

SCIENTIFIC REPORTS

OPEN

Engineered P450 biocatalysts show improved activity and regio-promiscuity in aromatic nitration

Ran Zuo¹, Yi Zhang¹, Chao Jiang¹ , John C. Hackett³, Rosemary Loria⁴, Steven D. Bruner⁵ & Yousong Ding¹

Nitroaromatics are among the most important and commonly used chemicals but their production often suffers from multiple unsolved challenges. We have previously described the development of biocatalytic nitration processes driven by an engineered P450 TxtE fusion construct. Herein we report the creation of improved nitration biocatalysts through constructing and characterizing fusion proteins of TxtE with the reductase domain of CYP102A1 (P450BM3, BM3R). The majority of constructs contained variable linker length while one was rationally designed for optimizing protein-protein interactions. Detailed biochemical characterization identified multiple active chimeras that showed improved nitration activity, increased coupling efficiency and higher total turnover numbers compared with TxtE. Substrate promiscuity of the most active chimera was further assessed with a substrate library. Finally, a biocatalytic nitration process was developed to nitrate 4-Me-DL-Trp. The production of both 4-Me-5-NO₂-L-Trp and 4-Me-7-NO₂-L-Trp uncovered remarkable regio-promiscuity of nitration biocatalysts.

The nitro (-NO₂) group acts as an essential unit in a number of pharmaceuticals¹, exemplified by anticancer drug nilutamine, antiparkinson agent tolcapone, and anti-infective agents chloramphenicol and the recently approved delamanid² and nifurtimox-eflornithine combination³. Drug candidates bearing the -NO₂ group also commonly appear in drug pipelines for treating a variety of existing and emerging diseases⁴⁻⁶. Additionally, the nitro group in particular is a versatile synthetic handle present in numerous building blocks in the synthesis of complex drug molecules⁷⁻⁹. The fundamental importance of the nitro group in pharmaceutical industry has driven the development of chemical nitration methods¹⁰. Indeed, syntheses of nitrochemicals, particularly nitroaromatics, are one of the most studied organic reactions¹¹, and classical electrophilic nitration methods with nitric acid as the nitrating reagent dominate current industrial processes. The limitations of the electrophilic method, however, is that it is notoriously non-selective, poorly tolerates other functional groups, poses safety concerns, and generates large quantities of acidic waste. Advanced nitration methods^{7,10,12-15} have recently been developed to account for these issues but none of them are successful on the scales required for pharmaceutical production.

Developing efficient catalysts that can sustainably produce a broad range of chemicals represents a major challenge in modern organic chemistry. Enzymes as biocatalysts are of great synthetic interest because their typically high stereo-, regio- and chemo-selectivity avoids lengthy protection/deprotection steps and the generation of impurities. Furthermore, biocatalytic processes are often non-toxic, require generally mild reaction conditions, and leave no residual heavy metal contamination¹⁶⁻¹⁸. As a result, the footprint of enzymes in the industrial production of chemicals is ever-increasing, exemplified by biocatalytic manufacturing of anti-diabetic drug sitagliptin¹⁹. In these applications, enzymes are capable of catalysing a breadth of organic transformations²⁰. However, despite Nature's successful application of several strategies to synthesize nitro-containing compounds²¹, there has been comparably little effort to exploit them in chemical industry²².

¹Department of Medicinal Chemistry, Center for Natural Products, Drug Discovery and Development, College of Pharmacy, University of Florida, Gainesville, Florida, 32610, USA. ²Department of Pharmaceutical Engineering, School of Chemical Engineering, Nanjing University of Science and Technology, Nanjing, Jiangsu, 210094, China. ³Department of Physiology and Biophysics and the Massey Cancer Center, Virginia Commonwealth University School of Medicine, Richmond, Virginia, 23298, USA. ⁴Department of Plant Pathology, Institute of Food and Agricultural Sciences, University of Florida, Gainesville, Florida, 32611, USA. ⁵Department of Chemistry, University of Florida, Gainesville, Florida, 32611, USA. Correspondence and requests for materials should be addressed to Y.D. (email: yding@cop.ufl.edu)

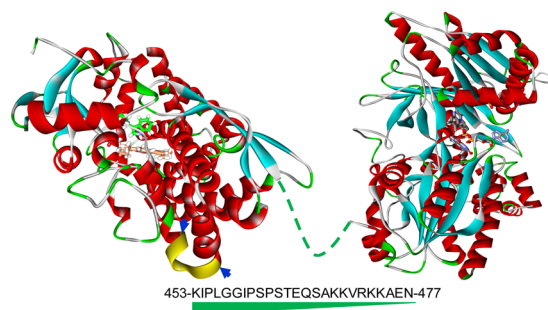


Figure 1. Schematic depiction of chimeric TxtE-BM3R constructs with variable linker length or a swapped loop (yellow). The structure of human NADPH-cytochrome P450 reductase (PDB: 3QE2, right) represented not-available BM3R structure along with TxtE (PDB: 4TPO, left). The 25-AA linker of P450BM3 is shown as a green dash line (middle) along with the amino acid sequence.

We have attempted to expand the biocatalytic toolbox by developing nitration biocatalysts. One example is TxtE, a cytochrome P450 enzyme that uses nitric oxide (NO) and O₂ to nitrate C4 of the *L*-tryptophan (Trp) indole in the thaxtomin biosynthetic pathway²³. It is the only known enzyme enabling direct nitration of a C-H bond in nature, rendering it to be a promising lead for developing nitration biocatalysts. P450s catalyse a wide array of oxygenation reactions under mild conditions and have impressive biotechnological potential^{17,24}. However, the requirement of auxiliary redox proteins and low activity and electron coupling efficiency are common limitations that constrain frequent industrial implementation of P450s^{25,26}. Inspired by previous research on artificial self-sufficient systems^{27–30}, we have recently addressed the need of redox partners in TxtE applications²² by fusing it with the reductase domain (BM3R) of naturally self-sufficient P450 P450BM3³¹. The chimera outperformed TxtE supplemented with spinach ferredoxin (Fer) and ferredoxin reductase (Frd) in terms of catalytic activity, and was subsequently utilized in the biocatalytic syntheses of two fluorinated nitro-Trp analogues²². However, both electron coupling efficiency and total turnover number (TTN) of the developed chimeric enzyme were 20% lower than wild type TxtE²². These must be addressed to attain our goal of developing efficient TxtE-based nitration biocatalysts for biotechnological applications.

Here we describe the characterization of 15 new chimeric TxtE-BM3R biocatalysts. These chimeras were developed by varying the length of a linker connecting TxtE (from *Streptomyces scabies*) and BM3R and swapping a putative interfacial loop on the TxtE to improve interactions with the reductase domain (Fig. 1). These studies have yielded TxtE-BM3R constructs with improved catalytic turnover, coupling efficiency, and broad substrate specificity. These advancements constitute essential steps on the path toward developing advanced nitration biocatalysts for industrial implementation.

Results and Discussions

Design and production of chimeric TxtE-BM3R variants. We previously reported three self-sufficient TxtE constructs²². Of the two reductase domains that we evaluated, the di-flavin reductase BM3R homologous to eukaryotic cytochrome P450 reductase³² conferred superior TxtE nitration activity compared with the P450RhF reductase domain (RhFRED), a natural fusion of Frd and Fer³³. In the latter case, the catalytic performance was notably dependent on the length of the linker connecting TxtE and RhFRED²². Indeed, profound impact of linker length on the performance of artificial fusions has appeared in several recent studies^{34–37}. Inspired by these results, we sought to create serial chimeric TxtE-BM3R variants by varying linker length in this work. Using a stepwise cloning approach, chimeric TxtE-BM3R variants with linker lengths of 3, 6, 9, 11, 14, 17, 19, 22, 24, and 27 amino acids (AA) were first constructed (see Supplementary Fig. S1) to quickly assess potential influences of linker length on enzyme nitration performance. To investigate the optimal length, a second set of variants with linker lengths of 12, 13, 15, and 16 was latter prepared by the same cloning method. Hereafter, these TxtE-BM3R variants are designated as TB-linker length, so for example the TxtE-BM3R variant connected by a 3-AA linker is denoted TB3. In all variants, two amino acids glutamate (E) and leucine (L) were appended to *N*-termini of the linkers as a result of the *SacI* digestion site facilitating molecular cloning. For instance, the linker of TB27 comprised EL and the entire 25-AA linker of P450BM3 (Fig. 1). The previous TxtE-BM3R construct contains a 13-AA linker that bears EQ at its *N*-terminus (Fig. S1)²², and is denoted TB13-Q in this study. Superimposition of the crystal structure of TxtE (PDB: 4TPO) with the P450 heme domain of P450BM3 (PDB: 1BVY)³² revealed an additional opportunity to potentially improve chimera's catalytic activity. A loop connecting two helices (J and K, Fig. S2A) of P450BM3 putatively contributes to the interface of heme and reductase domains. This basic-residue-rich loop (A291 to Y313) is about 7 Å away from the acidic residues of the FMN-binding domain and can provide a specific protein-protein interaction for efficient electron transfer (Fig. S2B)³². This stretch of amino acids is significantly different from those in TxtE (Fig. S2A). We thereby hypothesized that replacing those residues in TxtE with the corresponding residues in P450BM3 could increase the coupling efficiency between the domains (Fig. 1). The rationally swapped chimera (TB13S) carrying the JK loop from P450BM3 was constructed with overlapping PCR using *TB13-Q* and *P450BM3* as the templates. All constructs were expressed in *E. coli* and purified to homogeneity by a single Ni²⁺-NTA affinity chromatography²² (Fig. S3). All chimeras were similarly soluble, indicating the minimal effect of variable linker length on the protein solubility. Except TB3, TB6 and TB9, all chimeras showed a peak at around 450 nm in their reduced-carbon monoxide (CO) difference spectra

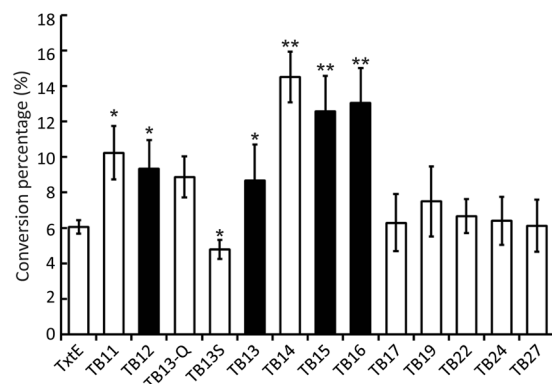


Figure 2. Relative nitration activity of TxtE and chimeric TxtE-BM3R variants. All reactions contained 0.5 mM Trp and 1.5 μ M P450. The TxtE reaction was further supplemented with 0.43 μ M spinach Fer and 0.33 μ M Frd. The reactions were incubated at 20 °C, 300 rpm for 30 minutes. All experiments were repeated at least three times. The results of TB12, TB13, TB15, and TB16 were shown as black bars. Chimeras showed significant difference in nitration activity compared to the wild type TxtE were indicated as * $P < 0.05$ or ** $P < 0.01$.

(Fig. S4)³⁸, indicating the production of active P450s. The concentrations of active chimeras and TxtE in purified proteins were quantitated using an extinction coefficient of $\Delta\epsilon_{450-490\text{nm}}$ of 91 $\text{mM}^{-1} \text{cm}^{-1}$ for the ferrous P450-CO complex.

Nitration performance of chimeric TxtE-BM3R variants. We first measured the catalytic performance of TB11, TB13-Q, TB13S, TB14, TB17, TB19, TB22, TB24, and TB27 to quickly assess the potential effects of linker length (Fig. 2). TxtE coupled with spinach Fer and Frd served as the control. HPLC analysis revealed that all fusion enzymes nitrated Trp to a different extent (Figs S5 and 2). After 30 min, TxtE (1.5 μ M) converted 6.1% of Trp (0.5 mM) into 4- NO_2 -L-Trp, and this rate was lower than all tested chimeric constructs (1.5 μ M). Compared with TxtE, the nitration activity of the most active chimera, TB14, was improved by 2.4 times to reach 14.5% conversion in 30 min. The same was true when the reaction time was extended to 60 min (12.5% for TxtE vs. 28.8% for TB14). Remarkably, TB14 converted 82% of the substrate into the nitro-product after 4 h and its reaction reached the completion at 8 h (Fig. S6). On contrast, the maximal substrate conversion in the TxtE reaction was about 78%, which was achieved after 8 h. The second most active enzyme was TB11 (10.2%), followed by TB13-Q (8.9%). Compared with TxtE, TB13-Q showed a higher initial reaction rate but a lower overall conversion rate (Fig. S6). The nitration activity of chimeras with longer linker length from 17 to 27 AAs were similar (6.2 to 7.5%) and around 2 times lower than TB14 (Fig. 2). Next, we characterized TB12, TB13, TB15, and TB16 to finely examine the optimal length from 11 to 17 AAs (Fig. S1B). The activity of both TB15 and TB16 was higher than TB11 and TB17 and was only slightly lower than TB14 (Fig. 2). On the other hand, TB12 and TB13 retained about 60% of TB14's activity. These results demonstrated the optimal linker length to be 14 to 16 AAs. Moreover, the similar activity of TB13 and TB13-Q (Fig. S1B) suggested that linker content played a minor role in determining activity (Fig. 2), consistent with the conclusions of several previous reports³⁵⁻³⁷. Additionally, standalone BM3R (Fig. S3) in solution was unable to support TxtE for nitration, indicating the necessity of the linker in modulating proper interactions between TxtE and BM3R.

The nitration activity of rationally designed TB13S (4.8%) was lower than TxtE and TB13-Q (Fig. 2). The observed catalytic activity, however, for the first time demonstrated that TxtE is a robust scaffold for chimerogenesis engineering that tolerates foreign structural elements and has potential to greatly expand fitness and improve catalytic properties of nitration biocatalysts^{39, 40}.

To yield mechanistic insights to account for the observed activity changes, we first assessed the binding affinities of Trp toward all active chimeras by UV-Vis difference spectroscopy (Table 1). All fusions showed Type I spectral shift of the Soret peak from 420 to around 390 nm after adding Trp, and demonstrated the overly similar level of binding affinity ($K_d = \sim 20 \mu\text{M}$). This result indicated that linker length or structural swapping minimally influenced substrate binding. We next determined coupling efficiency to evaluate the extent to which linker length and structural swapping may affect electron transfer compatibility during nitration reaction (Table 1). Coupling efficiency of TB11 (5.2%) and TB14 (5.3%) were about 2.2 times higher than TxtE (2.4%), possibly indicating that improvement in their activity was driven by more effective electron transfer. TB15 (3.9%) and TB16 (2.7%) also showed improved coupling efficiency in comparison with TxtE. Coupling efficiency of the other fusions was similar to that of TxtE. Of note, the positive influence of the swapped JK loop from P450BM3 was reflected by 1.3-time improvement in electron transfer efficiency of TB13S (2.6% vs 2.0% of TB13-Q). Furthermore, the NADPH consumption rate of TB13S was the lowest among all active chimeras and TxtE, and was more than 2 times lower than TB13 and TB13-Q. However, the coupling efficiency of even the most active chimera is comparably low^{35, 41}, indicating the great potential of further advancing nitration biocatalysts. Additionally, we determined total turnover number (TTN) as nmol product per nmol P450 in all reactions (Table 1). TB14 had the highest TTN at 707 that was 1.9 times and 2.3 times of values of TxtE and TB13-Q, respectively. TTNs of TB11, TB15, and TB16 were also substantially higher than

Chimeras	K_d (μM)	Coupling (%)	NADPH consumption ($\mu\text{M}/\text{min}$)	TTN
TxtE	24.8 \pm 1.1	2.4 \pm 0.3	42.3 \pm 4.2	378 \pm 17
TB11	16.2 \pm 0.6	5.2 \pm 0.5	32.8 \pm 5.0	535 \pm 28
TB12	22.3 \pm 1.4	1.8 \pm 0.2	86.3 \pm 3.9	332 \pm 40
TB13-Q	21.5 \pm 0.8	2.0 \pm 0.1	73.9 \pm 3.8	308 \pm 18
TB13S	19.3 \pm 0.6	2.6 \pm 0.2	30.8 \pm 3.1	202 \pm 20
TB13	21.3 \pm 0.4	2.2 \pm 0.3	65.7 \pm 4.1	342 \pm 29
TB14	17.4 \pm 1.0	5.3 \pm 0.5	45.5 \pm 4.8	707 \pm 16
TB15	21.4 \pm 0.9	3.9 \pm 0.4	53.6 \pm 5.1	658 \pm 33
TB16	18.9 \pm 1.0	2.7 \pm 0.2	80.5 \pm 4.3	351 \pm 39
TB17	15.5 \pm 0.6	2.1 \pm 0.3	50.1 \pm 4.6	348 \pm 24
TB19	16.0 \pm 0.8	2.4 \pm 0.5	52.1 \pm 6.7	464 \pm 21
TB22	27.1 \pm 1.3	2.4 \pm 0.5	46.3 \pm 6.3	408 \pm 32
TB24	14.9 \pm 0.3	2.3 \pm 0.1	46.4 \pm 3.4	381 \pm 15
TB27	16.2 \pm 1.1	2.0 \pm 0.6	51.1 \pm 7.2	548 \pm 17

Table 1. Binding affinity toward Trp, coupling efficiency, NADPH consumption and total turnover numbers of TxtE and its chimeras. All experiments were performed at least three times.

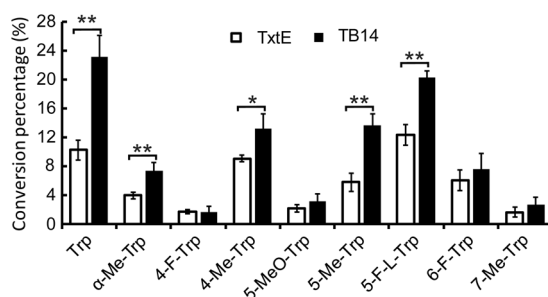


Figure 3. TxtE and TB14 nitrated Trp and its analogues to varying degrees. The reactions were prepared as described previously and incubated at 20 °C, 300 rpm for 45 minutes. All experiments were repeated at least three times. Substrates showed significant difference in the TxtE and TB14 reactions were indicated with * $P < 0.05$ or ** $P < 0.01$.

TxtE, presumably reflecting their higher nitration activity (Fig. 2). We also observed a high TTN of TB27. On the other hand, TTN of TB13S decreased from 308 of TB13-Q to 202, indicating the less efficient use of electrons in its nitration.

Substrate scope of TB14. Next, we attempted to characterize the substrate promiscuity of the most active chimera TB14 as well as TxtE with a library of 29 chemicals that carry substitutions and moderate alterations on the amine, carboxylate, or indole moieties of Trp (Fig. S7). Among these chemicals, 19 has been tested in a recent report with only wild type TxtE⁴². In spectroscopic analysis, 20 Trp analogues (Fig. S7A) induced the Type I spectral shift of TB14 indicating binding, while simple aromatics such as indole and other aromatic amino acids (Fig. S7B) had no detectable level of interaction with the enzymes. There was no significant difference between TB14 and TxtE in terms of binding affinities (Table S1, Figs S8 and S9), further confirming the minimal effect of fused BM3R on substrate binding²² (Table 1). Compounds with modified amine or carboxylate moiety generally showed weaker interactions than those with substitutions on the indole ring, the same observed previously⁴². We then examined the nitration activity of TB14 and TxtE toward all library members. To generate sufficient products from less favourable substrates for HPLC detection, the reactions were incubated for 45 min. Besides Trp and α -Me-DL-Trp, both enzymes nitrated seven Trp analogues with substitutions on the C4, C5, C6, and C7 of the indole (Figs S7A and 3). TB14 showed the higher activity toward all substrates. After 45 mins, both enzymes nitrated a significant amount of 4-Me-DL-Trp, 5-Me-DL-Trp, 5-F-L-Trp and 6-F-DL-Trp but not α -Me-DL-Trp, 4-F-DL-Trp, 5-MeO-DL-Trp and 7-Me-DL-Trp. 5-F-L-Trp emerged as the best substrate of TxtE and was comparable to Trp as the best in the TB14 reaction. Since enzymes nitrate only the L-conformers of racemic substrates, a higher enzyme activity can be expected with enantiopure substrates. The quantitative measurement of both TxtE and TB14 toward the select analogues supported improved performance of TB14 across a variety of substrate analogues and underlined its immediate applications to produce unnatural nitro-Trp compounds. These compounds can have broad applications in synthesizing numerous bioactive peptidic compounds and in analysing macromolecule structures and dynamics⁴³.

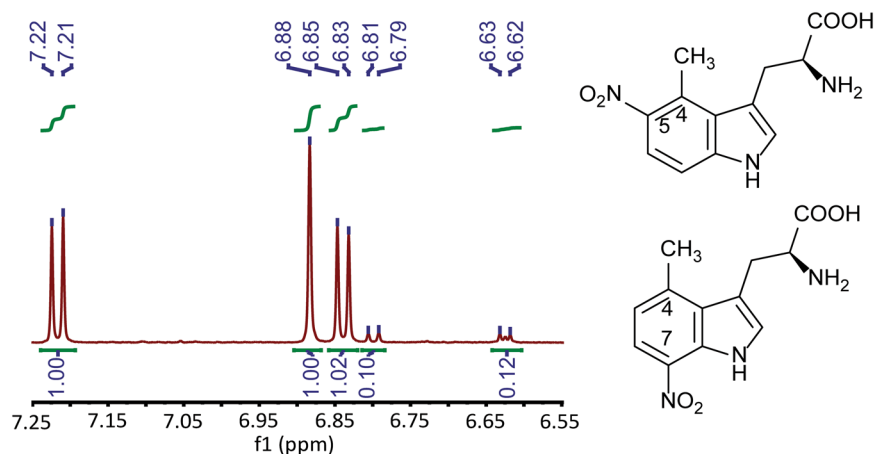


Figure 4. Both 4-Me-5-NO₂-L-Trp and 4-Me-7-NO₂-L-Trp were produced in the TB14 reaction with 4-Me-DL-Trp as substrate. The aromatic region of ¹H NMR spectrum of isolated product showed the chemical shifts of two sets of aromatic protons. Integrated values of these protons were also included.

Regio-promiscuity of TB14 in nitration. In previous study, we have observed remarkable, substrate-tuned regio-specificity of TB13-Q that selectively nitrates C7 when C4 of its substrate indole is occupied by an F substitution²². To further probe the extent to which physicochemical properties of the substituted group at C4 impact enzyme regio-selectivity, we employed TB14 in a scaled-up reaction to nitrate commercially available racemic 4-Me-DL-Trp. Compared with the fluorine replacement, the methyl group is larger in size and is electron donating in nature. Interestingly, the binding affinity of 4-Me-DL-Trp was 13 times higher than 4-F-DL-Trp (Table S1), and TB14 favoured 4-Me-DL-Trp about 8 times more over 4-F-DL-Trp in nitration (Fig. 3).

Despite technical challenges associated with production purification, we isolated about 80 μg of nitrated 4-Me-DL-Trp as a yellow powder by semi-preparative HPLC. By Marfey's derivatization, we observed the significant consumption (~85%) of 4-Me-L-Trp in the reaction and identified the nitro product with L-configuration (Fig. S10). This result further confirmed the strict stereo-selectivity of TB14 in nitration. LC-high resolution (HR) MS analysis of the isolated product confirmed the nitration on the substrate by giving one major peak with the expected molecular weight of nitro product ($m/z = 264.0892$, Fig. S11). Interestingly, one minor peak with the same molecular weight was eluted right after the major one (Fig. S11), suggesting the coexistence of two structural isomers in the product sample.

Next, we structurally characterized the isolated product by ¹H and ¹³C and 2D NMR analysis (Figs S12–S17). From the ¹H and COSY NMR spectra (Figs S12 and S13), we noticed two well separated AX coupling systems in the aromatic region, presumably indicating two nitration sites in the indole. One AX coupling system involved two large doublet peaks at δ 7.22 ppm and 6.84 ppm (Fig. 4) while the other was from another set of two small doublet peaks at δ 6.80 ppm and 6.62 ppm (Fig. 4). The large vicinal coupling constants (8–9 Hz) of the two AX coupling systems suggested the C5 and C7 of the indole as the nitration sites (Table S2, Fig. 4). Two additional AMX coupling systems occurring in the aliphatic region further provided details about the existence of two sets of α and β protons of Trp derivatives (Figs S12 and S13). Given the larger deshielding effect of C5-NO₂ than C7-NO₂ on the C4-Me group, the methyl signal of 4-Me-5-NO₂-L-Trp was in the lower field (δ 2.22 ppm vs 2.10 ppm in 4-Me-7-NO₂-L-Trp, Table S2). Using the integration values of two methyl groups, we determined the molar ratio of 4-Me-5-NO₂-L-Trp: 4-Me-7-NO₂-L-Trp to be about 10:1 (Figs S12 and 4). We were also able to observe the chemical shift signals of 4-Me-5-NO₂-L-Trp in the ¹³C NMR spectrum (Table S2, Fig. S14). The determination of the product structure was then assisted by HSQC and HMBC spectra (Figs S15 and S16). In the HMBC spectrum, the C4-Me group (2.22 ppm) correlated with a significantly deshielded aromatic carbon (142.18 ppm) that became possibly only by the C5-NO₂ group (Figs S16 and S17). Therefore, TB14 carried a striking regio-flexibility in nitrating 4-Me-DL-Trp to produce predominantly 4-Me-5-NO₂-L-Trp and 4-Me-7-NO₂-L-Trp as a minor product (Fig. 4). These results uncovered that different types of substituted groups at C4 of the indole ring can affect key parameters (activity and regio-selectivity) of TxtE biocatalysts. Interestingly, a single residue His176 in the F/G loop of TxtE was computationally identified as a potential determinant of enzyme regio-selectivity, and engineering of this site to some residues (e.g., Phe and Tyr) indeed created TxtE mutants that predominantly produce 5-NO₂-L-Trp⁴⁴. Our results provided new insights into intriguing and synthetically important TxtE's regio-selectivity.

Conclusions

In summary, we created 15 new chimeric TxtE fusion constructs by employing two protein engineering strategies in an attempt to develop novel biocatalytic nitration processes. Rationally swapping the JK loop of TxtE with that of P450BM3 noticeably improved the coupling efficiency of resulted variant (TB13S) but not its overall nitration performance. Varying the linker length between TxtE and BM3R led to the creation and identification of multiple chimeras, particularly TB14, which showed significantly higher catalytic activity than TxtE and TB13-Q²³. Remarkably, TB14 is the most active aromatic nitration biocatalyst ever developed. TB14 is also noteworthy in having broad substrate scope, suggesting its promising uses to synthesize a variety of nitro Trp analogues for

biomedical and biological applications. Furthermore, we demonstrated the production of both 4-F-5-NO₂-L-Trp and 4-Me-7-NO₂-L-Trp by the TB14-driven process. The demonstrated ability of TB14 to nitrate two sites of 4-Me-L-Trp marked TxtE as the promising material for developing nitration biocatalysts to synthesize structurally diverse nitroaromatics via protein engineering. Structural and mechanistic studies to clarify the molecular determinants for TxtE's substrate scope and regio-specificity are under way and can expand the fundamental understanding of enzyme evolution and mechanism and guide further engineering efforts.

Materials and Methods

General Chemicals, DNA Sub-cloning, and Bacterial Strains. Molecular biology reagents and enzymes were purchased from Fisher Scientific. Primers were ordered from Sigma-Aldrich. 4-Me-DL-Tryptophan was from MP Biomedical (Santa Ana, CA), while NOC-5 (3-(Aminopropyl)-1-hydroxy-3-isopropyl-2-oxo-1-triazene) was purchased from EMD Millipore. Other chemicals and solvents were purchased from Sigma-Aldrich and Fisher Scientific. *Escherichia coli* DH5 α and BL21-GOLD (DE3) (Agilent) were used for routine molecular biology studies and protein expression, respectively, and were grown in Luria-Bertani broth or Terrific broth. DNA sequencing was performed at Eurofins. A Shimadzu Prominence UHPLC system (Kyoto, Japan) fitted with an Agilent Poroshell 120 EC-C18 column (2.7 μ m, 3.0 \times 50 mm), coupled with a PDA detector was used for HPLC analysis. For semi-preparative HPLC, YMC-Pack Ph column (5 μ m, 4.6 \times 250 mm) was used. All NMR spectra were recorded in 100 mM DCl in D₂O on a Bruker 600 MHz spectrometer using a 5 mm TXI Cryoprobe in the AMRIS facility at the University of Florida, Gainesville, FL, USA. The instrument was operated at 600.17 MHz for ¹H NMR and 150.9 MHz for ¹³C NMR. Spectroscopy data were collected using Topspin 3.5 software. HRMS data were obtained using a Thermo Fisher Q Exactive Focus mass spectrometer equipped with electrospray probe on Universal Ion Max API source.

Construction of TxtE-BM3R variants. *TxtE* gene was amplified from genomic DNA of *S. scabies* 87.22 (NRRL B-24449) using a pair of SELKnco-F and SELKsac-R primers in PCR reactions (Table S3). The PCR product was analyzed by agarose gel and extracted with a GeneJET Gel Extraction Kit (Thermo). Purified PCR products and pET28b were digested with the restriction enzymes *Nco*I and *Sac*I, and corresponding linear DNAs were ligated to generate expression construct pET28b-TxtEs. To further create the TxtE-BM3R variants with variable linker length, BM3R domain with selected linker lengths was amplified from *P450BM3* gene by a set of primer pairs (Table S3). Purified PCR products and pET28b-TxtE construct were then digested with the restriction enzymes *Sac*I and *Xho*I, and corresponding linear DNAs were ligated to generate pET28b-TxtE-BM3R expression constructs (Fig. S1). To create the P450BM3 standalone reductase (BM3R) expression constructs, *BM3R* gene was amplified using a pair of BM3R-F and BM3R-R primers. Purified PCR products and pET28b were digested with the restriction enzymes *Nde*I and *Xho*I, and corresponding linear DNAs were ligated to create pET28b-BM3R. To create TB13S fusion variant, we used *TB13-Q* as the template and amplified the TxtES fragment with primers of TxtESF and TxtESR (Table S3), and the S13BM3R fragment with primers of FW-V1 and RV (Table S3) in PCR reactions. P450BM3 gene was used as the template to amplify JKL insert using primers of V1-3-F and V1-3-R. After purification, these fragments were fused by overlapping PCR technology. The full-length *TB13S* was cloned into the expression vector as described above. All inserts in the constructs were sequenced to exclude mutations introduced during PCR amplification and gene manipulation.

Heterologous expression and purification of recombinant proteins. Protein expression and purification followed our previous protocols²². The purified proteins were exchanged into storage buffer (25 mM Tris-HCl, pH 8.0, 100 mM NaCl, 3 mM β ME, and 10% glycerol) by PD-10 column, aliquoted and stored at -80 °C until needed. CO difference spectroscopy was used to measure the concentrations of functional P450s³⁸.

Spectral analysis of chimeric TxtE-BM3R variants. Purified TxtE and its chimeric fusions were spectrally analyzed following a previous protocol²². Briefly, the absorbance spectra (400–600 nm) of TxtE and its chimeric fusions in Tris-HCl (25 mM, pH 8) buffer were recorded by a Shimadzu UV2700 dual beam UV-Vis spectrophotometer. The ferric heme of enzymes was then saturated by bubbling carbon monoxide (Airgas) and the spectra of the saturated enzyme solutions were recorded. Sodium dithionite solution (30 μ L, 0.5 M) was then added to reduce ferric ion, and reduced spectra were taken subsequently. CO reduced difference spectra of all P450s were created by subtracting the CO binding spectra from the reduced spectra. Data were further analyzed by Excel. Substrate binding affinities to P450s were measured using 1.5 μ M of enzyme solutions in 25 mM Tris-HCl, pH 8.0²². The changes in absorbance (ΔA) were determined by subtracting the absorbance at ~ 420 nm from that at ~ 390 nm. Data were then fitted to the equation of $\Delta A = \Delta A_{\max}[L]/(K_d + [L])$ using GraphPad Prism 4.

Catalytic activities of chimeric TxtE-BM3R variants. P450 reactions (100 μ L) contained 0.5 mM substrate, 1 mM NADP⁺, 1 mM glucose, ~ 10 units/mL self-prepared glucose dehydrogenase crude extract, 1 mM NOC-5 in 100 μ L of Tris-HCl buffer (100 mM, pH 8.0). As the positive control, the TxtE reaction was also re-constructed in the above mixture further supplemented with 0.43 μ M spinach Fer and 0.33 μ M Frd. The reactions were initiated by adding 1.5 μ M P450s, and incubated at 20 °C, 300 rpm on a thermostat (Eppendorf) for 30 or 45 minutes. Methanol (200 μ L) was then added to stop the reactions. After centrifugation, 10 μ L solutions were analyzed by HPLC. The 4-NO₂-L-tryptophan was synthesized in a large-scale enzymatic reaction to establish a standard curve for product quantification. To determine the coupling efficiency, NADPH (2 mM) was used to replace the NADPH regeneration system in the reaction mixture. NADPH consumption in enzyme reactions was measured at 340 nm ($\epsilon = 6.22$ mM⁻¹cm⁻¹) with a Biotek Synergy HT Multi-Detection Microplate Reader. Non-enzymatic oxidation of NADPH was subtracted as the background. The quantity of nitrated product was

determined by HPLC analysis as described above. Coupling efficiency (%) was determined as product (nmol)/consumed NADPH (nmol) \times 100%. All reactions were independently repeated at least three times. In TTN and coupling efficiency studies, 0.5 μ M P450s and 0.5 mM substrate were used. Total turnover number (TTN) was reported as nmol product per nmol P450. Conversion rate (%) was calculated as product (nmol)/(product + substrate) (nmol) \times 100%. All experiments were performed at least in triplication.

Large-scale enzymatic synthesis of nitrated Trp analogs. To isolate sufficient amount of nitrated 4-Me-Trp analogues for structural determination, 18 μ M TB14 was used in a 10-mL reaction mixture containing 1.5 mM 4-Me-Trp, 3 mM NADP⁺, 3 mM glucose, \sim 30 units/mL self-prepared glucose dehydrogenase crude extract, 3 mM NOC-5 in 100 mM Tris-HCl buffer (pH 8.0). The reactions in a 200-ml flask were incubated at 20 °C, 250 rpm overnight, and then terminated by 20 mL methanol or acidification to pH 1.0 with 6 M HCl. After centrifugation, the supernatants were concentrated in vacuo and then freeze-dried. The products were redissolved in 3 mL methanol for semi-preparation²².

Analytical and semi-preparative HPLC analysis. For analytical analysis, the HPLC column kept at 40 °C. Solvent A and B were water with 0.1% formic acid and acetonitrile with 0.1% formic acid, respectively. The column was eluted first with 1% solvent B for 1 min and then with a linear gradient of 1–20% solvent B in 8 min, followed by another linear gradient of 20–99% solvent B in 2 min. The column was further cleaned with 99% solvent B for 2 min and then re-equilibrated with 1% solvent B for 2 min. The flow rate was set as 1 mL/min, and the products were detected at 211 nm with a PDA detector. For semi-preparative analysis, the column was first eluted with 20% solvent B (acetonitrile with 0.1% formic acid) for 3 min and then with a linear gradient of 20–40% solvent B for 3 min, followed by a linear gradient of 40–99% solvent B for 6 min. The column was then cleaned by 99% solvent B for 2 min and re-equilibrated with 20% solvent B for 3 min. The flow rate was set at 3 mL/min, and the products were detected at 211 nm with a PDA detector. All isolates were combined, concentrated, freeze-dried, and then weighed.

NMR analysis. In NMR analysis, chemical shifts were reported in parts per million (ppm) downfield from tetramethylsilane. Proton coupling patterns were described as singlet (s), doublet (d), double doublet (dd), triplet (t), and multiplet (m). 4-Me-5-nitro-L-tryptophan: ¹H NMR (600 MHz, 100 mM DCl in D₂O) δ 7.22 (d, J = 9.0 Hz, 1H), 6.88 (s, 1H), 6.84 (d, J = 8.9 Hz, 1H), 3.80 (dd, J = 10.1, 5.1 Hz, 1H), 3.22 (dd, J = 15.6, 5.1 Hz, 1H), 2.84 (dd, J = 15.7, 10.1 Hz, 1H), 2.23 (s, 3H). ¹³C NMR (151 MHz, D₂O) δ 170.56, 154.71, 142.20, 138.27, 128.39, 127.74, 124.17, 118.79, 109.77, 109.69, 58.96, 53.66, 27.77, 15.15. HRMS (ESI⁺): calc. for C₁₂H₁₃N₃O₄ [M + H]⁺: 264.0906, found: 264.0892. 4-Me-7-nitro-L-tryptophan: ¹H NMR (600 MHz, 100 mM DCl in D₂O) δ 6.80 (d, J = 8.3 Hz, 1H), 6.63 (d, J = 8.1 Hz, 1H), 3.92 (dd, J = 10.7, 5.4 Hz, 1H), 3.22 (dd, J = 16.2, 5.4 Hz, 1H), 2.98 (dd, J = 16.2, 10.8 Hz, 1H), 2.10 (s, 3H). HRMS (ESI⁺): calc. for C₁₂H₁₃N₃O₄ [M + H]⁺: 264.0906, found: 264.0893.

Marfey's derivatization. 4-Me-Trp and nitrated 4-Me-Trp from enzyme reactions were reacted with Marfey's reagent following manufacture manual (Thermo Scientific). Derivatized products were analyzed by LC-MS with a SHIMADZU Prominence UPLC system fitted with a Waters SymmetryShieldTM RP-C18 column (3.5 μ m, 4.6 \times 100 mm) and a Linear Ion Trap Quadrupole LC/MS/MS Mass Spectrometer system. The flow rate was 0.5 mL/min. The column was eluted with 90% solvent A (0.05 M triethylammonium acetate, pH 3.0), 10% solvent B (acetonitrile) for 2 min and then with a linear gradient of 10–50% solvent B for 60 min. The column was then cleaned by 50% solvent B for 5 min followed by a re-equilibration with 10% solvent B for 2 min. For MS detection, the turbo spray conditions used were: curtain gas: 30 psi; ion spray voltage: 5000 V; temperature: 550 °C; ion source gas 1: 30 psi; ion source gas 2: 20 psi.

Statistical analysis. Student's t-test was used to determine if the nitration activity of TxtE and its chimeras are significantly different from one another. To be considered as being significant, P is <0.05.

References

- Patterson, S. & Wyllie, S. Nitro drugs for the treatment of trypanosomatid diseases: past, present, and future prospects. *Trends Parasitol* **30**, 289–298 (2014).
- Ryan, N. J. & Lo, J. H. Delamanid: first global approval. *Drugs* **74**, 1041–1045 (2014).
- Priotto, G. *et al.* Nifurtimox-eflornithine combination therapy for second-stage African *Trypanosoma brucei* gambiense trypanosomiasis: a multicentre, randomised, phase III, non-inferiority trial. *Lancet* **374**, 56–64 (2009).
- Aneja, R. *et al.* Development of a novel nitro-derivative of noscapine for the potential treatment of drug-resistant ovarian cancer and T-cell lymphoma. *Mol Pharmacol* **69**, 1801–1809 (2006).
- Diacon, A. H. *et al.* Phase II dose-ranging trial of the early bactericidal activity of PA-824. *Antimicrob Agents Chemother* **56**, 3027–3031 (2012).
- Zhou, L., Stewart, G., Rideau, E., Westwood, N. J. & Smith, T. K. A class of 5-nitro-2-furancarboxylamides with potent trypanocidal activity against *Trypanosoma brucei* *in vitro*. *J Med Chem* **56**, 796–806 (2013).
- Ono, N. *The nitro group in organic synthesis*. (Ed. Ono), (Wiley-VCH, 2001).
- Flick, A. C. *et al.* Synthetic approaches to the 2014 new drugs. *Bioorg Med Chem* **24**, 1937–1980 (2016).
- Ding, H. X. *et al.* Synthetic approaches to the 2013 new drugs. *Bioorg Med Chem* **23**, 1895–1922 (2015).
- Yan, G. & Yang, M. Recent advances in the synthesis of aromatic nitro compounds. *Org Biomol Chem* **11**, 2554–2566 (2013).
- Ono, N. *The Nitro Group in Organic Synthesis*. 1, (Ed. Auflage), (John Wiley & Sons, 2003).
- Zolfigol, M. A. *et al.* Design of ionic liquid 3-methyl-1-sulfonic acid imidazolium nitrate as reagent for the nitration of aromatic compounds by *in situ* generation of NO₂ in acidic media. *J Org Chem* **77**, 3640–3645 (2012).
- Aridoss, G. & Laali, K. K. Ethylammonium nitrate (EAN)/Te₂O and EAN/TFAA: Ionic liquid based systems for aromatic nitration. *J Org Chem* **76**, 8088–8094 (2011).
- Laali, K. K. & Gettewert, V. J. Electrophilic nitration of aromatics in ionic liquid solvents. *J Org Chem* **66**, 35–40 (2001).

15. Lancaster, N. L. & Llopis-Mestre, V. Aromatic nitrations in ionic liquids: the importance of cation choice. *Chem Commun* 2812–2813 (2003).
16. Choi, J. M., Han, S. S. & Kim, H. S. Industrial applications of enzyme biocatalysis: Current status and future aspects. *Biotechnol Adv* **33**, 1443–1454 (2015).
17. Yang, G. & Ding, Y. S. Recent advances in biocatalyst discovery, development and applications. *Bioorgan Med Chem* **22**, 5604–5612 (2014).
18. Bornscheuer, U. T. *et al.* Engineering the third wave of biocatalysis. *Nature* **485**, 185–194 (2012).
19. Savile, C. K. *et al.* Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture. *Science* **329**, 305–309 (2010).
20. Clouthier, C. M. & Pelletier, J. N. Expanding the organic toolbox: a guide to integrating biocatalysis in synthesis. *Chem Soc Rev* **41**, 1585–1605 (2012).
21. Parry, R., Nishino, S. & Spain, J. Naturally-occurring nitro compounds. *Nat Prod Rep* **28**, 152–167 (2011).
22. Zuo, R. *et al.* An artificial self-sufficient cytochrome P450 directly nitrates fluorinated tryptophan analogs with a different regioselectivity. *Biotechnol J* **11**, 624–632 (2016).
23. Barry, S. M. *et al.* Cytochrome P450-catalyzed l-tryptophan nitration in thaxtomin phytotoxin biosynthesis. *Nat Chem Biol* **8**, 814–816 (2012).
24. McIntosh, J. A., Farwell, C. C. & Arnold, F. H. Expanding P450 catalytic reaction space through evolution and engineering. *Curr Opin Chem Biol* **19**, 126–134 (2014).
25. Hrycay, E. G. & Bandiera, S. M. The monooxygenase, peroxidase, and peroxygenase properties of cytochrome P450. *Arch Biochem Biophys* **522**, 71–89 (2012).
26. Bernhardt, R. & Urlacher, V. B. Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations. *Appl Microbiol Biot* **98**, 6185–6203 (2014).
27. Sadeghi, S. J. & Gilardi, G. Chimeric P450 enzymes: activity of artificial redox fusions driven by different reductases for biotechnological applications. *Biotechnol Appl Biochem* **60**, 102–110 (2013).
28. Fisher, C. W., Shet, M. S., Caudle, D. L., Martin-Wixtrom, C. A. & Estabrook, R. W. High-level expression in *Escherichia coli* of enzymatically active fusion proteins containing the domains of mammalian cytochromes P450 and NADPH-P450 reductase flavoprotein. *Proc Natl Acad Sci USA* **89**, 10817–10821 (1992).
29. Helvig, C. & Capdevila, J. H. Biochemical characterization of rat P450 2C11 fused to rat or bacterial NADPH-P450 reductase domains. *Biochemistry* **39**, 5196–5205 (2000).
30. Sibbesen, O., De Voss, J. J. & Montellano, P. R. Putidaredoxin reductase-putidaredoxin-cytochrome p450cam triple fusion protein. Construction of a self-sufficient *Escherichia coli* catalytic system. *J Biol Chem* **271**, 22462–22469 (1996).
31. Narhi, L. O. & Fulco, A. J. Identification and characterization of two functional domains in cytochrome P-450BM-3, a catalytically self-sufficient monooxygenase induced by barbiturates in *Bacillus megaterium*. *J Biol Chem* **262**, 6683–6690 (1987).
32. Sevrioukova, I. F., Li, H., Zhang, H., Peterson, J. A. & Poulos, T. L. Structure of a cytochrome P450-redox partner electron-transfer complex. *Proc Natl Acad Sci USA* **96**, 1863–1868 (1999).
33. Roberts, G. A. *et al.* A self-sufficient cytochrome P450 with a primary structural organization that includes a flavin domain and a [2Fe-2S] redox center. *J Biol Chem* **278**, 48914–48920 (2003).
34. Robin, A. *et al.* Engineering and improvement of the efficiency of a chimeric [P450cam-RhFRed reductase domain] enzyme. *Chem Commun* 2478–2480 (2009).
35. Hoffmann, S. M. *et al.* The impact of linker length on P450 fusion constructs: Activity, stability and coupling. *ChemCatChem* **8**, 1591–1597 (2016).
36. Belsare, K. D. *et al.* P-LinK: A method for generating multicomponent cytochrome P450 fusions with variable linker length. *Biotechniques* **57**, 13–20 (2014).
37. Govindaraj, S. & Poulos, T. L. Role of the linker region connecting the reductase and heme domains in cytochrome P450BM-3. *Biochemistry* **34**, 11221–11226 (1995).
38. Omura, T. & Sato, R. The carbon monoxide-binding pigment of liver microsomes. I. Evidence for its hemoprotein nature. *J Biol Chem* **239**, 2370–2378 (1964).
39. Rosic, N. N., Huang, W., Johnston, W. A., DeVoss, J. J. & Gillam, E. M. Extending the diversity of cytochrome P450 enzymes by DNA family shuffling. *Gene* **395**, 40–48 (2007).
40. Lautier, T. *et al.* Ordered chimerogenesis applied to CYP2B P450 enzymes. *Biochim Biophys Acta* **1860**, 1395–1403 (2016).
41. Munro, A. W., Daff, S., Coggins, J. R., Lindsay, J. G. & Chapman, S. K. Probing electron transfer in flavocytochrome P-450 BM3 and its component domains. *Eur J Biochem* **239**, 403–409 (1996).
42. Dodani, S. C. *et al.* Structural, functional, and spectroscopic characterization of the substrate scope of the novel nitrating cytochrome P450 TxtE. *Chembiochem* **15**, 2259–2267 (2014).
43. Royer, C. A. Probing protein folding and conformational transitions with fluorescence. *Chem Rev* **106**, 1769–1784 (2006).
44. Dodani, S. C. *et al.* Discovery of a regioselectivity switch in nitrating P450s guided by molecular dynamics simulations and Markov models. *Nat Chem* **8**, 419–425 (2016).

Acknowledgements

The authors wish to gratefully acknowledge financial support from the University of Florida Opportunity Fund (to Y.D., R.L., and S.B.). This study was partially supported by Air Force Office Scientific Research (FA9550-16-1-0186) and Oak Ridge Associated Universities Ralph E. Powe Junior Faculty Enhancement Award (to Y.D.). We are grateful to Mr. Julius Chai and Ms. Padmavathi Reddy from two local high schools for molecular cloning and protein expression.

Author Contributions

R.Z., and Y.D. conceived and designed experiments. R.Z., and Y.Z. performed the experiments. R.Z., Y.Z., C.J., J.H., R.L., S.B., and Y.D. analysed the data. R.Z., Y.Z., C.J., J.H., S.B. and Y.D. wrote the paper.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-00897-z

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017