Faraday Discussions

Cite this: *Faraday Discuss.*, 2024, **252**, 29

PAPER

Indigo production identifies hotspots in cytochrome P450 BM3 for diversifying aromatic hydroxylation†

Douglas J. Fansher, ‡[abc] Jonathan N. Besna ‡[bcd] and Joelle N. Pelletier *[abcd]

Received 1st February 2024, Accepted 5th February 2024
DOI: 10.1039/d4fd00017j

Evolution of P450 BM3 is a topic of extensive research, but screening the various substrate/reaction combinations remains a time-consuming process. Indigo production has the potential to serve as a simple high-throughput method for reaction screening, as bacterial colonies expressing indigo (+) variants can be visually identified *via* their blue phenotype. Indigo (+) single variants, indigo (−) single variants and a combinatorial library, containing mutations that enable the blue phenotype, were screened for their ability to hydroxylate a panel of 12 aromatic compounds using the 4-aminoantipyrine colorimetric assay. Recombination of indigo (+) single variants to create a multiple-variant library is a particularly useful strategy, as all top performing P450 BM3 variants with high hydroxylation activity were either indigo (+) single variants or contained multiple substitutions. Furthermore, active variants, as determined using the 4-AAP assay, were further characterized and several variants were identified that gave more than 90% conversion with 1,3-dichlorobenzene and predominantly formed 2,6-dichlorophenol; other variants showed significant substrate selectivity. This supports the hypothesis that substitution at positions that enable the indigo (+) phenotype, or hotspot residues, is a general mechanism for increasing aromatic hydroxylation activity. Overall, this research demonstrates that indigo (+) single variants, identified *via* colorimetric colony-based screening, may be recombined to generate a multiply-substituted variant library containing many variants with high aromatic hydroxylation activity. The combination of colony-based screening and other screening assays greatly accelerates enzyme engineering, as readily-identified indigo (+) single variants can be recombined to create a library of active multiple variants without extensive screening of single variants.

[a]*Chemistry Department, Université de Montréal, Montreal, QC, Canada. E-mail: joelle.pelletier@umontreal.ca*
[b]*PROTEO, The Québec Network for Research on Protein, Function, Engineering and Applications, Quebec, QC, Canada*
[c]*CGCC, Center in Green Chemistry and Catalysis, Montreal, QC, Canada*
[d]*Department of Biochemistry and Molecular Medicine, Université de Montréal, Montreal, QC, Canada*

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d4fd00017j

‡ Equally contributing (first) authors.

Introduction

Efforts to functionalize C–H bonds in late-stage processes are highly sought-after for the synthesis of small molecules and pharmaceuticals. However, achieving high yields and desired stereoselectivity using chemical synthesis methods can be challenging.[1] As the demand for cost-effective and environmentally friendly alternatives for small-molecule C–H functionalization increases, it is necessary to develop new biocatalytic methods. Enzyme catalysis offers several advantages over traditional chemical methods,[2–6] including performing reactions under mild conditions that do not require elevated temperatures, extreme pH, or toxic reagents. Enzymes have a main limitation in that their native substrate scope can be limited.[7] To overcome this challenge, enzyme engineering uses natural evolution as inspiration, introducing mutations to create variants with modified properties, such as altered substrate promiscuity, stereoselectivity, or improved stability under operating conditions.

Cytochrome P450 monooxygenases (P450s) are heme-containing enzymes that catalyze a broad range of complex oxidation reactions on organic compounds, including non-activated C–H bond reactions. This makes them ideal for generating fine chemicals such as chiral compounds used in flavors, fragrances, and pharmaceuticals, including new antibiotics and drug precursors.[8,9] The use of P450s for selective oxidation is highly appealing. CYP102A1, also referred to as P450 BM3, is a self-sufficient cytochrome P450 enzyme from *Bacillus megaterium*. The enzyme has gained attention as a biocatalyst for various synthetic applications due to its high stability and catalytic activity.[10,11] P450 BM3 naturally hydroxylates long-chain fatty acids such as lauric acid.[12] Enzyme engineering has expanded its reactivity in recent years, allowing it to catalyze oxidation reactions on non-native substrates, as well as new-to-nature reactions such as carbene[13] and nitrene transfer reactions.[14] Variants are generated and screened against the desired reaction/substrate combination, which is a time-consuming and resource-intensive aspect of enzyme engineering. Several colorimetric assays have been developed for screening P450 BM3 variants.[15–17] However, these assays still require multiple experimental steps before a color is visualized. In many cases, characterization with HPLC, LC-MS, or GC-MS is primarily used, which greatly limits screening throughput.[18–21]

Previous studies have demonstrated that specific P450 BM3 variants can effectively hydroxylate indole, resulting in the formation of indigo dye.[22–32] This is in contrast to the native enzyme, which does not perform this reaction at a readily-detectable level.[22] Indigo-producing variants can be easily identified as dark blue colonies when expressed in *E. coli*.[33] Furthermore, these variants have demonstrated a greater propensity to hydroxylate certain aromatic substrates.[34–37] More broadly, the production of indigo by cytochrome P450 variants can be a successful predictor of promiscuous activity, and some correlations have been established.[30,31,33,38,39] In our earlier work, only indigo-producing P450 BM3 variants were observed to hydroxylate 4-phenylbutan-2-one into 4-(4-hydroxyphenyl)butan-2-one, the prevalent raspberry aroma widely employed in the food and flavor industry,[22] priced near 10 000 USD per kg. A similar colorimetric colony-based screening was used to identify P450 BM3 variants with oxidation activity towards coumarin and 7-ethoxycoumarin.[33] These variants possess catalytic

properties similar to human cytochrome P450 enzymes, which are primarily involved in the metabolism of endogenous and xenobiotic compounds.[33] Because the wild-type P450 BM3 lacks visible indigo production, observation of indigo formation in *E. coli* can serve as a simple primary screen for new activities or as a starting point for further enzyme engineering.

Recently, we used an in-house P450 BM3 variant library to establish a correlation between indigo formation and hydroxylation of the aromatic compounds anisole and naphthalene.[40] In this work, 97 previously unreported variants of P450 BM3 substituted at any of 18 positions resulted in a blue phenotype, with 80 among the variants being well-expressed. These indigo-positive (indigo (+)) variants were significantly more likely than similarly well-expressed indigo-negative (indigo (−)) variants to favor the hydroxylation of anisole and naphthalene, providing up to a three-fold increase in activity compared to wild-type P450 BM3. These limited experiments demonstrated that colony-based screening is an efficient, high-throughput approach to identify P450 BM3 variants having the potential to exhibit new or improved reactivity.

The limited number of substrates used in previous correlation studies highlights the need to examine the broader applicability of indigo production as a proxy assay for predicting the hydroxylation of specific aromatic substrates by P450 BM3 variants. Based on prior information, we hypothesized that indigo formation is associated with generalized substrate promiscuity. Here, we report our screening efforts against a panel of structurally diverse aromatic substrates using a series of indigo (+) and indigo (−) variants. We applied a sequence of three assays that serve different purposes. The indigo formation is a colony-based, high-throughput screen that allowed us to readily identify the top indigo-producers. The limitation of this cellular assay is that only indole can be used as a substrate, such that it does not allow evaluating substrate promiscuity. To evaluate substrate promiscuity, we used the more general 4-AAP assay: it measures phenol formation, thus allowing investigation within the class of aromatic compounds. Finally, an NADPH consumption assay complements the 4-AAP assay since it measures overall reactivity of the enzyme; together they allow determination of the "coupling efficiency" – the fraction of NADPH used for product formation rather than for unproductive conversion of H_2O to H_2O_2. Thus, variants that produce equal levels of phenolic product may do so with different levels of NADPH consumption, demonstrating different coupling efficiencies. Both the 4-AAP and NADPH assays are of moderate throughput. Differences in activity between indigo (+) and (−) variants will be discussed. Finally, regioselectivity and reaction conversions for the hydroxylation of 1,3-dichlorobenzene by P450 BM3 variants will be highlighted.

Results and discussion

P450 BM3 variant-library design and project workflow

With the increasing interest in broadening the scope of reactions that P450 BM3 and its variants can catalyze, there is a growing need for high-throughput screening approaches. Indigo production has the potential to serve as a simple high-throughput method for reaction screening, as bacterial colonies expressing indigo (+) variants can be visually identified *via* their blue phenotype. Variants capable of hydroxylating indole to form indigo have shown the ability to

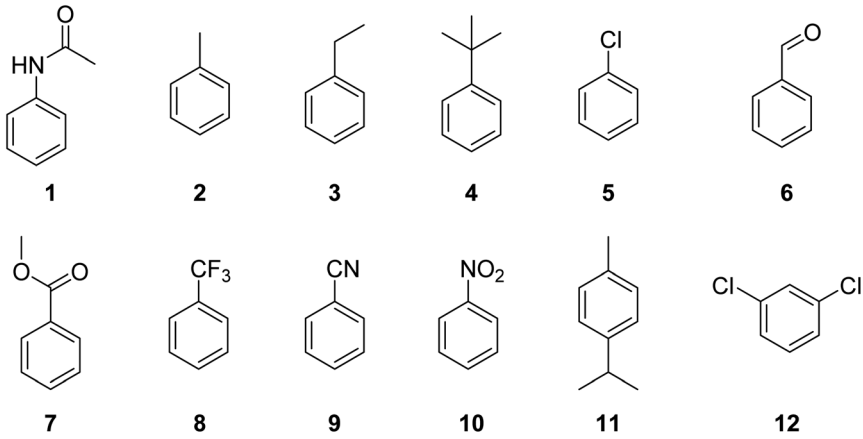


Fig. 1 Panel of substituted benzenes tested as substrates with indigo (+) and indigo (−) P450 BM3 variants.

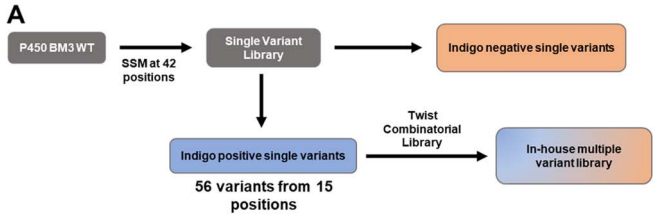
hydroxylate a limited number of other aromatic substrates.[22,34–37] Furthermore, previous studies have shown some success by screening only these indigo (+) variants[33] but the few variants and substrates assayed preclude generalizing a correlation between indigo formation and aromatic hydroxylation.

In this study, we further investigate the importance of indigo-producing variants relative to their indigo-negative counterparts by testing a library of both indigo (+) and indigo (−) variants against a panel of aromatic compounds (Fig. 1). Substituted benzenes with varied steric and electronic properties were selected to test for tolerance of P450 BM3 variants to a wide range of functional groups. Certain P450 BM3 variants can accept these substituted benzenes, most notably **1–5** [37,41–52] and 41–52) and **11**.[18,37,52] However, apart from **11**, only a few variants have been shown to accept these compounds as substrates (ESI Table S1†), suggesting that hydroxylation of substituted benzenes is under-explored. Alternatively, these compounds and their derivatives may not be adequate substrates of previously investigated P450 BM3 variants and have therefore not been reported.

To maximize the probability of identifying P450 BM3 variants capable of accepting **1–12** as substrates for hydroxylation, we used an in-house library comprising point-substituted (single) variants and multiply substituted variants. Investigating variants with diverse mutational patterns and differing numbers of substitutions allows sampling the sequence landscape that enables increased hydroxylation of these aromatic substrates.

Singly-substituted variants were generated using site-saturation mutagenesis (SSM) at 42 distinct active-site residues that lie above the plane of the heme, in the vicinity of the active site.[40] This allowed identification of "single variants" that have the indigo (+) or indigo (−) phenotype when expressed in *E. coli* (Fig. 2A). Among the indigo (+) single variants, 56 were selected that include a substitution at any of 15 positions, including the well-known mutational hot-spot residues A82, F87, A180, M212, R255 and I263 (Fig. 2A). These 56 substitutions were used to design a multiply-substituted, combinatorial variant library (Fig. 2B). The

A



B

Region	Native Residue	Possible Substitutions
B′-helix	L75	F, H, R
	D80	C, N, P
	F81	R, Y
	A82	E, Q
87	F87	G, L, P
F-helix	F173	E, H, P
	M177	H, L, P
	A180	E, F, M, P, W
	L181	F, I, P
	A184	E, F, H, K, R
G-helix	M212	H, K, P, R, T, W
I-helix	R255	D, G, I, M, P, T
	Y256	H, K, V
	I259	G, K, P
	I263	G, H, N, P, Q

C

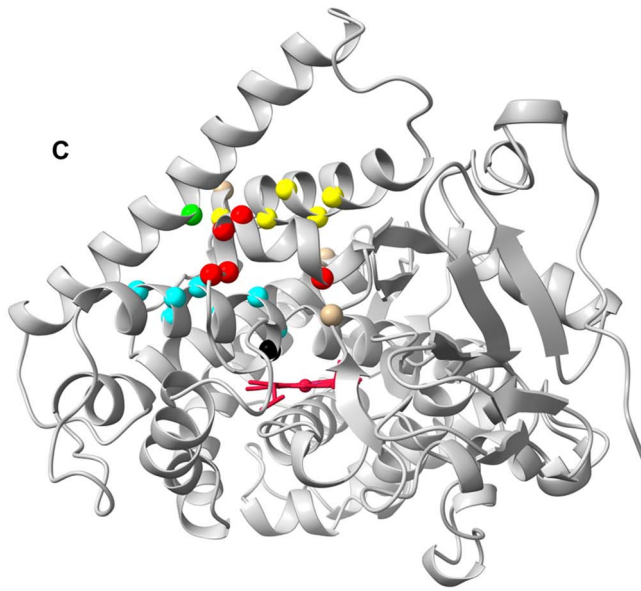
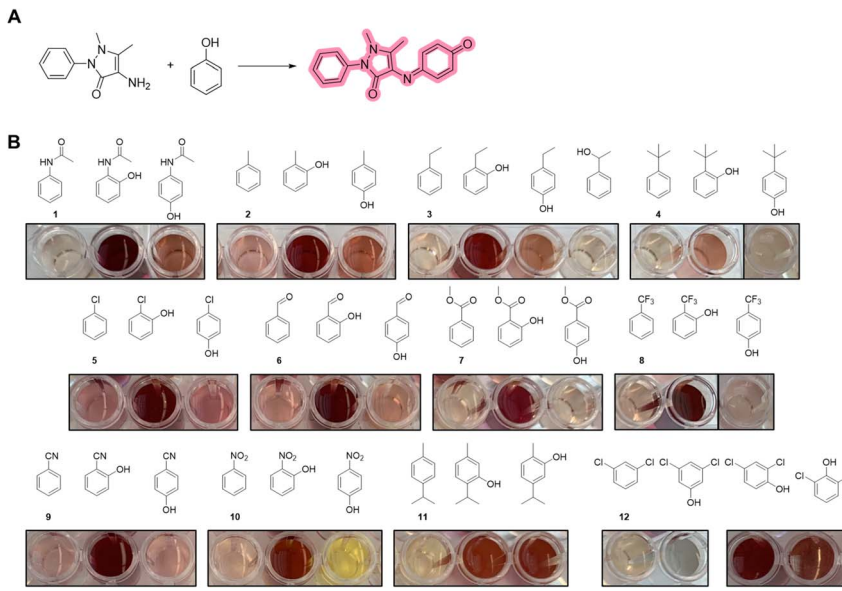


Fig. 2 (A) Engineering strategy for creating the P450 BM3 variant library. Site-saturation mutagenesis (SSM) was performed at 42 positions in the vicinity of the active site; indigo (+) singly-substituted variants (single variants) were then recombined to produce multiply-substituted variants (multiple variants). As a result, multiple variants are all derived from indigo (+) single variants but may have an indigo (+) or indigo (–) phenotype. Indigo (+) and indigo (–) single and multiple variants were then investigated for aromatic hydroxylation. (B) Indigo (+) P450 BM3 substitutions used to create the multiply-substituted variant library. (C) Crystal structure of P450 BM3 (PDB: 1FAG)[53] representing all substituted positions discussed in this article as colored spheres. The colors correspond to the secondary structure elements listed in (B): B′-helix (red), F-helix (yellow), G-helix (green) and I-helix (blue). Position 87 and other unintended positions not listed in (B) are represented by black and beige spheres, respectively.

combinatorial library was designed to afford an average of 4 substitutions per multiply-substituted variant with the majority (96% of the library) having between 2 and 7 substitutions. The positions of all planned substitutions investigated in this study, whether singly or in combination, are highlighted in Fig. 2C; the complete list of variants investigated in this study, including their indigo (+) or (−) phenotype, is found in ESI Table S2.† In addition, three unplanned substitutions were observed that result from errors during library creation. One such example is substitution E143D in variant B10 (identity F81Y/E143D/A180M/M212K; indigo (+) phenotype); the others are T88P (in variant F9) and P172H (in variant G8) (ESI Table S2†).

Expression of the multiply-substituted library in *E. coli* and selection of ∼1050 colonies for the indigo (+)/(−) phenotype showed that only ∼33% of variants exhibited the indigo (+) phenotype. This was unexpected, since all substitutions included in the multiple-variant library were originally selected for their indigo (+) phenotype; we had expected that most of the multiple variants would display straightforward additivity of the indigo (+) phenotype, or that positive sign epistasis would amplify their combination. The observed predominant negative sign epistasis illustrates the complexity of indigo (+) substitutions.

The DNA sequence was analyzed for 192 of the ∼1050 colonies. This confirmed integrity of the library design (barring the 3 unexpected substitutions mentioned above) and the expected distribution of substitutions per variant. The panel of 192 multiple variants sequenced contains mutations at all 15 of the hot spot positions (Fig. 2B and C). Furthermore, many of the indigo (+) variants contained within the single- and multiple-variant libraries are novel and differ considerably from



Scheme 1 4-AAP colorimetric assay. (A) Conjugation of 4-aminoantipyrine (4-AAP) with the representative compound, phenol, to form a colored 4-AAP-phenol conjugate (red/pink) observed at 485 nm (ref. 17). (B) Colors of conjugation products with the substrates **1–12** and their respective hydroxylated products observed at 485 nm.

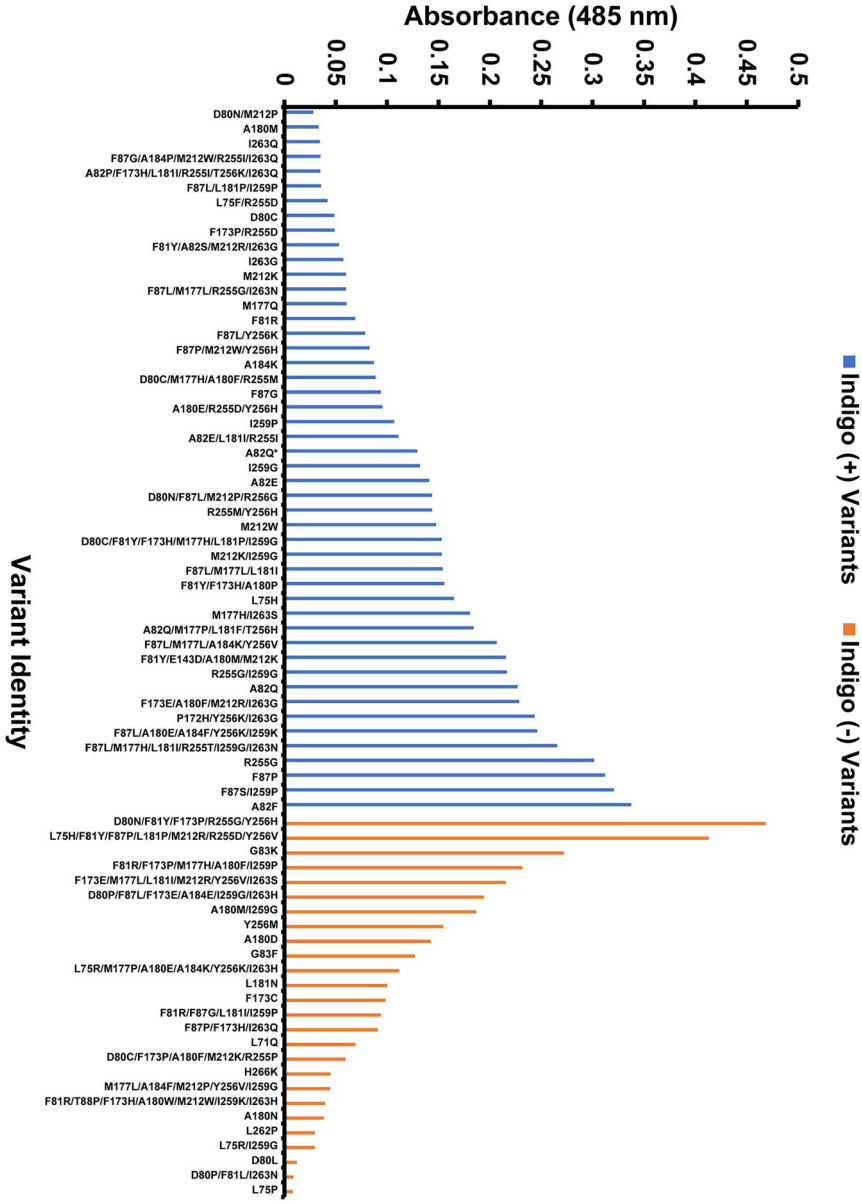


Fig. 3 Activity of indigo (+) (blue) and indigo (−) (orange) P450 BM3 variants towards hydroxylation of **1**. The conjugation of phenolic products to 4-aminoantipyrine was measured at 485 nm. For clarity, of the 91 variants assayed, only the 74 variants showing activity are represented; no wild-type P450 BM3 activity was observed. Note: A82Q* = C-terminal His-tag. All experiments were performed in at least duplicate.

previously reported indigo (+) variants (ESI Table S3†). Substitutions at any of the 15 positions, either singly or in combination, can produce either an indigo (+) or indigo (−) phenotype (ESI Table S2†).

From these 192 variants, we brought 37 unique indigo (+) and 19 unique indigo (−) multiple variants forward for further analysis. The sequence composition of this panel of multiple variants is representative of the larger combinatorial library (ESI Table S2†). We note that A82Q* is a C-terminally His-tagged variant. The His-tag enables purification and immobilization of the enzyme in continuous flow reactions. We included this variant to determine whether addition of the His-tag affects enzyme activity under our operational conditions.

The panel of 56 indigo (+)/(−) multiple variants was complemented by addition of 35 single variants (21 indigo (+) and 14 indigo (−)). The 91 variants of P450 BM3 were assayed against compounds **1–12** to determine the phenolic product(s) formed. We initially performed a rapid colorimetric screen, where the production of phenolic products was estimated using the 4-aminoantipyridine (4-AAP assay) (Scheme 1A).[17] Prior to performing the enzymatic hydroxylation reactions, we verified that none among compounds **1–12** give rise to significant color formation when reacted with 4-AAP (Scheme 1B), thus confirming that unreacted **1–12** will not cause a significant background signal. Using a panel of commercially-sourced phenols representing the most likely products of hydroxylation of compounds **1–12**, we observed that most form colored 4-AAP conjugates (Scheme 1B). Because the intensity and color differ, this assay will serve to compare the reactivity of variants toward one substrate, but should not be used to compare the reactivity of any variant between different substrates.

The results of such a screen are expected to conform to one of the three following scenarios: (A) a demonstration that indigo (+) variants are more active in aromatic hydroxylation than indigo (−) variants; (B) no difference in aromatic hydroxylation activity being seen between the two groups; or (C) indigo (+) variants being less active in aromatic hydroxylation than indigo (−) variants. In addition to determining any correlation between indigo production and aromatic hydroxylation, this analysis enables us to identify variants of interest for further characterization using more sensitive analytical techniques, such as liquid chromatography (LC).

Screening for aromatic hydroxylation

Using the rapid 4-AAP assay, we found that **7**, **11**, and **12** are substrates for the wild-type P450 BM3. These results differ from literature reports demonstrating that **2–5** and **11** are substrates for the wild-type enzyme (ESI Table S1†). This discrepancy is due to many of these studies using LC-MS or GC-MS detection techniques, which are more sensitive than the 4-AAP assay at detecting low levels of phenolic products.[18] Additionally, the 4-AAP assay is not equally sensitive at detecting all phenolic products; the efficiency of 4-aminoantipyrine conjugation

varies as a result of phenolic substitution, as does the absorbance of the resulting conjugate (Scheme 1B). As a result, differences in reaction regioselectivity can lead to different levels of detection. Another important distinction is that WT P450 BM3 mainly hydroxylates the isopropyl group of *p*-cymene (**11**), rather than carrying out ring hydroxylation. As the 4-AAP assay only measures phenolic product formation, it cannot serve as a reporter for these other sites of hydroxylation (see ethylbenzene α-hydroxylation product, Scheme 1B). These results indicate that the 4-AAP assay should be used solely as a semi-quantitative indicator of reactivity in this context.

Nonetheless, the 4-AAP reporter assay allowed ready comparison of the reactivity of variants toward a given substrate. Using the 4-AAP assay, the panel of 91 single and multiple variants (in addition to WT and His-tagged P450BM3), containing both indigo (+) and indigo (−) variants, were tested for reactivity with **1–12** ESI Table S4.† Fig. 3 shows the resulting activities with **1**, stratified according to the (+)/(−) indigo phenotype; the variant identities are given. The reactivities with **2–12** are in ESI Fig. S1–S11.† Furthermore, the data is also provided in a machine-readable format (ESI Table S4†).

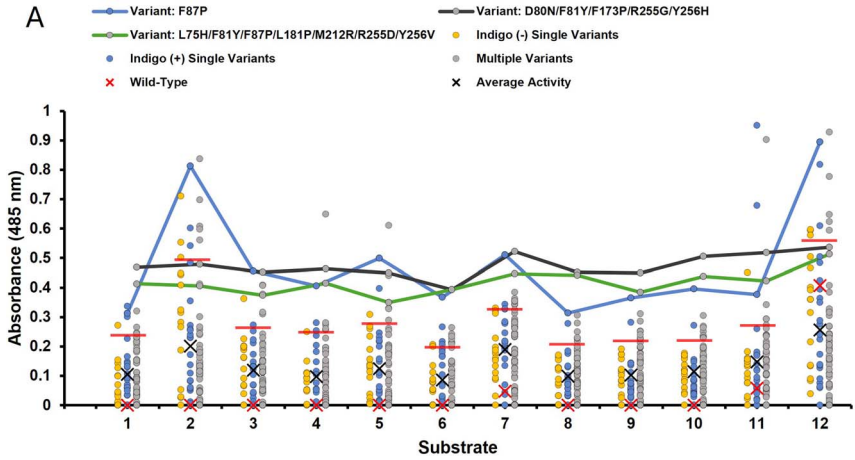
Analysis of these results shows only modest differences in activity between the indigo (+) and indigo (−) populations when assayed against **1–12** and reported with the 4-AAP assay (ESI Fig. S12†). Indigo (+) and indigo (−) variants displayed an equal ability to hydroxylate **1**, **8**, **9**, and **12**. Indigo (+) variants had the highest activity for **2–5**, **11**, and **12**. In contrast, for **1** and **6–10** an indigo (−) variant exhibited the highest level of activity. Compound **11** showed the largest difference in activity between the indigo (+) and indigo (−) variants (ESI Fig. S12†).

Nonetheless, in all cases the differences in average activity observed between the indigo (+) and indigo (−) populations was modest (ESI Fig. S12†). Statistical tests showed no significant difference between these two populations for reaction with any of the 12 compounds. Furthermore, no statistically significant difference was found when comparing the top 5% most active indigo (+) and indigo (−) variants to the rest of the indigo (+) and indigo (−) population for each substrate. This was unsurprising; despite the frequent observation of indigo (+) variants being the most active for a given substrate, many indigo (−) variants exhibited high levels of activity (ESI Fig. S12†).

However, our mutational dataset is biased: all positions that were mutated in this work were initially selected because substitution at each of these positions can give rise to one or more indigo (+) variants. Our results suggest that the positions that were substituted are hotspots where substitution – regardless of the outcome for indigo production – tends to increase reactivity for aromatic hydroxylation relative to the wild-type activity (ESI Fig. S13†). This is suggested upon comparing the single variants: the indigo (+) single variants showed marginally higher activity than the indigo (−) single variants for **1**, **4**, **6**, **7**, and **11**, whereas for **3**, **5**, and **8**, there was little difference in activity between the single variants, and for **2**, **9**, **10**, and **12**, the indigo (−) variants exhibited greater activity than the indigo (+) single variants.

Nonetheless, we observed many cases (substrates **1**, **4**, **8** and **9–11**) where the multiple variants had higher activities than either the indigo (+) or indigo (−) single variants. Furthermore, in all cases except for **3** and **11**, the most active variants all contained more than one substitution (ESI Fig. S13†). To investigate the impact of creating multiple variants, we further stratified the multiple-variant

A



B

	Substrate											
	1	2	3	4	5	6	7	8	9	10	11	12
Indigo (-) Single Variants	1	2	1	-	1	-	-	-	-	-	1	2
Indigo (+) Single Variants	3	3	2	3	2	3	2	2	2	2	3	3
Multiple Variants	6	5	7	7	7	7	8	8	8	8	6	5

Fig. 4 Aromatic hydroxylation activity of P450 BM3 variants as reported with the 4-AAP assay. (A) Activity of P450 BM3 variants towards substituted benzenes (substrates **1–12**) is given as absorbance at 485 nm. Mean absorbance values for each substrate are displayed with a black "X". Wild-type activity is represented with a red "X". Variants above each red line are the 10 most active for that substrate. Dots are colored according to the variant category with indigo (–) single variants, indigo (+) single variants and multiple variants being gold, blue and grey, respectively. The top three variants overall are F87P, D80N/F81Y/F173P/R255G/Y256H, and L75H/F81Y/F87P/L181P/M212R/R255D/Y256V and are represented as blue, black, and green lines, respectively. All experiments were performed at least in duplicate and the average value is shown. (B) Classification of the 10 most active variants for each substrate with the number of indigo (–) single variants (gold), indigo (+) single variants (blue) and multiple variants (grey) for each substrate displayed in the respective substrate columns. The sum of each column is 10.

group into indigo (+) and indigo (–) and compared them to the indigo (+) and indigo (–) single variants (Fig. 4). Many moderately to highly active variants capable of accepting **1–12** as substrates were discovered in the indigo (+) and (–) multiple-variant libraries. Interestingly, indigo (–) single variants typically displayed moderate activity, yet were rarely found to be amongst the most active variants for a particular substrate. Although no clear trend was observed regarding whether indigo (+) variants are more active than indigo (–) variants, this data demonstrates that recombining indigo (+) mutations leads to variants with increased aromatic hydroxylation activities over the respective single variants, regardless of whether the resulting multiple variant is indigo (+) or indigo (–). This supports the hypothesis that substitution at these hotspot residues is a general mechanism for increasing aromatic hydroxylation activity.

Impact of specific substitutions on aromatic hydroxylation of 1–12

Using **1** as a substrate, indigo (+) single variants were, on average, more active than indigo (−) variants and were outperformed by the multiple variants; the most active variant is D80N/F81Y/F173P/R255G/Y256H, which is an indigo (−) multiple variant (Fig. 4). With regards to **2**, indigo (−) single variants were most active on average, yet the most active variant is an indigo (+) multiple variant containing three substitutions, F87P/A180M/I263G. The different variant groups all had similar patterns of activity with **3**, with the two most active variants being F87P and (as for **1**) D80N/F81Y/F173P/R255G/Y256H, which are indigo (+) and indigo (−), respectively.

For **4**, indigo (+) multiple variants had higher average activity than the other three variant groups, with F81Y/E143D/A180M/M212K having the highest activity. The two variants with the next highest activity for **4** were D80N/F81Y/F173P/R255G/Y256H (as for **3**) and L75H/F81Y/F87P/L181P/M212R/R255D/Y256V, which are both indigo (−) multiple variants. The most active variants for **5** and **6** were F87P/A180M/I263G (as for **2**) and D80N/F81Y/F173P/R255G/Y256H (as for **1** and **3**), which are indigo (+) and indigo (−) multiple variants, respectively.

Nearly all 91 variants tested showed a significant improvement in activity for **7** compared to the wild-type P450BM3. This demonstrates that one or a few substitutions at the hot-spot positions suffice to result in a measurable increase in activity relative to the wild-type enzyme. Furthermore, it demonstrates that substitution at many different positions enables this increase in activity. The indigo (−) multiple variant D80N/F81Y/F173P/R255G/Y256H displayed the highest activity with **7** (similar to **1**, **3**, **4** and **6**), outperforming all single variants tested.

Multiple variants all had higher activity with **8–10** than single variants (Fig. 4). In all three cases, the most active variant was again the indigo (−) D80N/F81Y/F173P/R255G/Y256H. With **11**, the top three most active variants were L75H, F87L/M177L/A184K/Y256V, and R255G. These three variants are indigo (+), but there are also many indigo (−) multiple variants that displayed high activities for **11**. Finally, for **12**, the three most active variants were F87P/A180M/I263G (as for **2**, **5** and **6**), F87P (as for **3**) and A184K, which are all indigo (+) variants. Overall, these results demonstrate that many variants can accept a wide range of aromatic compounds having different steric and electronic properties.

We observed several examples of variants that possess enhanced selectivity for certain substrates (ESI Fig. S14A†). If a variant's ranking for a compound is considerably higher (lower number in ranking) than the ranking for other compounds, then we assume that the variant is more selective for that compound. Variant L75H has low activity towards **1–10**, and **12** and ranks poorly when compared the other variants tested (ESI Fig. S14B†). However, it is the most active variant tested against **11**, demonstrating that L75H possesses enhanced selectivity for **11** over the other compounds tested. R255P follows a similar trend, displaying high activity towards **2** (ESI Fig. S14B†). For this variant, little to no activity was detected with other compounds. I258F is selective for **11** in addition to compound **2** (ESI Fig. S14B†). F87P/A180M/I263G is moderately selective for compounds **1–7**, **9** and **12** but displays low activity for the other compounds tested. Furthermore, this variant possesses the highest activities, ranking first for **2**, **5**, and **12** out of all the variants tested. F81Y/E143D/A180M/M212K displayed high selectivity for **4**

and 7 where it was the first and fifth most active variant tested (ESI Fig. S14B†). However, it is ranked 44th with **12**, displaying activity lower than that of wild-type. F81R/F87L/M177L displayed remarkable selectivity for **10**, mainly due its poor reactivity with the other compounds tested. These variants can be highly useful for selective hydroxylation of certain substituted benzenes. This analysis demonstrates that certain variants are selective for specific compounds while F87P, D80N/F81Y/F173P/R255G/Y256H, and L75H/F81Y/F87P/L181P/M212R/R255D/Y256V (Fig. 4A) are universally active across all substrates tested and display promiscuity rather than selectivity. This demonstrates that recombination of indigo (+) mutations is a beneficial approach to enhancing hydroxylation of benzene derivatives to form phenols.

To assess the impact of the substitutions from another angle, we performed comprehensive ranking of the variants based on the 4-AAP assay data with each substrate (ESI Table S5†). The top P450 BM3 variant rankings for each substrate, according to the 4-AAP absorbance results, are listed in (ESI Table S6†). The results reveal clear insights into the impact of substitutions toward different substrates. Notably, the indigo (+) single variant F87P emerged as the top-ranked variant overall (ESI Tables S5 and S7†), showcasing exceptional performance across the entire substrate panel. Further analysis indicates a noteworthy trend, as substitutions at position F87 are consistently included in the multiple variants ranked as the 3rd, 4th, 5th, 7th, and 8th top variants for the different substrates. This suggests a crucial role for substitutions at position 87 in influencing the substrate promiscuity of P450 BM3 variants across a diverse range of aromatic substrates.

In addition to the recurrence of F87 substituents, many of the multiple variants highlighted above had substitutions at R255, F81, Y256, F173 and I263 (in descending order of frequency). These results demonstrate that screening indigo (+) single variants and multiple variants that are derived from indigo (+) mutations is beneficial, as these variants are most active, whether their indigo phenotype is (+) or (−) (ESI Tables S5 and S6†). Furthermore, many highly active variants contain multiple mutations, demonstrating that recombining mutations found in active single variants is a valid strategy for increasing aromatic hydroxylation activity. Many of these substitutions are novel to P450 BM3-catalyzed aromatic hydroxylation reactions (ESI Table S1†). The approach of using indigo-positive variants allowed us to identify amino acid positions (255, 81, 256, and 173) that are infrequently studied compared to positions that are frequently reported in the literature with respect to aromatic hydroxylation (47, 51, 74, 75, 87, 330, and 263) (ESI Table S1†).[54] Combining these mutations has led to variants with high activity towards substituted benzenes, demonstrating utility of the approach. This work showcases that identification of positions where indigo (+) substitutions occur serves to identify hotspots for aromatic hydroxylation; screening single variants made at those positions is likely to yield improvement relative to wild-type activity, and recombining these substitutions leads to many highly active variants.

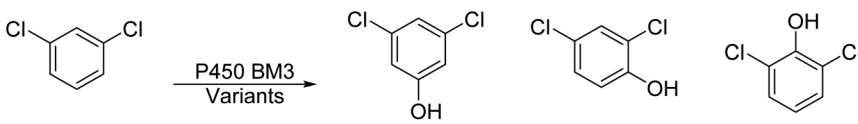
NADPH consumption

NADPH consumption serves as a valuable metric for assessing the activity level of P450 BM3 variants. The rate of NADPH consumption provides insights into the

overall catalytic activity of the enzyme. However, its efficacy as a standalone indicator may be influenced by the coupling efficiency of the reactions; where NADPH consumption is poorly coupled with product formation, correlation with the desired reactivity is weakened. Nonetheless, integrating NADPH consumption data with colorimetric assays measuring product formation offers a high-throughput approach for characterizing P450 BM3 variants.

Analysis of P450 BM3 NADPH consumption rates over 10 minutes yielded no clear difference between the indigo (+) and indigo (−) variants for reaction with substrates **1–12**. These results are consistent with the results of the 4-AAP assay and no statistically significant differences were observed. Indigo (+) variants most often had the highest NADPH consumption rates with **2**, **3**, and **5–12** (ESI Fig. S15†). For example, the highest NADPH consumption rate measured was for variant A184K with compound **6**, at 0.156 μM NADPH per min; this variant was ranked in 14th for activity with compound **6** based on the 4-AAP assay (ESI Table S8†). With **12**, variant A184K is ranked in the top 3 of all variants for both NADPH consumption rates and the 4-AAP assay, but the correlation between the two assays does not always hold. For example, A184K ranked 3rd highest in NADPH consumption for **1** and **11**, but ranked low (46 and 53 of the 91 variants tested) according to the 4-AAP assay. High NADPH consumption and low 4-AAP results suggests low coupling efficiency or no phenolic product formation, as is possible for **11**, where oxidation of the alkyl substituents could occur rather than aromatic hydroxylation. Further observations of high NADPH consumption and lower 4-

Table 1 Reaction conversion and product distribution for the oxidation of 1,3-dichlorobenzyne (**12**) by P450 BM3 variants[a]



12 **12a** **12b** **12c**

Variant identity	Conversion[b] (%)	**12a** (%)	**12b** (%)	**12c** (%)
M212K	98	0	8	92
L71N	93	0	10	90
D80C	91	0	13	87
A180F	70	0	9	91
I263G	68	0	10	90
F87P/A180M/I263G	67	0	16	84
A184K	66	0	9	91
Wild-type P450 BM3	63	0	25	75
A180M	61	0	18	82
F87P	12	0[c]	70	30
His-A82Q[d]	8	0	80	20
F87L/M177L/R255G/I263N	5	0	61	39

[a] Reactions were performed in duplicate with shaking for 3 hours using an NADPH recycling system (G-6-PDH, 9 eq. glucose-6-phosphate). [b] Sum of products. [c] Product **12a** was detected after 6 hours reaction time. [d] C-terminal His_6-tagged A82Q.

AAP activity with A184K were seen for **6–10**, suggesting reaction uncoupling. Low NADPH consumption with moderate activity based on the 4-AAP assay was also seen for variant A184K and substrate **2**. Overall, NADPH consumption is complementary to the 4-AAP assay in providing a distinct assessment of catalytic capacity; both assays offer strengths and have weaknesses, and are readily amenable to a high-throughput format.

P450 BM3 variant expression

If our results suggest that substitution at these 15 hotspots positions tends to increase reactivity for aromatic hydroxylation relative to the wild-type, it could be that these positions have an impact on expression of well-folded, active enzyme. Higher expression of indigo (+) variants over indigo (−) has been observed previously, on a small scale.[55] Here, we observed that the expression of well-folded indigo (+) variants was significantly higher than that of indigo (−) variants (ESI Table S9 and ESI Fig. S16†); well-folded enzyme is observed *via* the CO-difference assay, which quantifies the concentration of heme-bound enzyme. As the scale of enzymatic reactions grows, so does the need to produce higher amounts of protein. Our analysis of 91 variants of P450 BM3 shows that on average, indigo (+) variants provide higher expression of well-folded protein and are thus better suited for reaction scale-up over indigo (−) variants, assuming similar activity levels.

We note that, throughout this work, the indigo (+) phenotype was essentially stable. In rare cases, a variant having been previously observed to be indigo (+) in the presence of indole appeared indigo (−); we minimized the impact of such observations by performing a sufficient number of experimental replicas. Such rare observations are likely related to expression level: when expressed in liquid culture in deep-well plates – where oxygenation is poor – expression can be weaker in some experiments as measured by the CO spectrophotometric assay, indicating that it is a systematic error and not a property of the variants. Under those conditions, a direct link is seen between poor expression and apparent "loss" of the indigo phenotype.

Overall, by screening a library of P450 BM3 variants, we identified highly active variants capable of hydroxylating compounds **1–12**. Variants containing multiple mutations were often the most active for a particular substrate, with indigo (+) single variants occasionally being the most active.

Reaction conversion and product distribution of active P450 BM3 variants

We selected substrate **12** (1,3-dichlorobenzene) to determine reaction conversion and regioselectivity with some of the active variants identified during initial screening. Compound **12** was chosen, since many variants were active with this substrate and three hydroxylation products can be formed, allowing for the possibility of identifying variants with differing regioselectivity. We utilized a cofactor-recycling system to regenerate NADPH, ensuring that the cofactor was not limiting. This regeneration process was carried out by glucose-6-phosphate dehydrogenase, using glucose-6-phosphate as a substrate to convert $NADP^+$ into NADPH and 6-phospho-gluconolactone.[56,57]

Reactions were performed using a 9-fold excess of glucose-6-phosphate (36 mM) over 1,3-dichlorobenzene (4.0 mM). A side-product of many P450 reactions,

including P450 BM3, is uncoupling of the reaction cycle to produce H_2O_2. As mentioned above, this uncoupling reaction is unproductive, as it consumes NADPH yet no product is formed; in addition, the H_2O_2 produced decreases enzyme stability.[58] The large excess of glucose-6-phosphate used to recycle NADPH allows variants that are active and stable, but have lower coupling efficiencies, to produce phenolic products without being limited by NADPH.

We selected P450 BM3 variants that were active for **12** (ESI Fig. S11†). The panel of variants consisted of nine single variants (M212K, L71N, D80C, A180F, I263G, A184K, A180M, F87P and A82Q) and two multiple variants (F87P/A180M/I263G and F87L/M177L/R255G/I263N). L71N is an indigo (−) variant while the others are indigo (+) variants. Under the reaction conditions, wild-type P450 BM3 displayed modest conversion (63%) using 1,3-dichlorobenzene, with 2,6-dichlorophenol (**12c**) being the major product (Table 1). However, the regioselectivity of the wild-type enzyme was significantly lower than that of the majority of variants tested, with as much as 20% of the products formed being 3,5-dichlorophenol. M212K proved to be the most active variant, with a reaction conversion of 98%, predominantly forming 2,6-dichlorophenol (Table 1). L71N and D80C also showed high reaction conversions (93% and 91%, respectively). The regioselectivity was similar to that of M212K, where 2,6-dichlorophenol is the major product and 2,4-dichlorophenol (**12b**) is the minor product. The reaction conversions for variants A180F, I263G, F87P/A180M/I263G, A184K, and A180M ranged from 61–70%. Variants F87P, His-A82Q, and F87L/M177L/R255G/I263N had significantly lower reaction conversions, ranging from 5–12%. However, for these variants, the major product was 2,4-dichlorophenol, in contrast with the other variants tested. No 3,5-dichlorophenol (**12a**) was detected for any of the variants after 3 hours of reaction time. However, detectable levels of 3,5-dichlorophenol were observed for F87P after 6 hours of reaction time.

Overall, most of the variants had moderate to excellent reaction conversions and predominantly formed compound 2,6-dichlorophenol (**12c**). The differences in conversion and regioselectivity between the wild-type enzyme P450 BM3 and the variants tested demonstrates that some substitutions enhance the selectivity for hydroxylation of 1,3-dichlorophenol, as is the case for M212K, L71N, D80C, and I263G. Substitutions at position 180, such as A180M, have a null effect on activity and selectivity as these values were like those of the wild-type enzyme. However, A180F did see modest improvements in both conversion and regioselectivity.

The process we adopted in this work is a balance of screening throughput and how quantitative a test is. We initially focused on identifying variants unlikely to be of interest (indigo colony screen) and then used more quantitative methods (4-AAP and NADPH consumption) on fewer variants but with a greater variety of compounds. Finally, we measured reaction conversion and regioselectivity for the variants of interest with analytical methods. When comparing reaction conversions to the 4-AAP assay results, no discernible trend was observed regarding an increase in absorbance at 485 nm compared to reaction conversion (ESI Table S10†). The amount of conjugation of phenolic compounds to 4-aminoantipyrine differs between regioisomers and this can lead to differences in the level of detection (Scheme 1B). Furthermore, differences in the extinction coefficients of the 4-aminoantipyrine-phenolic product conjugates are likely to exist. This means that as the ratio of regioisomers produced by P450 BM3 variants changes, the measurements from the 4-AAP assay may differ from those determined *via* LC.

However, this cannot account for the range of 4-AAP assay values obtained for the four variants giving the highest conversion, since they exhibit similar regioselectivity. Another possible explanation is that the 4-AAP reaction conditions use lower levels of NADPH (400 μM) and shorter reaction times than the glucose-6-phosphate dehydrogenase-coupled assay. Longer reaction times and excess NADPH, used to measure conversion, provide more time for less active but more stable variants to outcompete the more active but less stable ones. Further analysis of the variants is necessary to determine the underlying reason for the difference in activity levels. This will be explored in the future. Despite its shortcomings, the 4-AAP assay provides a quick assessment of activity and separates variants into active and inactive groups. The active group, or a subset, can be further characterized using more accurate analytical techniques, as shown here, ultimately allowing identification of variants that were highly active for hydroxylation of 1,3-dichlorobenzene and selectively formed 2,6-dichlorophenol as the major product.

Conclusions

In this study, over 2000 datapoints (including duplicates) relating sequence and enzyme activity were generated *via* the 4-AAP reporter assay. The data are provided in a machine-readable format (ESI Table S4†) for easier incorporation into enzymology databases, such as the RetroBioCat database[59] or P450 BM3 variant database.[54] Such datasets relating sequence and a proxy for activity prove invaluable to researchers aiming to create predictive machine learning algorithms for enzyme engineering, as they require accurate and high-quality datasets.[59]

The prospect of using a colorimetric colony-based screening assay to detect variants capable of aromatic hydroxylation would greatly enhance our ability to screen variants in a high-throughput fashion. This work provides valuable insight into the properties of indigo (−) and indigo (+) variants towards hydroxylation of aromatic substrates. No significant correlation in activity was observed for indigo (+) variants being more active than indigo (−) variants. This conclusion is similar to a previous report that found no correlation between blue colonies and hydroxylation of valencene, but that compound differs structurally from the compounds used in this study as well as from indigo.[60] Nonetheless, we have demonstrated that recombination of indigo (+) single variants to create a multiple-variant library is a particularly useful strategy, as all top performing P450 BM3 variants were either indigo (+) single variants or contained multiple substitutions generated by combining indigo (+) single variants. Furthermore, in most cases, multiple variants had the highest activity, with both indigo (−) and indigo (+) multiple variants being equally represented among the top variants. Active variants, as determined using the 4-AAP assay, should be further characterized using analytical approaches as demonstrated here.

By using this approach, we readily identified variants giving significantly higher conversion than the wild-type and more than 90% hydroxylation of 1,3-dichlorobenzene. Overall, this research demonstrates the indigo (+) single variants, identified by colorimetric colony-based screening, may be recombined to generate a multiple-variant library that yields many variants with high aromatic hydroxylation activity toward reagents having diverse substituents. This approach

greatly accelerates enzyme engineering, as readily-identified indigo (+) single variants can be recombined to create a library of active multiple variants without extensive screening of single variants.

Experimental

General information

The WT P450 BM3 in the pcWORI vector was provided by Prof. Frances Arnold (Caltech, USA). Substituted benzenes, phenolic products, and glucose-6-phosphate were purchased from AK Scientific or Fisher Scientific. NADPH was purchased from Enzo Life Sciences. Catalase and glucose-6-phosphate dehydrogenase were purchased from Sigma Aldrich. DNA sequencing was performed at the Génome Québec Innovation Center of McGill University. Box plots were created with Excel using the default parameters. The "whiskers" extend to the largest value that is less than or equal to 1.5 times the interquartile range. Outliers, plotted as circle data points, represent data greater than 1.5 times the interquartile range.

Variant-library generation and storage

The general engineering strategy for P450 BM3 is depicted in Fig. 2A. Initially, a library of singly-substituted variants of P450 BM3 was generated through site-saturation mutagenesis (SSM) at 42 active-site residues, encompassing residues within a 20 Å radius above the heme-iron. The megaprimer method was employed, utilizing four sets of primers (NDT, VMA, ATG, or TGG) per position to encode each amino acid without redundancy or stop codons, according to a previously reported procedure.[22] Expression in *Escherichia coli* DH5α allowed identification of well-expressed indigo producers. Among these, 56 indigo (+) variants were selected; each includes a single substitution at any of 15 positions, as listed in Fig. 2B. The identity of the 56 indigo (+) single variants was utilized to construct a combinatorial library that was cloned into the pCWori vector (Twist Bioscience, USA). Knowing that the library was designed to encode 73% wild-type at each of the 15 target positions, the theoretical diversity of the combinatorial library was 8×10^9 variants; the average number of substitutions per variant is expected to be near 4 with the majority (96% of variants) having 2 to 7 substitutions. Following transformation of the combinatorial library DNA, approximately 1×10^9 primary transformants were obtained by the supplier, from which the pooled plasmid DNA was obtained. The combinatorial library plasmid DNA was transformed into electrocompetent *E. coli* DH5α and plated on Luria–Bertani (LB) agar with ampicillin (100 μg mL^{-1}). Over 10 000 colonies were obtained. Over 1000 well-isolated colonies were selected for phenotypic determination (indigo (+) or indigo (−)) by colorimetric colony screening, of which 192 were analyzed by DNA sequencing.

Site-directed mutagenesis of P450 BM3 (CYP102A1)

Site-directed mutagenesis was performed to introduce the A82Q mutation into the wild-type P450 BM3 (CYP102A1) gene. The mutagenesis was conducted using traditional PCR methods. The forward primer used was A82Q FWD with the following sequence: ACAAATTTAAGCGCTTGACTTAAGTTTTTATCAAAGCGTG

and the reverse primer used was A82Q REV with the following sequence: ACGTGATTTTCAGGGAGACGGGTTATTTACAAGC. PCR amplification was carried out using Phusion Plus polymerase. The reaction conditions were set with an annealing temperature of 60 °C and a total of 35 cycles.

Expression of His-tagged A82Q in *E. coli* C41 DE3

The mutated P450 BM3 gene was cloned into the pET-15b vector and transformed into *E. coli* C41 DE3 for expression. Starter cultures were grown in LB medium with ampicillin (100 μg mL^{-1}) at 37 °C with shaking at 250 rpm for 16 hours. For protein expression, the starter culture was transferred to 500 mL of Terrific Broth (TB) supplemented with ampicillin (100 μg mL^{-1}), δ-aminolevulinic acid (12.5 μg mL^{-1}), thiamine (5 μg mL^{-1}), trace metals, and Na_2MoO_4. The culture was grown until an optical density (OD) of 0.7–1.0 was reached, followed by induction with 1 mM IPTG and further incubation for 24 hours at 25 °C with shaking. Cells were harvested by centrifugation, and the pellets were frozen at −80 °C for at least 2 hours. These steps were performed according to a previously reported procedure.[61]

Colorimetric colony screening for indigo formation

Cultures of *E. coli* DH5α harbouring P450 BM3 variants (incubated overnight at 37 °C) were plated on ZYP5052 agar plates (autoinducing plates) containing indole (0.5 mM in DMSO, 0.2% final DMSO concentration) and ampicillin (100 μg mL^{-1}). Plates were incubated at 37 °C for 8 hours followed by further incubation at 30 °C for 40 hours. Plates were stored at 4 °C to allow color development.

Enzyme expression and storage

Enzyme expression was performed according to a previously reported procedure[22] with the following changes: (1) Erlenmeyer flasks (125 or 250 mL) containing 30 mL of TB media were used instead of deep-well plates. This increased expression per mL and reduced batch-to-batch variations in variant expression. (2) The expression temperature was set at 30 °C for 18 hours. (3) Cultures were separated into two tubes (15 mL culture per tube) and pelleted for storage until use.

Enzyme preparation

The cell pellet, from 15 mL of *E. coli* culture, was thawed on ice and resuspended in potassium phosphate buffer (100 mM, pH 8.0, 2.5 mL) containing lysozyme (0.5 mg mL^{-1}) and incubated at room temperature for 30 min with shaking. The suspension was sonicated using a Branson Digital Sonifier 250 (micro-tip, 30% amplitude, 45 s ON/45 s OFF, 2 × 30 s ON/30 s OFF) for a total of 3 cycles. Cell debris was removed by centrifugation at 13 623*g* for 10 minutes at 4 °C and the resulting clarified cell lysate was used for subsequent steps.

CO-difference assay

The concentration of well-folded P450 enzyme in clarified cell lysates was determined using a CO binding assay as previously described.[22] In brief, in a 96-well plate, 40 μL of 0.3 M sodium dithionite was added to clarified *E. coli* lysate (160 μL) and a measurement at 450 nm was taken. The plate was placed in a CO-purged chamber for 5 minutes, after which a second measurement at 450 nm was taken.

Enzyme concentrations were determined using the difference between the measurements and an extinction coefficient (0.091 M^{-1} cm^{-1}). The variant concentrations were diluted to 0.13 μM for use in further experiments.

General procedure for screening P450 BM3 activity with substituted benzenes: NADPH consumption and 4-AAP assay

Assays were conducted at room temperature in 96-well plates. Clarified cell lysate (0.13 μM P450 BM3, 100 μL), substrate (440 mM stock in DMSO, 20 mM final concentration) and potassium phosphate buffer (100 mM, pH 8.0, 95 μL) were added to a well and incubated at room temperature for 5 min. NADPH (15 μL of 440 mM stock, 409 μM final concentration) was added and NADPH consumption was measured at 340 nm for 10 min. After a further 50 min of reaction time (60 min total), the enzymatic reactions were quenched with 25 μL urea (4 M urea in 0.1 M NaOH). 4-Aminoantipyrine[17] (20 μL, 5 mg mL^{-1} stock) and $K_2S_2O_8$ (20 μL, 5 mg mL^{-1} stock) were then added in rapid succession and incubated for 60 min for the color to develop. Absorbance was measured at 485 using a microplate reader.

LC characterization of P450 BM3-catalyzed hydroxylation of 1,3-dichlorobenzene (12)

Sample preparation. P450 BM3 variants identified during initial screening with the 4-AAP assay were characterized with 1,3-dichlorobenzene (**12**) using an enzymatic NADPH recycling system. Each reaction contained clarified cell lysate (100 mM KPO_4 buffer, with 0.13 μM P450 BM3), glucose-6-phosphate (35.7 mM), catalase (1260 U mL^{-1}), $NADP^+$ (5.64 mM), NADPH (780 μM) and glucose-6-phosphate dehydrogenase (2.7 U mL^{-1}), and was initiated with addition of 1,3-dichlorobenzene (4.0 mM, 440 mM stock in DMSO). The final volume for the reaction was 1115 μL, and total DMSO in the reaction was less than 1% v/v. Aliquots (100 μL) were removed at various time points and quenched with HPLC-grade methanol (900 μL). The reaction mixture was incubated at room temperature for 10 min to precipitate the enzyme followed by centrifugation at 13 623*g* for 10 min. From the clarified reaction mixture, 700 μL was transferred to an HPLC vial for characterization by LC.

LC characterization. LC was performed at University de Montréal Regional Mass Spectrometry Centre using an Agilent 1200 coupled to an Agilent

Table 2 LC method for characterization of the enzymatic conversion of 1,3-dichlorobenzene (**12**) into 3,5-dichlorophenol, 2,4-dichlorophenol or 2,6-dichlorophenol by P450 BM3 variants

Time	H_2O (%)	MeOH (%)	Flow (mL min^{-1})
0	70	30	0.8
8	5	95	0.8
9	5	95	0.8
9.5	70	30	0.8
14	70	30	0.8

Technologies 6110 Quadrupole LC/MS using a Symmetry C18 column (3.5 μm, 4.6 × 75 mm, Part No. WAT066224). Only LC was used; no mass analysis was performed. Substrate consumption and product(s) formation were determined by integrating peak areas at 214.4 nm. Integration values were compared to calibration curve(s) of substrate and products to determine conversion (ESI Fig. S17†). The LC method is depicted in Table 2.

Author contributions

DF contributed to conceptualization. DF and JB contributed to formal analysis, investigation, methodology, visualization, and writing – original draft. JP contributed to formal analysis and funding acquisition. All authors contributed to review and editing and have given approval to the final manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This project was funded by the Natural Science and Engineering Research Council of Canada (NSERC) discovery grant RGPIN-N-2018-04686 and the Canada Research Chair in Engineering of Applied Proteins CRC-2020-00171 to JNP. We thank Alexandra Furtos for access to LC-MS instrumentation and Ali Fendri for fruitful discussion. JNB acknowledges the support of the NSERC-funded CREATE-APRENTICE program.

Notes and references

1 R. K. Zhang, X. Huang and F. H. Arnold, *Curr. Opin. Chem. Biol.*, 2019, **49**, 67–75.
2 M. S. Barber, U. Giesecke, A. Reichert and W. Minas, *Molecular Biotechnolgy of Fungal Beta-Lactam Antibiotics and Related Peptide Synthetases*, 2004, pp. 179–215.
3 S. V. Ranganathan, S. L. Narasimhan and K. Muthukumar, *Bioresour. Technol.*, 2008, **99**, 3975–3981.
4 S. Wu, R. Snajdrova, J. C. Moore, K. Baldenius and U. T. Bornscheuer, *Angew. Chem., Int. Ed.*, 2021, **60**, 88–119.
5 R. N. Patel, *Expert Opin. Drug Discovery*, 2008, **3**, 187–245.
6 A. Bezborodov and N. Zagustina, *Appl. Biochem. Microbiol.*, 2016, **52**, 237–249.
7 S. M. Thomas, R. DiCosimo and V. Nagarajan, *Trends Biotechnol.*, 2002, **20**, 238–242.
8 S. Schulz, M. Girhard and V. B. Urlacher, *ChemCatChem*, 2012, **4**, 1889–1895.
9 V. V. Shumyantseva, T. V. Bulko, V. B. Lisitsyna, V. B. Urlacher, A. V. Kuzikov, E. V. Suprun and A. I. Archakov, *Biophysics*, 2013, **58**, 349–354.
10 D. J. Cook, J. D. Finnigan, K. Cook, G. W. Black and S. J. Charnock, in *Advances in Protein Chemistry and Structural Biology*, ed. C. Z. Christov, Academic Press, 2016, vol. 105, pp. 105–126.

11 E. O'Reilly, V. Köhler, S. L. Flitsch and N. J. Turner, *Chem. Commun.*, 2011, **47**, 2490–2501.
12 C. F. Butler, C. Peet, A. E. Mason, M. W. Voice, D. Leys and A. W. Munro, *J. Biol. Chem.*, 2013, **288**, 25387–25399.
13 P. S. Coelho, E. M. Brustad, A. Kannan and F. H. Arnold, *Science*, 2013, **339**, 307–310.
14 J. A. McIntosh, P. S. Coelho, C. C. Farwell, Z. J. Wang, J. C. Lewis, T. R. Brown and F. H. Arnold, *Angew Chem. Int. Ed. Engl.*, 2013, **52**, 9309–9312.
15 M. W. Peters, P. Meinhold, A. Glieder and F. H. Arnold, *J. Am. Chem. Soc.*, 2003, **125**, 13442–13450.
16 M. Alcalde, E. T. Farinas and F. H. Arnold, *SLAS Discovery*, 2004, **9**, 141–146.
17 T. S. Wong, N. Wu, D. Roccatano, M. Zacharias and U. Schwaneberg, *SLAS Discovery*, 2005, **10**, 246–252.
18 J. Ikebe, M. Suzuki, A. Komori, K. Kobayashi and T. Kameda, *Sci. Rep.*, 2021, **11**, 19004.
19 Y. Zhang, Z. Xiong, Y. Li, M. Wilson, K. E. Christensen, E. Jaques, P. Hernández-Lladó, J. Robertson and L. L. Wong, *Nat. Synth.*, 2022, **1**, 936–945.
20 R.-J. Li, K. Tian, X. Li, A. R. Gaikaiwari and Z. Li, *ACS Catal.*, 2022, **12**, 5939–5948, DOI: **10.1021/acscatal.1c06011**.
21 J. Reinen, J. S. van Leeuwen, Y. Li, L. Sun, P. D. Grootenhuis, C. J. Decker, J. Saunders, N. P. Vermeulen and J. N. Commandeur, *Drug Metab. Dispos.*, 2011, **39**, 1568–1576.
22 O. Rousseau, M. C. C. J. C. Ebert, D. Quaglia, A. Fendri, A. H. Parisien, J. N. Besna, S. Iyathurai and J. N. Pelletier, *ChemCatChem*, 2020, **12**, 837–845.
23 Q. S. Li, J. Ogawa, R. D. Schmid and S. Shimizu, *Biosci., Biotechnol., Biochem.*, 2005, **69**, 293–300.
24 Y. Lu and L. Mei, *J. Ind. Microbiol. Biotechnol.*, 2007, **34**, 247–253.
25 S. Meng, Z. Li, Y. Ji, A. J. Ruff, L. Liu, M. D. Davari and U. Schwaneberg, *Chin. J. Catal.*, 2023, **49**, 81–90.
26 L. Ma, T. Sun, Y. Liu, Y. Zhao, X. Liu, Y. Li, X. Chen, L. Cao, Q. Kang, J. Guo, L. Du, W. Wang and S. Li, *Synth. Syst. Biotechnol.*, 2023, **8**, 452–461.
27 L. Ma, F. Li, X. Zhang, H. Chen, Q. Huang, J. Su, X. Liu, T. Sun, B. Fang, K. Liu, D. Tang, D. Wu, W. Zhang, L. Du and S. Li, *Sci. China: Life Sci.*, 2022, **65**, 550–560.
28 J. Mendoza-Avila, K. Chauhan and R. Vazquez-Duhalt, *Dyes Pigm.*, 2020, **178**, 108384.
29 F. Kong, J. Chen, X. Qin, C. Liu, Y. Jiang, L. Ma, H. Xu, S. Li and Z. Cong, *ChemCatChem*, 2022, **14**, e202201151.
30 Q.-S. Li, U. Schwaneberg, P. Fischer and R. D. Schmid, *Chem.–Eur. J.*, 2000, **6**, 1531–1536.
31 H. M. Li, L. H. Mei, V. B. Urlacher and R. D. Schmid, *Appl. Biochem. Biotechnol.*, 2008, **144**, 27–36.
32 Z. Pengpai, H. Sheng, M. Lehe, L. Yinlin, J. Zhihua and H. Guixiang, *Appl. Biochem. Biotechnol.*, 2013, **171**, 93–103.
33 S. H. Park, D. H. Kim, D. Kim, D. H. Kim, H. C. Jung, J. G. Pan, T. Ahn, D. Kim and C. H. Yun, *Drug Metab. Dispos.*, 2010, **38**, 732–739.

34 H. Zhou, B. Wang, F. Wang, X. Yu, L. Ma, A. Li and M. T. Reetz, *Angew. Chem., Int. Ed.*, 2019, **58**, 764–768.
35 A. B. Carmichael and L. L. Wong, *Eur. J. Biochem.*, 2001, **268**, 3117–3125.
36 M. Landwehr, L. Hochrein, C. R. Otey, A. Kasrayan, J.-E. Bäckvall and F. H. Arnold, *J. Am. Chem. Soc.*, 2006, **128**, 6058–6059.
37 C. J. Whitehouse, S. G. Bell, H. G. Tufton, R. J. Kenny, L. C. Ogilvie and L. L. Wong, *Chem. Commun.*, 2008, 966–968, DOI: **10.1039/b718124h**.
38 W. T. Sulistyaningdyah, J. Ogawa, Q. S. Li, R. Shinkyo, T. Sakaki, K. Inouye, R. D. Schmid and S. Shimizu, *Biotechnol. Lett.*, 2004, **26**, 1857–1860.
39 M. Budde, M. Morr, R. D. Schmid and V. B. Urlacher, *Chembiochem*, 2006, **7**, 789–794.
40 D. Valikhani, J. N. Besna, O. Rousseau, C. Lemay-St-Denis, G. Lamoureux and J. Pelletier, *ChemRxiv*, 2024, DOI: **10.26434/chemrxiv-2024-5tjrz**.
41 J. A. O'Hanlon, X. Ren, M. Morris, L. L. Wong and J. Robertson, *Org. Biomol. Chem.*, 2017, **15**, 8780–8787.
42 C. J. Whitehouse, W. Yang, J. A. Yorke, B. C. Rowlatt, A. J. Strong, C. F. Blanford, S. G. Bell, M. Bartlam, L. L. Wong and Z. Rao, *ChemBioChem*, 2010, **11**, 2549–2556.
43 A. Dennig, N. Lulsdorf, H. Liu and U. Schwaneberg, *Angew. Chem., Int. Ed. Engl.*, 2013, **52**, 8459–8462.
44 S. D. Munday, O. Shoji, Y. Watanabe, L. L. Wong and S. G. Bell, *Chem. Commun.*, 2016, **52**, 1036–1039.
45 S. D. Munday, S. Dezvarei, I. C. K. Lau and S. G. Bell, *ChemCatChem*, 2017, **9**, 2512–2522.
46 S. D. Munday, S. Dezvarei and S. G. Bell, *ChemCatChem*, 2016, **8**, 2789–2796.
47 M. Karasawa, K. Yonemura, J. K. Stanfield, K. Suzuki and O. Shoji, *Angew. Chem., Int. Ed.*, 2022, **61**, e202111612.
48 S. Dezvarei, O. Shoji, Y. Watanabe and S. G. Bell, *Catal. Commun.*, 2019, **124**, 97–102.
49 K. Suzuki, J. K. Stanfield, O. Shoji, S. Yanagisawa, H. Sugimoto, Y. Shiro and Y. Watanabe, *Catal. Sci. Technol.*, 2017, **7**, 3332–3338.
50 C. J. C. Whitehouse, W. Yang, J. A. Yorke, H. G. Tufton, L. C. I. Ogilvie, S. G. Bell, W. Zhou, M. Bartlam, Z. Raob and L.-L. Wong, *Dalton Trans.*, 2011, **40**, 10383–10396.
51 A. Dennig, E. Busto, W. Kroutil and K. Faber, *ACS Catal.*, 2015, **5**, 7503–7506.
52 C. J. Whitehouse, S. G. Bell and L. L. Wong, *Chem.–Eur. J.*, 2008, **14**, 10905–10908.
53 C. J. Whitehouse, S. G. Bell and L. L. Wong, *Chem. Soc. Rev.*, 2012, **41**, 1218–1260.
54 D. J. Fansher, J. N. Besna, A. Fendri and J. N. Pelletier, *ACS Catal.*, 2024, **14**, 5560–5592.
55 N. A. Nguyen, J. Jang, T. K. Le, T. H. H. Nguyen, S. M. Woo, S. K. Yoo, Y. J. Lee, K. D. Park, S. J. Yeom, G. J. Kim, H. S. Kang and C. H. Yun, *J. Agric. Food Chem.*, 2020, **68**, 6683–6691.
56 J. Brummund, M. Muller, T. Schmitges, I. Kaluzna, D. Mink, L. Hilterhaus and A. Liese, *J. Biotechnol.*, 2016, **233**, 143–150.
57 G. Grogan, *JACS Au*, 2021, **1**, 1312–1329.
58 P. C. Cirino and F. H. Arnold, *Angew. Chem., Int. Ed.*, 2003, **42**, 3299–3301.

59 W. Finnigan, M. Lubberink, L. J. Hepworth, J. Citoler, A. P. Mattey, G. J. Ford, J. Sangster, S. C. Cosgrove, B. Z. da Costa, R. S. Heath, T. W. Thorpe, Y. Yu, S. L. Flitsch and N. J. Turner, *ACS Catal.*, 2023, **13**, 11771–11780.
60 T. de Rond, J. Gao, A. Zargar, M. de Raad, J. Cunha, T. R. Northen and J. D. Keasling, *Angew. Chem., Int. Ed.*, 2019, **58**, 10114–10119.
61 A. Fendri, D. Valikhani and J. N. Pelletier, *React. Chem. Eng.*, 2024, **9**, 803–815.