Automated Flow Synthesis of Artificial Heme Enzymes for Enantioselective Biocatalysis

Authors: Giulio Fittolani,[1]† Dennis A. Kutateladze,[1]† Andrei Loas,[1] Stephen L. Buchwald,[1]* Bradley L. Pentelute[1,2,3,4]*

Affiliations:
[1]Department of Chemistry, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
[2]The Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, 500 Main Street, Cambridge, MA 02142, USA
[3]Center for Environmental Health Sciences, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
[4]Broad Institute of MIT and Harvard, 415 Main Street, Cambridge, MA 02142, USA

*Email: sbuchwal@mit.edu, blp@mit.edu

†These authors contributed equally to this work.

Abstract:
The remarkable efficiency with which enzymes catalyze small molecule reactions has driven their widespread application in organic chemistry. Here, we employ automated fast-flow solid-phase synthesis to access full-length enzymes without restrictions on the number and structure of non-canonical amino acids incorporated. We demonstrate the total syntheses of Fe-dependent *Bacillus subtilis* myoglobin (BsMb) and sperm whale myoglobin (SwMb), which displayed excellent enantioselectivity and yield in carbene transfer reactions. Absolute control over enantioselectivity in styrene cyclopropanation was achieved using L- and D-BsMb mutants which delivered each enantiomer of cyclopropane product in identical and opposite enantiomeric enrichment. BsMb mutants outfitted with non-canonical amino acids were used to facilitate detailed structure-activity relationship studies, revealing a previously unrecognized hydrogen-bonding interaction as the primary driver of enantioselectivity in styrene cyclopropanation.

Main text:

Introduction

Enzymatic catalysis has exerted a tremendous impact on the field of synthetic organic chemistry. Current research has focused on identifying and optimizing enzymes that catalyze small-molecule organic reactions,(*1*) pointing to a wide range of opportunities for improving chemistry catalyzed by both natural and engineered proteins. Towards this goal, directed evolution approaches have excelled, leading to a generalizable platform for the development of enzymes for a variety of enantioselective transformations.(*2*, *3*) Indeed, evolved enzymes (*3*–*7*) have been deployed on industrial scales, including in the syntheses of approved drugs sitagliptin,(*8*) molnupiravir,(*9*) and islatravir.(*10*) While efficacious, recombinant expression techniques are currently best suited to accessing protein sequences based on naturally occurring amino acids. The presence of non-canonical amino acids (ncAAs) in the primary sequence of enzymes has in several cases been demonstrated to exert beneficial effects on reaction performance.(*11*, *12*) Genetic code reprogramming(*13*–*16*) techniques have expanded access to proteins containing ncAAs, however, current technologies remain somewhat laborious and are only able to simultaneously incorporate up to two distinct ncAAs with specific steric constraints.(*11*, *13*, *17*–*20*) These limitations highlight the importance of developing general strategies to install multiple ncAAs into the primary amino acid sequence of proteins.

As a complement to recombinant expression techniques, we considered the total chemical synthesis of enzymes which would allow for straightforward incorporation of ncAAs without restraints on their structure and number. Rapid synthetic access to protein catalysts outfitted with ncAAs would allow for exploration of under-investigated non-canonical sequence space, providing enzymes with potential non-biological reactivity or enhanced efficiency.(*13*, *14*) Furthermore, installation of ncAAs by chemical synthesis could be used to enhance mechanistic understanding of biocatalytic reactions with precise control over steric and electronic properties of key amino acid residues.

Recent efforts in our laboratories have resulted in the development of automated fast-flow peptide synthesis (AFPS) technology to produce full-length single domain proteins, and is well suited to those with primary sequences exceeding the length limit of conventional solid phase peptide synthesis (SPPS).(*21*) Key to the efficacy of AFPS is the execution of coupling reactions in a flow format at elevated temperature, which results in fast (~2.5 min) and highly efficient (typically >99.7% yield) coupling/deprotection cycles. Using AFPS, full-length proteins (up to 213 AAs)(*22*) and abiotic mirror-image D-proteins(*23*) can be synthesized on the timescale of hours through successive amino acid couplings on a solid support.

In this work, we describe the first application of AFPS to the total chemical synthesis of full-length artificial heme enzymes and their use in catalytic carbene transfer reactions (**Fig. 1A**). Our findings demonstrate that AFPS provides a new tool for accessing artificial enzymes, paving the way for exploring non-canonical sequence space in enzyme engineering.

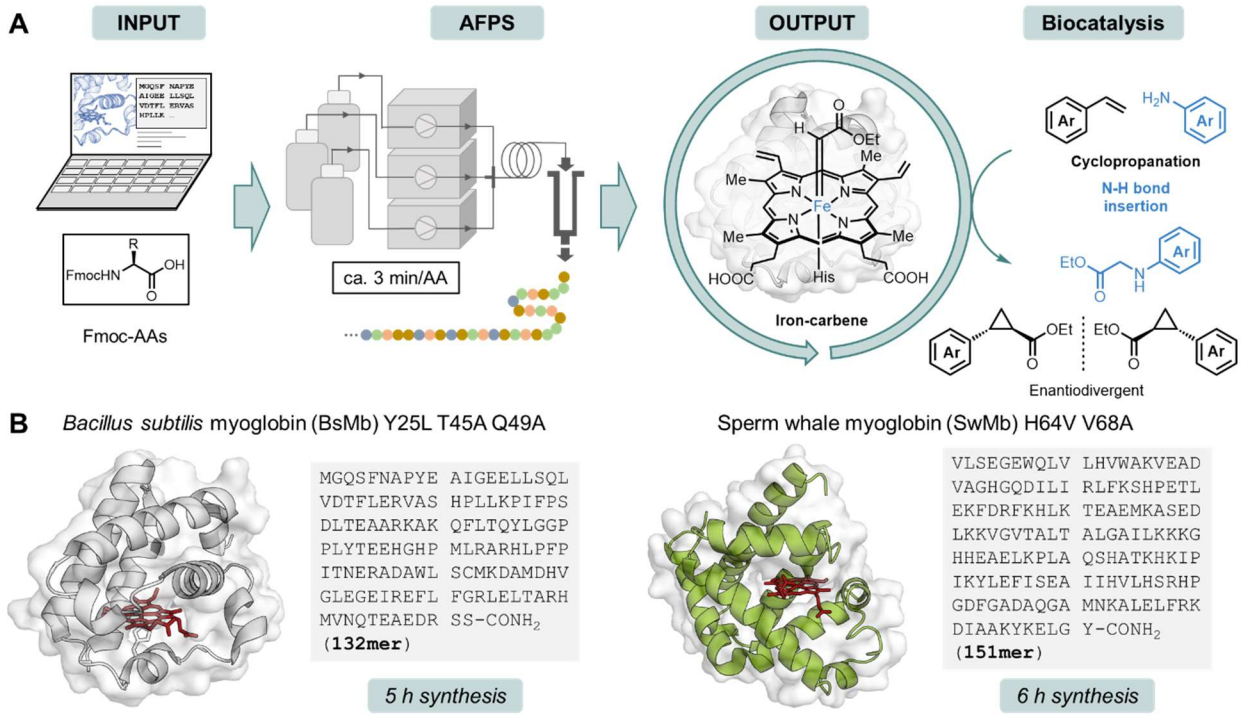


Fig. 1 Streamlined approach to obtain artificial heme enzymes using AFPS. A) The AFPS technology uses a sequence input to provide full-length artificial proteins that are competent in biocatalytic reactions. B) Amino acid sequences of BsMb (PDB 1UX8) and SwMb (PDB 1A6K). All proteins were synthesized as C-terminal primary amides.

Heme enzymes in biocatalysis

In nature, heme-dependent enzymes are involved in an extraordinary number of catalytic oxidative transformations and have proven to be highly successful platforms for synthetic small-molecule biocatalysis.(*24*) Significant efforts have been directed towards engineering heme enzymes for biocatalysis, with a particular focus on the chemistry performed by iron-carbene intermediates.(*25*) In 2013, Arnold demonstrated that cytochrome P450 enzymes from *Bacillus megaterium* could be evolved to catalyze highly enantioselective cyclopropanation reactions on styrene substrates with ethyl diazoacetate.(*26*) Since then, a variety of other heme-dependent enzymes have been shown to be competent for enantioselective styrene cyclopropanation,(*27*) including those harboring abiological noble metal cofactors.(*28*) Beyond cyclopropanation, enzyme-supported iron carbene intermediates derived from diazo compounds have also been leveraged to perform insertion reactions into C–H, C–N, and Si–H bonds.(*29*)

Total chemical syntheses of heme enzymes are rare and can be challenging, requiring multiple native chemical ligation steps or the use of ligation auxiliaries.(*30*, *31*) We hypothesized that the AFPS technology could provide rapid access to artificial heme enzymes in a single-shot format that are active in biocatalytic reactions. To test this, we targeted the syntheses of two heme-dependent enzyme mutants, *Bacillus subtilis* myoglobin (BsMb)(*32*) and sperm whale myoglobin (SwMb),(*33*) which have previously been evolved to catalyze enantioselective olefin cyclopropanation and N–H insertion reactions, respectively (**Fig. 1B**). We began with the synthesis of the triple mutant BsMb Y25L T45A Q49A, which was reported by Arnold to catalyze the cyclopropanation of 3,4-difluorostyrene using ethyl diazoacetate in high yield (79%) and enantiomeric ratio (e.r.) (99:1 e.r.$_{trans}$).(*32*)

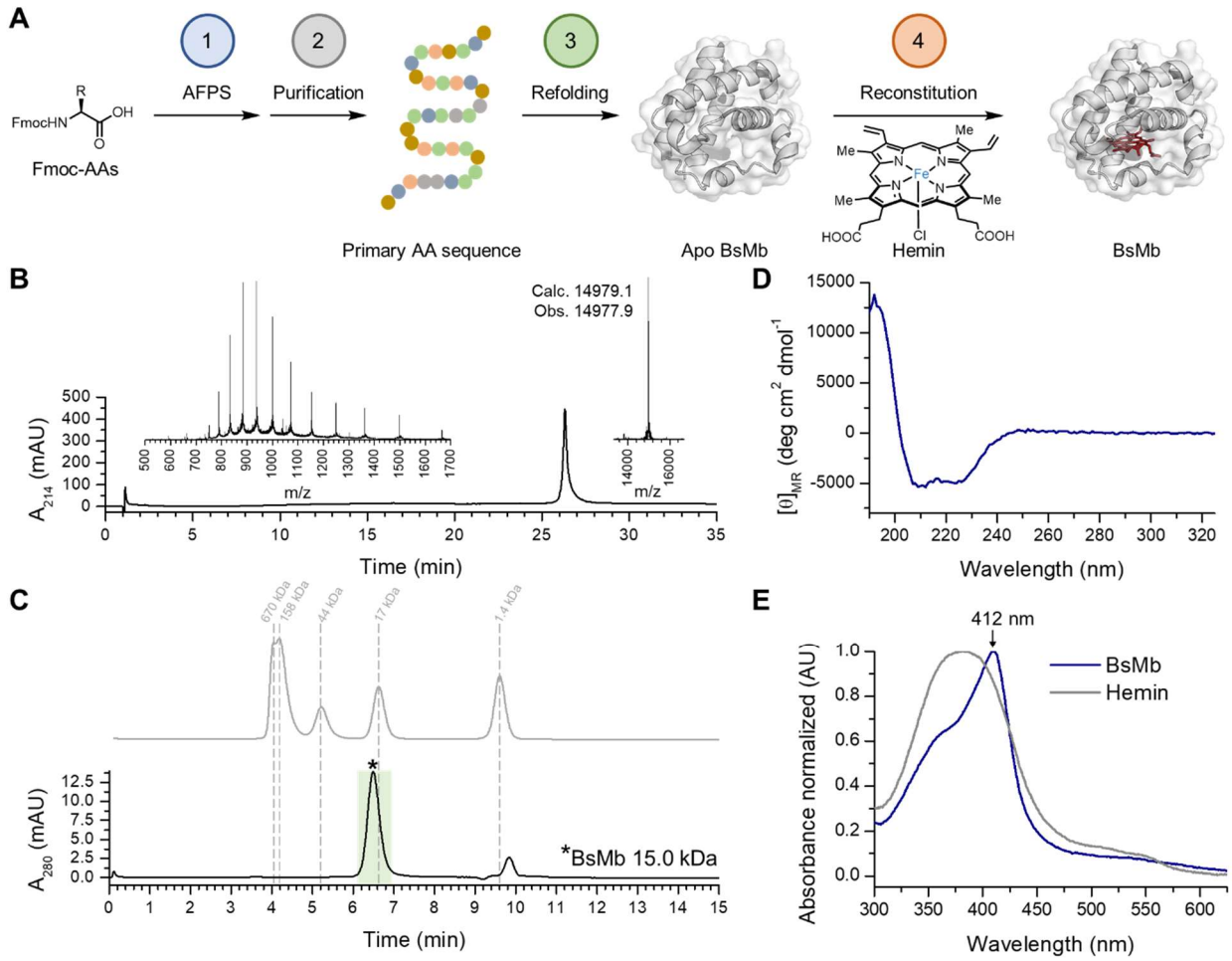


Fig. 2 Synthetic heme enzymes can be refolded and reconstituted with a heme cofactor. A) Summary of the production workflow to obtain catalytically active heme-based enzymes. B) Analytical HPLC chromatogram (*bottom*), mass-to-charge (*left inset*), and deconvoluted (*right inset*) mass spectra of purified BsMb (detailed methods can be found in Section 2 of the Supplementary Materials). C) Analytical size exclusion chromatography (SEC) trace of refolded synthetic BsMb (highlighted in green) compared to a molecular weight standard (*shown in gray*) confirms the monomeric form of BsMb. D) Circular dichroism (CD) spectrum of *apo*-BsMb obtained in 10 mM TRIS, pH 8.0 at 25 °C supports alpha-helical character. E) UV-Vis spectrum of synthetic BsMb reconstituted with heme cofactor shows a red-shift relative to hemin only supporting cofactor incorporation.

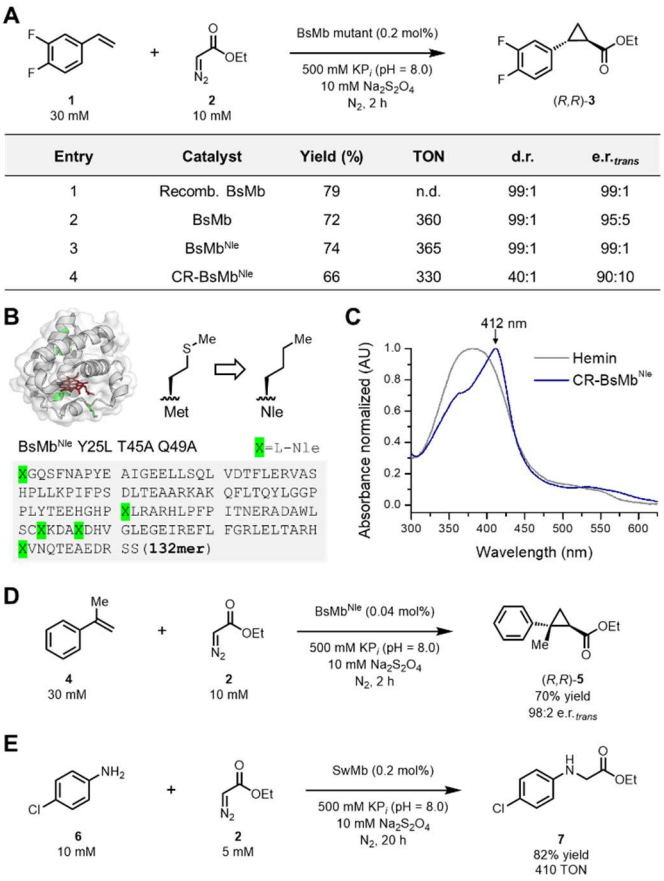
Synthetic heme enzymes produced with AFPS

Access to catalytically active synthetic heme enzymes was accomplished through a four-step protocol consisting of: (1) synthesis of the primary amino acid sequence by AFPS, (2) HPLC purification, (3) refolding by dialysis, and (4) reconstitution of the folded protein with the heme cofactor (**Fig. 2A**). Enzymes were synthesized by AFPS according to our previously reported automated flow methods(*21*) utilizing a solid support (24 μmol typical scale), *in situ* activated amino acids, and Fmoc deprotection. The syntheses of BsMb and SwMb proceeded in 5 h (264 steps) and 6 h (302 steps), respectively, with an average time of 2.5 min for the incorporation of each amino acid residue. Following cleavage and deprotection of the amino acid chain from the peptidyl resin and purification of the resulting material by preparative reverse phase HPLC, the unfolded polypeptides were isolated in up to multi-milligram quantities. The identity and homogeneity of the amino acid sequences of *apo*-BsMb and *apo*-SwMb were characterized by

HPLC and liquid-chromatography mass-spectrometry (LC-MS) methods (**Fig. 2B**) (for SwMb, see Section 3.4 in the Supplementary Materials).

A two-step protocol consisting of refolding and subsequent reconstitution with heme was applied to access functional synthetic enzymes from the purified primary amino acid sequence. Ultimately, we found that dialysis-based methods for refolding provided optimal results in accessing the *apo* forms (see Section 4.1.2 and 4.1.3 of the Supplementary Materials). The synthetic polypeptides were dissolved in a denaturing buffer (typically 6 M guanidinium·HCl, 50 mM TRIS, pH 8.0) and dialyzed for 16 h against an aqueous buffer at 4 °C (20 mM TRIS·HCl, 200 mM NaCl, pH 7.5). Following these procedures, soluble monomeric forms of folded *apo*-BsMb and *apo*-SwMb were obtained as verified with size exclusion chromatography (SEC) (**Fig. 2C** and **S71**). Circular dichroism (CD) spectra of the *apo* forms of BsMb and SwMb obtained following dialysis protocols displayed minima at approximately 210 nm and 220 nm, consistent with largely alpha helical secondary structures, matching with previously reported data for recombinant BsMb(*34*) and SwMb(*28*) (**Fig. 2D** and **S76**).

With access to *apo* enzymes, we examined methods for reconstitution with the heme cofactor, exploiting its tight binding with the globin scaffold.(*35*) Following incubation of Fe-protoporphyrin IX possessing an axial chloride with *apo*-BsMb, successful reconstitution was verified with analytical SEC and in-line UV-Vis spectroscopy (**Fig. S77**). Importantly, the UV-Vis spectrum of reconstituted BsMb displays a significant red-shift of the Soret band (λ_{max} = 412 nm) relative to the heme cofactor alone, consistent with productive binding (**Fig. 2E**).(*36*) SwMb was prepared through a similar protocol using the corresponding AFPS-generated polypeptide, with the resulting synthetic enzyme displaying a comparable red-shift in its UV-Vis spectrum (**Fig. S78**). Additionally, a negative control experiment involving incubation of bovine serum albumin (BSA) with hemin chloride exhibited negligible changes in the UV-Vis profile, supporting highly specific binding between the heme proteins and their cofactor (**Fig. S78**).(*37*)



Entry	Catalyst	Yield (%)	TON	d.r.	e.r.$_{trans}$
1	Recomb. BsMb	79	n.d.	99:1	99:1
2	BsMb	72	360	99:1	95:5
3	BsMb^{Nle}	74	365	99:1	99:1
4	CR-BsMb^{Nle}	66	330	40:1	90:10

Fig. 3 Artificial heme enzymes catalyze stereoselective olefin cyclopropanation and N–H bond insertion. A) Enantioselective cyclopropanation catalyzed by BsMb. Reaction conditions: 20 μM Mb mutant, 30 mM **1**, 10 mM **2**, 10 mM $Na_2S_2O_4$, 10% MeOH, 2 h, r.t., 250 μL scale, reaction performed under N_2 atmosphere. (TON = turnover number, n.d. = not determined, d.r. = diasteromeric ratio, e.r. = enantiomeric ratio) . B) BsMb^{Nle} with mutations highlighted in green (PDB 1UX8). BsMb^{Nle} was obtained as C-terminal amide. C) UV-Vis spectrum of reconstituted CR-BsMb^{Nle} relative to the spectrum of the hemin cofactor. D) BsMb^{Nle} catalyzed cyclopropanation employing α-methylstyrene. E) SwMb-catalyzed N–H bond insertion reaction. Reaction conditions: 30 μM SwMb variant, 10 mM **4**, 5 mM **5**, 10 mM $Na_2S_2O_4$, 5% MeOH, 12 h, r.t., 250 μL scale, reaction performed under N_2 atmosphere. All yields, d.r. and e.r. were evaluated based on GC analysis using dodecane as an internal standard, relative to **2**. All reactions were performed in triplicate.

Carbene transfer reactions catalyzed by synthetic heme enzymes

With access to reconstituted synthetic enzymes, we assessed their performance in catalytic carbene transfer reactions. We began by subjecting BsMb to similar conditions reported by Arnold for the reaction of 3,4-difluorostyrene **1** and ethyldiazoacetate **2**. Employing a low catalyst loading of BsMb (0.2 mol%) in aqueous potassium phosphate buffer, cyclopropyl product (*R*,*R*)-**3** was formed in good yield and with excellent enantioselectivity (72% and 95:5 e.r.$_{trans}$) in 2 h at room temperature. (**Fig. 3A**, entry 1). These values compare well with results reported by Arnold obtained with whole-cell lysates harboring mutant BsMb (79% yield and 99:1 e.r.$_{trans}$).(*32*)

We hypothesized that potential undesired methionine *S*-oxidation or modification by carbene transfer(*38*) might adversely impact the performance of the cyclopropanation reaction. Hence, we synthesized a mutant in which all methionine residues were substituted with the ncAA L-norleucine (Nle), a close structural analog of methionine that features a methylene group in place of the sulfur atom (BsMb^{Nle}) (**Fig. 3B**). Refolding and reconstitution were performed using

identical protocols (**Fig. S70**) and the CD spectrum of BsMb^{Nle} confirmed the overall alpha helical character of the protein (**Fig. S76**). UV-Vis spectroscopy was used to demonstrate successful incorporation of heme (**Fig. S78**). When assayed under standard reaction conditions, BsMb^{Nle} displayed improved enantioselectivity compared to BsMb delivering (*R,R*)-**3** in 99:1 e.r.$_{trans}$ with comparable yield (74%) (**Fig. 3A**, entry 3). Taken together, these data support that global incorporation of Nle does not impair catalytic activity and likely improves enzyme stability.(*39*) These results demonstrate a clear advantage to the incorporation of ncAAs and make a compelling case for enzyme synthesis by AFPS.

Direct folding bypassed HPLC purification

With the aim of reducing the time and resources needed to access active enzymes, we explored the possibility of bypassing HPLC purification of the primary amino acid sequence following synthesis by AFPS. This simplification of the production workflow would accelerate the experimental evaluation of synthetic enzymes, akin to using whole cell lysates in the screening of enzymes produced by directed evolution.(*40*) Following AFPS synthesis of the 132 AA sequence of BsMb^{Nle} and cleavage of the polypeptide from resin, we directly subjected the resulting material to dialysis refolding conditions previously optimized for synthetic BsMb^{Nle}. After reconstitution with the heme cofactor, crude BsMb^{Nle} (termed CR-BsMb^{Nle}) was fully characterized and displayed the presence of a Soret band shift in the UV-Vis spectrum indicated successful incorporation of heme (**Fig. 3C**). When deployed in the cyclopropanation reaction between **1** and **2**, CR-BsMb^{Nle} demonstrated good performance, providing (*R,R*)-**3** in 66% yield and slightly reduced 90:10 e.r.$_{trans}$ (**Fig. 3A**, entry 4). Despite the relatively low purity of CR-BsMb^{Nle} (estimated 19% based on absorbance, **Fig. S74** and **S75**), good enantioselectivity was still achieved, demonstrating the utility of this strategy in bypassing the significant time and resource bottleneck associated with HPLC purification. Hence, synthetic enzymes can be used to perform efficient stereoselective catalysis without purification, which we anticipate could be used for the expedited discovery of biocatalysts.

Expanding the biocatalysis scope of synthetic heme enzymes

We sought to demonstrate the ability of BsMb^{Nle} to catalyze cyclopropanation reactions on unreported substrates for this enzyme. We found that BsMb^{Nle} promoted highly enantioselective conversion of α-methyl styrene **4** to cyclopropanated product **5** (70% yield, 98:2 e.r.$_{trans}$) possessing a quaternary center at low loading (0.04 mol%) (**Fig. 3D**). Notably, enzymatic cyclopropanation reactions of 1,1-disubstituted styrenes are rare, (*33*, *41–43*) highlighting AFPS as an excellent discovery tool for biocatalysis.

To further demonstrate the practical utility of synthetic enzymes, we evaluated the performance of BsMb^{Nle} on a preparative scale (0.1 mmol with respect to **2**) in the cyclopropanation reaction between **2** and **4**. The reaction was found to provide (*R,R*)-**5** in high isolated yield (84%) and excellent enantioselectivity (96:4 e.r.$_{trans}$) using a reduced catalyst loading (0.02 mol%).

In addition to enantioselective cyclopropanation reactions, we investigated the chemistry catalyzed by another heme-dependent enzyme, sperm whale myoglobin (SwMb). Synthetic SwMb was found to be active towards N–H insertion reactions using ethyl diazoacetate, as reported by Fasan and coworkers.(*44*) Incubating 4-chloroaniline **6** and ethyldiazoacetate **2** with synthetic SwMb (H64V V68A) at 0.2 mol% loading led to the formation of product **7** in 82% yield (**Fig. 3E**). These results compare well with the reported performance of SwMb in this reaction,

and further demonstrate that AFPS can be used generally to produce active heme-based enzymes for biocatalysis.(*44*)

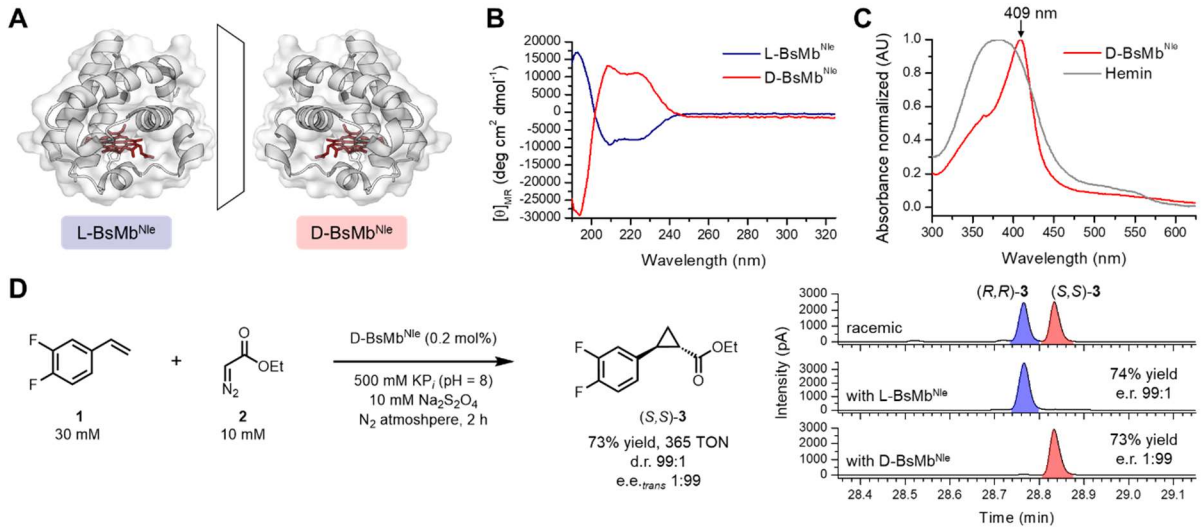


Fig. 4 Artificial D-BsMb^{Nle} is active for enantiodivergent cyclopropanation. A) A schematic comparison of the enantiomeric forms of BsMb^{Nle}. Both proteins were obtained as C-term primary amides. B) CD spectrum of *apo*-D-BsMb^{Nle} and *apo*-L-BsMb^{Nle} obtained in 10 mM TRIS, pH 8.0 at 25 °C. C) UV-Vis spectrum of reconstituted D-BsMb^{Nle} confirms successful incorporation of the heme cofactor. D) Mirror image enantioselective cyclopropanation catalyzed by D-BsMb^{Nle}. Reaction conditions: 20 μM D-BsMb^{Nle}, 30 mM **1**, 10 mM **2**, 10 mM $Na_2S_2O_4$, 5% MeOH, 2 h, r.t., 250 μL scale, reaction performed under N_2 atmosphere (*left*). Stacked chiral GC-MS traces obtained from crude reaction mixtures with $Rh_2(AcO)_4$ yielding a racemic mixture, L-BsMb^{Nle}, and D-BsMb^{Nle} (*right*). All yields, d.r., and e.r. were evaluated based on GC analysis using dodecane as an internal standard, relative to **2**. All reactions were performed in triplicate.

Enantiodivergent catalysis with synthetic D-heme enzymes

While directed evolution campaigns have excelled at delivering enzymes for the production of a specific enantiomer of a product, evolving catalysts from the same starting sequence that generate the opposite enantiomer of product is often challenging and not generalizable.(*45*, *46*) In practice, such an undertaking essentially involves performing directed evolution workflows twice, and though the sense of asymmetric induction can in some cases be inverted, enantioselectivity values are generally not equal and opposite.(*47*) Computational tools have made important progress in predicting enantiocomplementary mutants, however, these techniques have not yet been demonstrated to be generalizable.(*48*, *49*) In principle, synthetic mirror-image D-enzymes(*50*, *51*) could predictably deliver the opposite enantiomer of product with unaltered efficiency and selectivity, bypassing multiple rounds of evolution.

Hence, we aimed to demonstrate that mirror image D-BsMb^{Nle} could be accessed using AFPS and applied to the cyclopropanation reaction of **1** to generate the (*S,S*) enantiomer of cyclopropane **3** (**Fig. 4A**). Notably, recombinant methods to express D-proteins have not yet been developed, necessitating total chemical synthesis as the only current means to access mirror-image proteins.(*23*, *50*, *52*) Using AFPS technology, D-BsMb^{Nle} was synthesized in ca. 5 h starting from commercially available D-amino acids. The homogeneity and purity of the primary amino acid sequence of synthetic D-BsMb^{Nle} were confirmed by HPLC and LC-MS methods (see Section 3.3 in the Supplementary Materials). Refolding by dialysis and reconstitution with the heme cofactor proceeded smoothly to provide D-BsMb^{Nle} which was characterized by CD to display opposite enantiomeric features compared to its L-analog (**Fig. 4B**), while effectively binding heme (**Fig. 4C**). In the model cyclopropanation reaction between styrene **1** and ethyl diazoacetate **2**, D-

BsMb^{Nle} was observed to provide the opposite cyclopropyl product (*S*,*S*)-**3** in identical and opposite enantioselectivity (1:99 e.r.$_{trans}$) without erosion of yield (73%; **Fig. 4D**). The synthetic D-BsMb^{Nle} variant is readily available as an enantiocomplementary mutant, requiring no additional engineering efforts for its generation. To the best of our knowledge, this represents the first example of the use of a full-length D-enzyme in small-molecule biocatalysis.

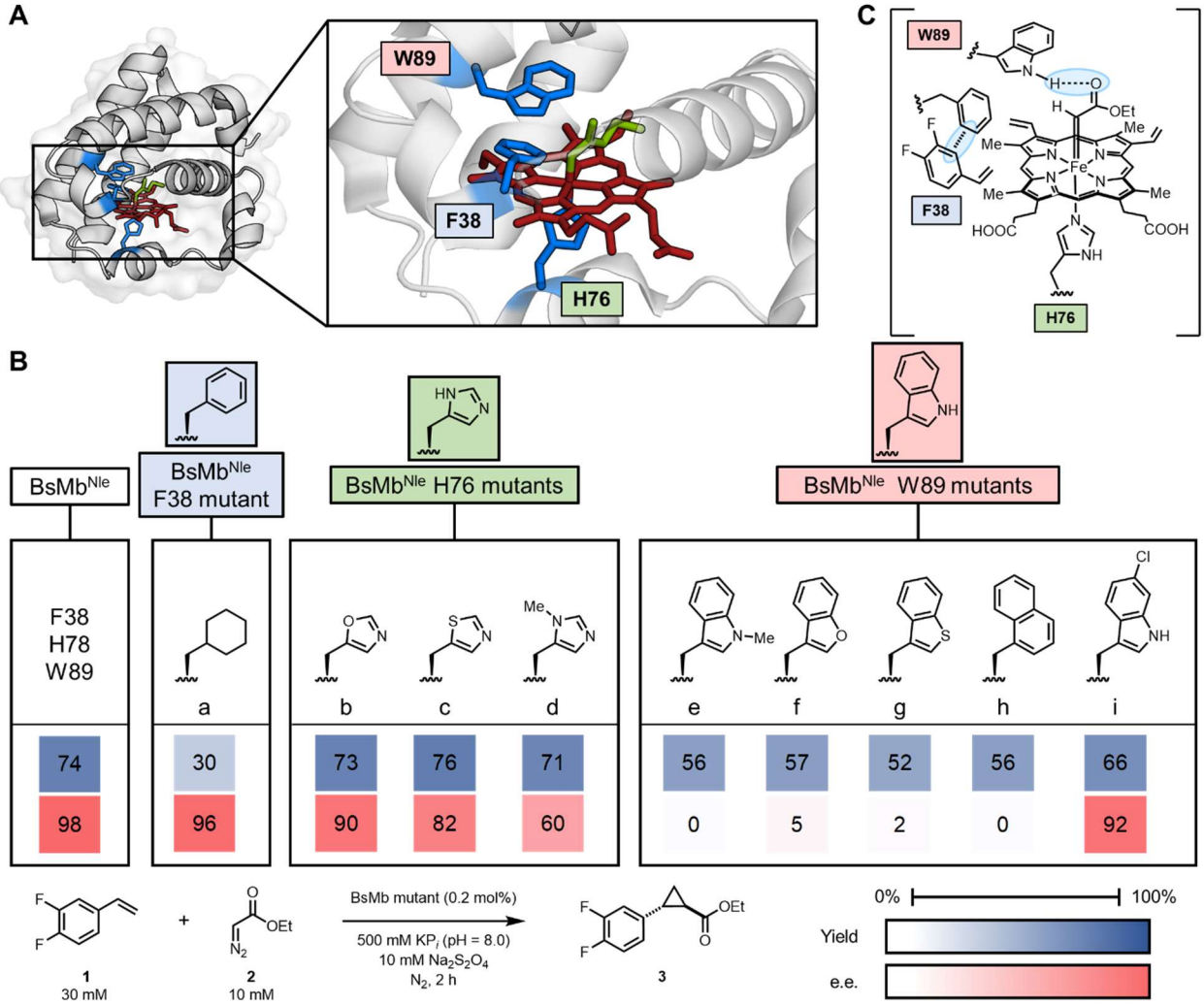


Fig. 5 Non-canonical amino acids inform the origins of enantioselectivity in BsMb-catalyzed cyclopropanation. A) 3D model of BsMb^{Nle} with the iron carbene intermediate modelled in the active site (adapted from crystal structure PDB 1UX8) (*left*). Enlarged depiction of the active site region with the side chains selected for engineering with ncAAs (*right*). B) Enantioselective cyclopropanation catalyzed by BsMb F38, H76, and W89 mutants. Reaction conditions for model reaction: 20 μM Mb mutant, 30 mM **1**, 10 mM **2**, 10 mM $Na_2S_2O_4$, 5% MeOH, 2 h, r.t., 250 μL scale, reaction performed under N_2 atmosphere. All yields, d.r. and e.e. were evaluated based on GC analysis using dodecane internal standard, relative to **2**. All reactions were performed in triplicate. (e.e. = enantiomeric excess) C) Proposed roles of F38, H76, and W89 during the myoglobin catalyzed carbene insertion reaction.

ncAA-containing BsMb mutants offer mechanistic insight

The limited structural diversity of canonical AAs restricts the quality and type of mechanistic insight achievable through mutagenesis studies. We considered that ncAAs could enable highly-targeted examination of specific non-covalent interactions by precisely tuning of electronic and/or steric effects. Historically, mechanistic enzymology has benefitted greatly from

studies involving mutants possessing ncAA in place of native residues, however ncAAs are often challenging to install using recombinant expression means.(*53*) Ultimately, we envisioned that AFPS synthesis could produce novel, ncAA-bearing BsMb analogs required for precise structure-activity characterization, an advance that could offer unprecedented insight into enzymatic carbene transfer behavior.

Guided by crystallographic information available for wild type BsMb (PDB 1UX8), we identified amino acids located in proximity to the Fe-heme cofactor that conceivably are involved in enantioinduction (**Fig. 5A**). Using AFPS, we incorporated a panel of ncAAs at F38, H76, and W89 of BsMb^{Nle} to elucidate the origins of enantioselectivity in the model cyclopropanation reaction for the formation of (*R,R*)-**3** (**Fig. 5B**). In all cases, mutant enzymes were characterized by HPLC and LC-MS analysis (see Sections 3.5 to 3.13 in the Supplementary Materials) and were found to exist as soluble monomers (**Fig. S72** and **S73**) that bind heme (**Fig. S79**).

Fasan and Zhang have provided computational evidence to support the presence of stabilizing aromatic interactions between the styrene substrate and a key phenylalanine residue within the active site of a related myoglobin that catalyzes enantioselective styrene cyclopropanation.(*54*) BsMb possesses an analogous phenylalanine residue (F38) in proximity to the heme binding site that could play a similar role. To test this hypothesis, we used AFPS to synthesize a mutant in which F38 was replaced with a ncAA bearing a cyclohexyl substituent—which cannot engage in such a π-interaction—in place of the corresponding phenyl ring (**Fig. 5A**). When used in the model reaction, BsMb^{Nle} F38a led to a significant decrease in the yield of **3** while maintaining high levels of enantioselectivity (98:2 e.r.$_{trans}$). These results are consistent with the proposal that F38 plays an insignificant role in asymmetric induction but is involved in effective catalysis, possibly by preorganization of the styrene nucleophile via attractive aromatic interactions. Additionally, given the known impact of the axial ligand on the chemistry catalyzed by heme-dependent cytochrome enzymes, we investigated mutations at H76.(*27*) Electronic tuning of the axial ligand on Fe was achieved through replacement of the imidazole of H76 with oxazole (BsMb^{Nle} H76b), thiazole (BsMb^{Nle} H76c), and *N*-methylimidazole (BsMb^{Nle} H76d) (**Fig. 5B**). All mutants proved to be catalytically active towards cyclopropanation to deliver **3** with varying levels of enantiomeric purity.

Finally, we probed the role of W89 in asymmetric induction using a series of ncAAs, ultimately demonstrating the importance of the heterocyclic N–H bond. Incorporating an *N*-methylated tryptophan analog in catalyst BsMb^{Nle} W89e enabled precise removal of the indole N–H bond. When used in the model reaction, the W89e mutant delivered **3** in racemic form and only moderately suppressed yield (56%). Similarly, replacement of the indole of W89 with benzofuran (BsMb^{Nle} W89f), benzothiophene (BsMb^{Nle} W89g), and 1-naphthyl (BsMb^{Nle} W89h) analogs lacking the N–H bond also led to the formation of **3** in good yield and essentially racemic form. In light of the observation of an insignificant racemic background reaction involving free heme (see Section 5.5 in the Supplementary Materials), these results strongly implicate the indole N–H bond of W89 in enantiodiscrimination in this reaction.(*12*) Indeed, high enantioselectivity was recovered when using BsMb^{Nle} W89i, which possesses a 6-chloro indole substituent with an unaltered heterocyclic N–H bond. Taken together, these structure-activity data are consistent with an anchoring hydrogen-bonding interaction between the carbonyl of the metal-carbene intermediate and W89 as being the primary driver of enantioselectivity (**Fig. 5C**). Consistent with the presence of such a hydrogen-bonding interaction, decreasing the coordinating ability of the axial ligand on

Fe with mutants BsMb^{Nle} H76b and BsMb^{Nle} H76c resulted in diminished enantioselectivity, presumably by lowering the electron density on the carbonyl oxygen and rendering it a weaker hydrogen-bond acceptor. Overall, these studies further support the utility of AFPS as a highly enabling tool for mechanistic enzymology.

Conclusions

In summary, this work demonstrates that artificial heme enzymes obtained through total chemical synthesis using AFPS technology exhibit catalytic performance on-par with traditionally produced recombinant enzymes. In a streamlined workflow consisting of polypeptide synthesis, purification, refolding, and reconstitution with heme, fully synthetic enzymes were demonstrated to be accessible in up to milligram quantities. Synthetic enzymes displayed excellent performance in model enantioselective cyclopropanation and N–H bond insertion reactions, providing access to highly enantiomerically enriched products. Unpurified enzymes could be used efficaciously as catalysts, highlighting total synthesis by AFPS as a viable discovery tool for the rapid evaluation of enzyme mutants. D-configured BsMb^{Nle} was successfully shown to catalyze the formation of the opposite enantiomer of cyclopropane product and straightforward incorporation of ncAA at programmed sites in BsMb^{Nle} was leveraged to provide valuable mechanistic insight into the origins of asymmetric induction, revealing the role of an active site hydrogen-bond. Heme-dependent myoglobin enzymes are remarkably important biocatalysts which are known to catalyze a wide range of non-biological reactions, and our current efforts are directed towards exploring new chemistry performed by these enzymes. Beyond this, we anticipate that the strategies disclosed herein will prove enabling for the field of biocatalysis in general by providing a means to access fully synthetic enzymes, especially those possessing ncAAs.

Acknowledgments: We thank Dr. Ahmet Yesilçimen, Dr. Nathan Dow, Dr. Kira Podolsky, Gino Occhialini, Dr. Elaine Raguram, and Dr. Alex Callahan for insightful discussions.

Funding: G.F. acknowledges support from a postdoctoral fellowship from the Ludwig Center at MIT's Koch Institute for Integrative Cancer Research. D.A.K. acknowledges support from a postdoctoral fellowship from the National Institutes of Health (F32GM150211-01). S.L.B. thanks the National Institutes of Health (R35 GM122483) for partial support of this work.

Competing interests: The authors declare no competing interests.

Data and materials availability: All data generated during this study are available either in the main text or supplementary materials.

References:

1. C. K. Winkler, J. H. Schrittwieser, W. Kroutil, Power of Biocatalysis for Organic Synthesis. *ACS Cent. Sci.* **7**, 55–71 (2021).
2. C. Zeymer, D. Hilvert, Directed Evolution of Protein Catalysts. *Annu. Rev. Biochem.* **87**, 131–157 (2018).
3. U. T. Bornscheuer, G. W. Huisman, R. J. Kazlauskas, S. Lutz, J. C. Moore, K. Robins,

Engineering the third wave of biocatalysis. *Nature* **485**, 185–194 (2012).

4. S. Wu, R. Snajdrova, J. C. Moore, K. Baldenius, U. T. Bornscheuer, Biocatalysis: Enzymatic Synthesis for Industrial Applications. *Angew. Chemie Int. Ed.* **60**, 88–119 (2021).

5. S. Simić, E. Zukić, L. Schmermund, K. Faber, C. K. Winkler, W. Kroutil, Shortening Synthetic Routes to Small Molecule Active Pharmaceutical Ingredients Employing Biocatalytic Methods. *Chem. Rev.* **122**, 1052–1126 (2022).

6. D. L. Hughes, Biocatalysis in Drug Development—Highlights of the Recent Patent Literature. *Org. Process Res. Dev.* **22**, 1063–1080 (2018).

7. M. T. Reetz, Biocatalysis in Organic Chemistry and Biotechnology: Past, Present, and Future. *J. Am. Chem. Soc.* **135**, 12480–12496 (2013).

8. C. K. Savile, J. M. Janey, E. C. Mundorff, J. C. Moore, S. Tam, W. R. Jarvis, J. C. Colbeck, A. Krebber, F. J. Fleitz, J. Brands, P. N. Devine, G. W. Huisman, G. J. Hughes, Biocatalytic Asymmetric Synthesis of Sitagliptin Manufacture. *Science* **329**, 305–310 (2010).

9. J. A. McIntosh, T. Benkovics, S. M. Silverman, M. A. Huffman, J. Kong, P. E. Maligres, T. Itoh, H. Yang, D. Verma, W. Pan, H.-I. Ho, J. Vroom, A. M. Knight, J. A. Hurtak, A. Klapars, A. Fryszkowska, W. J. Morris, N. A. Strotman, G. S. Murphy, K. M. Maloney, P. S. Fier, Engineered Ribosyl-1-Kinase Enables Concise Synthesis of Molnupiravir, an Antiviral for COVID-19. *ACS Cent. Sci.* **7**, 1980–1985 (2021).

10. M. A. Huffman, A. Fryszkowska, O. Alvizo, M. Borra-Garske, K. R. Campos, K. A. Canada, P. N. Devine, D. Duan, J. H. Forstater, S. T. Grosser, H. M. Halsey, G. J. Hughes, J. Jo, L. A. Joyce, J. N. Kolev, J. Liang, K. M. Maloney, B. F. Mann, N. M. Marshall, M. McLaughlin, J. C. Moore, G. S. Murphy, C. C. Nawrat, J. Nazor, S. Novick, N. R. Patel, A. Rodriguez-Granillo, S. A. Robaire, E. C. Sherer, M. D. Truppo, A. M. Whittaker, D. Verma, L. Xiao, Y. Xu, H. Yang, Design of an in vitro biocatalytic cascade for the manufacture of islatravir. *Science* **366**, 1255–1259 (2019).

11. A. J. Burke, S. L. Lovelock, A. Frese, R. Crawshaw, M. Ortmayer, M. Dunstan, C. Levy, A. P. Green, Design and evolution of an enzyme with a non-canonical organocatalytic mechanism. *Nature* **570**, 219–223 (2019).

12. M. Ortmayer, F. J. Hardy, M. G. Quesne, K. Fisher, C. Levy, D. J. Heyes, C. R. A. Catlow, S. P. de Visser, S. E. J. Rigby, S. Hay, A. P. Green, A Noncanonical Tryptophan Analogue Reveals an Active Site Hydrogen Bond Controlling Ferryl Reactivity in a Heme Peroxidase. *JACS Au* **1**, 913–918 (2021).

13. F. Agostini, J. Völler, B. Koksch, C. G. Acevedo-Rocha, V. Kubyshkin, N. Budisa, Biocatalysis with Unnatural Amino Acids: Enzymology Meets Xenobiology. *Angew. Chemie Int. Ed.* **56**, 9680–9703 (2017).

14. I. Drienovská, G. Roelfes, Expanding the enzyme universe with genetically encoded unnatural amino acids. *Nat. Catal.* **3**, 193–202 (2020).

15. A. R. Nödling, N. Santi, T. L. Williams, Y.-H. Tsai, L. Y. P. Luk, Enabling protein-hosted organocatalytic transformations. *RSC Adv.* **10**, 16147–16161 (2020).

16. J. W. Chin, Expanding and reprogramming the genetic code. *Nature* **550**, 53–60 (2017).

17. I. Drienovská, C. Mayer, C. Dulson, G. Roelfes, A designer enzyme for hydrazone and oxime formation featuring an unnatural catalytic aniline residue. *Nat. Chem.* **10**, 946–952 (2018).

18. Z. Zhou, G. Roelfes, Synergistic catalysis in an artificial enzyme by simultaneous action of two abiological catalytic sites. *Nat. Catal.* **3**, 289–294 (2020).

19. J. S. Trimble, R. Crawshaw, F. J. Hardy, C. W. Levy, M. J. B. Brown, D. E. Fuerst, D. J. Heyes, R. Obexer, A. P. Green, A designed photoenzyme for enantioselective [2+2] cycloadditions. *Nature* **611**, 709–714 (2022).

20. L. Longwitz, R. B. Leveson-Gower, H. J. Rozeboom, A.-M. W. H. Thunnissen, G. Roelfes, Boron catalysis in a designer enzyme. *Nature* **41**, 39–41 (2024).

21. N. Hartrampf, A. Saebi, M. Poskus, Z. P. Gates, A. J. Callahan, A. E. Cowfer, S. Hanna, S. Antilla, C. K. Schissel, A. J. Quartararo, X. Ye, A. J. Mijalis, M. D. Simon, A. Loas, S. Liu, C. Jessen, T. E. Nielsen, B. L. Pentelute, Synthesis of proteins by automated flow chemistry. *Science* **368**, 980–987 (2020).

22. A. Saebi, J. S. Brown, V. M. Marando, N. Hartrampf, N. M. Chumbler, S. Hanna, M. Poskus, A. Loas, L. L. Kiessling, D. T. Hung, B. L. Pentelute, Rapid Single-Shot Synthesis of the 214 Amino Acid-Long N-Terminal Domain of Pyocin S2. *ACS Chem. Biol.* **18**, 518–527 (2023).

23. A. J. Callahan, S. Gandhesiri, T. L. Travaline, R. M. Reja, L. Lozano Salazar, S. Hanna, Y.-C. Lee, K. Li, O. S. Tokareva, J.-M. Swiecicki, A. Loas, G. L. Verdine, J. H. McGee, B. L. Pentelute, Mirror-image ligand discovery enabled by single-shot fast-flow synthesis of D-proteins. *Nat. Commun.* **15**, 1813 (2024).

24. T. L. Poulos, Heme Enzyme Structure and Function. *Chem. Rev.* **114**, 3919–3962 (2014).

25. E. Danelius, N. J. Porter, J. Unge, F. H. Arnold, T. Gonen, MicroED Structure of a Protoglobin Reactive Carbene Intermediate. *J. Am. Chem. Soc.* **145**, 7159–7165 (2023).

26. P. S. Coelho, E. M. Brustad, A. Kannan, F. H. Arnold, Olefin Cyclopropanation via Carbene Transfer Catalyzed by Engineered Cytochrome P450 Enzymes. *Science* **339**, 307–310 (2013).

27. Y. Yang, F. H. Arnold, Navigating the Unnatural Reaction Space: Directed Evolution of Heme Proteins for Selective Carbene and Nitrene Transfer. *Acc. Chem. Res.* **54**, 1209–1225 (2021).

28. H. M. Key, P. Dydio, D. S. Clark, J. F. Hartwig, Abiological catalysis by artificial haem proteins containing noble metals in place of iron. *Nature* **534**, 534–537 (2016).

29. C. K. Prier, F. H. Arnold, Chemomimetic Biocatalysis: Exploiting the Synthetic Potential of Cofactor-Dependent Enzymes To Create New Catalysts. *J. Am. Chem. Soc.* **137**, 13992–14006 (2015).

30. C. Di Bello, C. Vita, L. Gozzini, Total synthesis of horse heart cytochrome c. *Biochem. Biophys. Res. Commun.* **183**, 258–264 (1992).

31. D. W. Low, M. G. Hill, M. R. Carrasco, S. B. H. Kent, P. Botti, Total synthesis of cytochrome b562 by native chemical ligation using a removable auxiliary. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 6554–6559 (2001).

32. K. E. Hernandez, H. Renata, R. D. Lewis, S. B. J. Kan, C. Zhang, J. Forte, D. Rozzell, J. A. McIntosh, F. H. Arnold, Highly Stereoselective Biocatalytic Synthesis of Key Cyclopropane Intermediate to Ticagrelor. *ACS Catal.* **6**, 7810–7813 (2016).

33. M. Bordeaux, V. Tyagi, R. Fasan, Highly Diastereoselective and Enantioselective Olefin Cyclopropanation Using Engineered Myoglobin-Based Catalysts. *Angew. Chemie Int. Ed.* **54**, 1744–1748 (2015).

34. W. Zhang, J. S. Olson, G. N. Phillips, Biophysical and Kinetic Characterization of HemAT, an Aerotaxis Receptor from Bacillus subtilis. *Biophys. J.* **88**, 2801–2814 (2005).

35. M. S. Hargrove, D. Barrick, J. S. Olson, The association rate constant for heme binding to globin is independent of protein structure. *Biochemistry* **35**, 11293–11299 (1996).

36. L. Giangiacomo, A. Ilari, A. Boffi, V. Morea, E. Chiancone, The Truncated Oxygen-avid Hemoglobin from Bacillus subtilis. *J. Biol. Chem.* **280**, 9192–9202 (2005).

37. L. Zhou, M. Luo, R. Tian, X. P. Zeng, Y. Y. Peng, N. Lu, Generation of a Bovine Serum Albumin-Diligand Complex for the Protection of Bioactive Quercetin and Suppression of Heme Toxicity. *Chem. Res. Toxicol.* **34**, 920–928 (2021).

38. X. Xu, C. Li, Z. Tao, Y. Pan, Aqueous hemin catalyzed sulfonium ylide formation and subsequent [2,3]-sigmatropic rearrangements. *Green Chem.* **19**, 1245–1249 (2017).

39. P. C. Cirino, Y. Tang, K. Takahashi, D. A. Tirrell, F. H. Arnold, Global incorporation of norleucine in place of methionine in cytochrome P450 BM-3 heme domain increases peroxygenase activity. *Biotechnol. Bioeng.* **83**, 729–734 (2003).

40. E. L. Bell, W. Finnigan, S. P. France, A. P. Green, M. A. Hayes, L. J. Hepworth, S. L. Lovelock, H. Niikura, S. Osuna, E. Romero, K. S. Ryan, N. J. Turner, S. L. Flitsch, Biocatalysis. *Nat. Rev. Methods Prim.* **1**, 46 (2021).

41. A. Tinoco, V. Steck, V. Tyagi, R. Fasan, Highly Diastereo- and Enantioselective Synthesis of Trifluoromethyl-Substituted Cyclopropanes via Myoglobin-Catalyzed Transfer of Trifluoromethylcarbene. *J. Am. Chem. Soc.* **139**, 5293–5296 (2017).

42. A. L. Chandgude, R. Fasan, Highly Diastereo- and Enantioselective Synthesis of Nitrile-Substituted Cyclopropanes by Myoglobin-Mediated Carbene Transfer Catalysis. *Angew. Chemie Int. Ed.* **57**, 15852–15856 (2018).

43. P. Bajaj, G. Sreenilayam, V. Tyagi, R. Fasan, Gram-Scale Synthesis of Chiral Cyclopropane-Containing Drugs and Drug Precursors with Engineered Myoglobin Catalysts Featuring Complementary Stereoselectivity. *Angew. Chemie* **128**, 16344–16348 (2016).

44. G. Sreenilayam, R. Fasan, Myoglobin-catalyzed intermolecular carbene N-H insertion with arylamine substrates. *Chem. Commun.* **51**, 1532–1534 (2015).

45. A. M. Knight, S. B. J. Kan, R. D. Lewis, O. F. Brandenberg, K. Chen, F. H. Arnold, Diverse Engineered Heme Proteins Enable Stereodivergent Cyclopropanation of Unactivated Alkenes. *ACS Cent. Sci.* **4**, 372–377 (2018).

46. P. F. Mugford, U. G. Wagner, Y. Jiang, K. Faber, R. J. Kazlauskas, Enantiocomplementary enzymes: Classification, molecular basis for their enantiopreference, and prospects for mirror-image biotransformations. *Angew. Chemie - Int. Ed.* **47**, 8782–8793 (2008).

47. W. Fu, N. M. Neris, Y. Fu, Y. Zhao, B. Krohn-Hansen, P. Liu, Y. Yang, Enzyme-controlled stereoselective radical cyclization to arenes enabled by metalloredox biocatalysis. *Nat. Catal.* **6**, 628–636 (2023).

48. Z. Wu, S. B. J. Kan, R. D. Lewis, B. J. Wittmann, F. H. Arnold, Machine learning-assisted directed protein evolution with combinatorial libraries. *Proc. Natl. Acad. Sci.* **116**, 8852–8858 (2019).

49. I. Kalvet, M. Ortmayer, J. Zhao, R. Crawshaw, N. M. Ennist, C. Levy, A. Roy, A. P. Green, D. Baker, Design of Heme Enzymes with a Tunable Substrate Binding Pocket Adjacent to an Open Metal Coordination Site. *J. Am. Chem. Soc.*, doi: 10.1021/jacs.3c02742 (2023).

50. K. Harrison, A. S. Mackay, L. Kambanis, J. W. C. Maxwell, R. J. Payne, Synthesis and applications of mirror-image proteins. *Nat. Rev. Chem.* **7**, 383–404 (2023).

51. R. C. deL. Milton, S. C. F. Milton, S. B. H. Kent, Total Chemical Synthesis of a D-Enzyme: The Enantiomers of HIV-1 Protease Show Reciprocal Chiral Substrate Specificity. *Science* **256**, 1445–1448 (1992).

52. Y. Xu, T. F. Zhu, Mirror-image T7 transcription of chirally inverted ribosomal and functional RNAs. *Science* **378**, 405–412 (2022).

53. N. Morikubo, Y. Fukuda, K. Ohtake, N. Shinya, D. Kiga, K. Sakamoto, M. Asanuma, H. Hirota, S. Yokoyama, T. Hoshino, Cation−π Interaction in the Polyolefin Cyclization Cascade Uncovered by Incorporating Unnatural Amino Acids into the Catalytic Sites of Squalene Cyclase. *J. Am. Chem. Soc.* **128**, 13184–13194 (2006).

54. A. Tinoco, Y. Wei, J.-P. Bacik, D. M. Carminati, E. J. Moore, N. Ando, Y. Zhang, R. Fasan, Origin of High Stereocontrol in Olefin Cyclopropanation Catalyzed by an Engineered Carbene Transferase. *ACS Catal.* **9**, 1514–1524 (2019).