolvability. Proc Natl Acad Sci U S A 103(15):5869–5874
27. Lane AN, Kirschner K (1983) The catalytic mechanism of tryptophan synthase from *Escherichia coli*. Eur J Biochem 129(3):571–582
28. Buller AR et al (2015) Directed evolution of the tryptophan synthase β-subunit for stand-alone function recapitulates allosteric activation. Proc Natl Acad Sci 112(47):14599–14604
29. Caulkins BG et al (2015) Catalytic roles of beta Lys87 in tryptophan synthase: N-15 solid state NMR studies. Biochim Biophys Acta Protein Proteomics 1854(9):1194–1199
30. Sol A, Tsai C-J, Ma B, Nussinov R (2009) The origin of allosteric functional modulation: multiple pre-existing pathways. Structure 17(8):1042–1050
31. Gerek ZN, Ozkan SB (2011) Change in allosteric network affects binding affinities of PDZ domains: analysis through perturbation response scanning. PLoS Comput Biol 7(10):e1002154
32. McLeish Tom CB, Cann Martin J, Rodgers Thomas L (2015) Dynamic transmission of protein allostery without structural change: spatial pathways or global modes? Biophys J 109(6):1240–1250
33. Suel GM, Lockless SW, Wall MA, Ranganathan R (2003) Evolutionarily conserved networks of residues mediate allosteric communication in proteins. Nat Struct Mol Biol 10(1):59–69
34. Woods KN, Pfeffer J (2016) Using THz spectroscopy, evolutionary network analysis methods, and MD simulation to map the evolution of allosteric communication pathways in c-type lysozymes. Mol Biol Evol 33(1):40–61
35. Raboni S, Bettati S, Mozzarelli A (2005) Identification of the geometric requirements for allosteric communication between the α- and β-Subunits of tryptophan synthase. J Biol Chem 280(14):13450–13456
36. Weyand M, Schlichting I, Herde P, Marabotti A, Mozzarelli A (2002) Crystal structure of the βSer178 → Pro mutant of tryptophan synthase: a "knock-out" allosteric enzyme. J Biol Chem 277(12):10653–10660
37. Shi L, Kay LE (2014) Tracing an allosteric pathway regulating the activity of the HslV protease. Proc Natl Acad Sci 111(6):2140–2145
38. Yoo SJ et al (1996) Purification and characterization of the heat shock proteins HslV and HslU that form a new ATP-dependent protease in *Escherichia coli*. J Biol Chem 271(24):14035–14040
39. Gao X et al (2009) Directed evolution and structural characterization of a simvastatin synthase. Chem Biol 16(10):1064–1074
40. Jiménez-Osés G et al (2014) The role of distant mutations and allosteric regulation on LovD active site dynamics. Nat Chem Biol 10(6):431–436
41. Hettwer S, Sterner R (2002) A novel tryptophan synthase β-subunit from the hyperthermophile *Thermotoga maritima*: quaternary structure, steady-state kinetics, and putative physiological role. J Biol Chem 277(10):8194–8201
42. Hiyama T, Sato T, Imanaka T, Atomi H (2014) The tryptophan synthase β-subunit paralogs TrpB1 and TrpB2 in *Thermococcus kodakarensis* are both involved in tryptophan biosynthesis and indole salvage. FEBS J 281(14):3113–3125
43. Dunn MR, Otto C, Fenton KE, Chaput JC (2016) Improving polymerase activity with unnatural substrates by sampling mutations in homologous protein architectures. ACS Chem Biol 11(5):1210–1219
44. Khanal A, Yu McLoughlin S, Kershner JP, Copley SD (2015) Differential effects of a mutation on the normal and promiscuous activities of orthologs: implications for natural and directed evolution. Mol Biol Evol 32(1):100–108

45. Lehmann M, Wyss M (2001) Engineering proteins for thermostability: the use of sequence alignments versus rational design and directed evolution. Curr Opin Biotechnol 12(4):371–375
46. Brinkmann-Chen S et al (2013) General approach to reversing ketol-acid reductoisomerase cofactor dependence from NADPH to NADH. Proc Natl Acad Sci U S A 110(27):10946–10951
47. Kuriyan J, Eisenberg D (2007) The origin of protein interactions and allostery in colocalization. Nature 450(7172):983–990
48. Merkl R (2007) Modelling the evolution of the archaeal tryptophan synthase. BMC Evol Biol 7(1):1–20
49. Murciano-Calles J, Romney DK, Brinkmann-Chen S, Buller AR, Arnold FH (2016) A panel of TrpB biocatalysts derived from tryptophan synthase through the transfer of mutations that mimic allosteric activation. Angew Chem Int Ed 55(38):11577–11581
50. Durak LJ, Payne JT, Lewis JC (2016) Late-stage diversification of biologically active molecules via chemoenzymatic C–H functionalization. ACS Catal 6(3):1451–1454
51. Pathak TP, Miller SJ (2013) Chemical tailoring of teicoplanin with site-selective reactions. J Am Chem Soc 135(22):8415–8422
52. Roy AD, Grüschow S, Cairns N, Goss RJM (2010) Gene expression enabling synthetic diversification of natural products: chemogenetic generation of pacidamycin analogs. J Am Chem Soc 132(35):12243–12245
53. Blaser G, Sanderson JM, Batsanov AS, Howard JAK (2008) The facile synthesis of a series of tryptophan derivatives. Tetrahedron Lett 49(17):2795–2798
54. Konda-Yamada Y et al (2002) Convenient synthesis of 7′ and 6′-bromo-D-tryptophan and their derivatives by enzymatic optical resolution using D-aminoacylase. Tetrahedron 58(39): 7851–7861
55. Ma C, Liu X, Yu S, Zhao S, Cook JM (1999) Concise synthesis of optically active ring-A substituted tryptophans. Tetrahedron Lett 40(4):657–660
56. Fowler VG et al (2006) Daptomycin versus standard therapy for bacteremia and endocarditis caused by *Staphylococcus aureus*. N Engl J Med 355(7):653–665
57. Nguyen KT et al (2006) Combinatorial biosynthesis of novel antibiotics related to daptomycin. Proc Natl Acad Sci U S A 103(46):17462–17467
58. Herger M, van Roye P, Romney DK, Brinkmann-Chen S, Buller AR, Arnold FH (2016) Synthesis of β-branched tryptophan analogs with an engineered variant of tryptophan synthase. J Am Chem Soc 138(27):8388–8391