Biocatalytic Strategy for Construction of sp3-Rich Polycyclic Compounds from Directed Evolution and Computational Modeling

Rudi Fasan (rfasan@ur.rochester.edu)
University of Rochester https://orcid.org/0000-0003-4636-9578
David A. Vargas
University of Rochester
Xinkun Ren
University of Rochester
Arkajyoti Sengupta
University of California, Los Angeles
Ledong Zhu
Shandong University
Marc Garcia-Borràs
University of Girona https://orcid.org/0000-0001-9458-1114
Kendall Houk
University of California Los Angeles https://orcid.org/0000-0002-8387-5261

Article

Keywords:

Posted Date: May 13th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1639676/v1

Additional Declarations: There is **NO** Competing Interest.

Version of Record: A version of this preprint was published at Nature Chemistry on February 13th, 2024.
See the published version at https://doi.org/10.1038/s41557-023-01435-3.

Abstract

Catalysis with engineered enzymes has provided more efficient and streamlined routes for the synthesis and manufacturing of drug molecules. Despite landmark achievements, the potential of biocatalysis toward assisting early-stage drug discovery campaigns remains largely untapped. We have developed a novel biocatalytic strategy for the construction of sp3-rich polycyclic compounds via an intramolecular cyclopropanation of benzothiophenes and related heterocycles. Despite the inherent challenge presented by this reaction, two regiocomplementary carbene transferases were evolved to catalyze the highly stereoselective cyclization of benzothienyl substrates bearing diazo ester groups at either the C2 or C3 position of the heterocycle. The detailed mechanisms of these reactions were obtained by a combination of crystallographic and computational (quantum mechanics calculations and molecular dynamics simulations) analyses. Leveraging these insights, the narrow substrate scope of one of the biocatalysts could be expanded to include previously unreactive substrates, highlighting the value of integrating evolutionary and rational strategies for the development of enzymes useful for new-to-nature enzyme transformations. The molecular scaffolds made accessible by the present strategy feature a combination of 3D and stereochemical complexity with 'rule-of-3' properties, which should make them highly valuable for fragment-based drug discovery campaigns.

Introduction

Biocatalysis with engineered enzymes has been covering an increasing important role toward enabling and streamlining the stereoselective synthesis of drug molecules and other high-value compounds.[1-6] In addition, the reaction scope of biocatalysis has been recently expanded to include new-to-nature transformations.[7-10] A highly attractive, but currently underexploited, role for biocatalysis lies in providing access to stereochemically rich, 3D 'fragments' for fragment-based drug discovery (FBDD) campaigns (**Figure 1a,b**).[11] Over the past two decades, FBDD has represented a key strategy for drug discovery, yielding several candidates entering clinical trials, some of which have been approved for commercialization.[12] This drug discovery approach relies on the availability of libraries of small and diverse organic molecules ('fragments') which are screened to identify weak binders of a protein of interest and then linked to generate potent inhibitors of such targets.[12,13] In terms of physico-chemical properties, best suited fragments for FBDD applications are molecules that adhere to the "rule of three" (Ro3), corresponding to a molecular weight below 300 Da, less than 3 rotatable bonds, and an n-octanol/water partition coefficient (cLogP) lower than 3.[14] While several readily available molecules meet these requirements, medicinal chemists and chemical biologists have highlighted the need for more unique, stereochemically rich, and three-dimensional 'fragments' to expand opportunities for drug discovery via FBDD.[15-18] Motivated by this context, we have developed a biocatalytic strategy to access a series of previously inaccessible sp^3-rich sulfur-containing polycyclic scaffolds which exhibit these highly sought-after features and could thus serve as potentially valuable additions to fragment libraries for FBDD applications.

Heme-dependent enzymes and proteins have emerged as promising biocatalysts for mediating carbene transfer reactions such as olefin cyclopropanations that are not known in nature.[19-26] Artificial metalloenzymes useful for this type of reactions have also been reported.[27-34] More recently, the scope of these strategies has been extended to include the first examples of enzyme-catalyzed intramolecular cyclopropanations.[35,36] Herein, we report a novel biocatalytic methodology based on engineered myoglobin variants for the asymmetric intramolecular cyclopropanation of diazoester-functionalized benzothiophenes. This reaction has no chemocatalytic precedent and attempts to realize similar transformations using organometallic catalysts were met with limited success.[37,38] Upon optimization of the biocatalyst via protein engineering, both C2- and C3-functionalized benzothiophenes, along with other heterocycles, could be cyclized with excellent enantioselectivity to furnish stereochemically-rich, three-dimensional scaffolds (Fsp^3: 0.36, 75% of which are stereogenic centers) with physical-chemical properties adhering to the rule of three (Avg. MW: 230 ± 31 Da, Avg. Rot. Bonds: 0.1 ± 0.3, Avg cLogP: 2.5 ± 0.5, Avg Polar Surf. Area: 28.4 ± 3.1; **Figure 1c**). Crystallographic analysis of the engineered enzymes in combination with computational studies based on DFT calculations and MD simulations, provide insights into the mechanism and origin of protein-mediated stereocontrol in these reactions, which were further leveraged to enhance the activity and expand the substrate scope of these biocatalysts.

Results And Discussion

Biocatalyst Evolution for C2-functionalized Benzothiophene. In initial studies, a diverse set of heme-containing enzymes and proteins was screened, including wild-type myoglobin (Mb), its distal histidine variant Mb(H64V), $P450_{BM3}$, CYP119, P411-CHF,[41] different cytochrome *c* proteins, and others, for their ability to promote the intramolecular cyclopropanation of benzo[b]thiophen-2-ylmethyl 2-diazoacetate (**1a**) to give the target tetracyclic scaffold **2a**. However, none of the enzymes produced any detectable amount of the desired product (**Figure 2a** and **Table S1**). Of note, the same reaction also failed or gave minimal yields in the presence of several transition metal catalysts commonly used for carbene transfer reactions such as $Rh_2(OAc)_4$, Fe(TPP)(Cl), and Co(TPP) (**Figure 2a and Table S1**). We then extended our screening to an in-house collection of ~100 diverse Mb variants containing a range of single to quadruple mutations at residues surrounding the heme cofactors. While the large majority of these variant showed no activity, Mb(H64F) displayed basal activity for formation of the desired product **2a** (2% assay yield) with modest enantioselectivity (61% *ee*; **Figure 2c**). Mb(H64F) was thus selected as the parent scaffold for iterative rounds of directed evolution, in which active site residues most proximal to the iron center, i.e. Leu29, Phe43, Phe46, Val68, and Ile107, were randomized via site-saturation mutagenesis, followed by library screening in multi-well plates and as whole cells under anaerobic conditions. The improved hits identified after each round were validated as purified protein in the reaction with **1a**.

Through this process, five beneficial mutations were accumulated resulting in the quintuple mutant Mb(F43I,F46L,H64F,V68G,I107A), called $Mb_{BTIC\text{-}C2}$, with significantly increased activity for generation of **2a** from **1a** (2%→75% assay yield). In addition, $Mb_{BTIC\text{-}C2}$ displays excellent enantioselectivity (>99% *ee*) producing a single stereoisomer whose configuration was determined to be 3a*S*,3b*S*,8a*R* via single-crystal

X-ray diffraction (**Figure 3c**). Interestingly, structure-activity analyses of the $Mb_{BTIC-C2}$ lineage revealed a clear synergistic effect of mutations F43I and F46L on the catalytic activity (TON) of the enzyme (**Figure 2c**), without affecting enantioselectivity. These analyses also indicated a distinct benefit of a His→Phe mutation at the level of the distal histidine (His64) and large-to-small mutations at multiple sites within the heme pocket (Phe43→Ile/Leu; Phe46→Ile/Leu; Ile107→Ala) to favor the target reaction. The latter findings are generally consistent with the high steric demands inherent to mediating an intramolecular cyclopropanation reaction with the confines of myoglobin heme pocket.[35]

Cyclopropanation of C3-functionalized Benzothiophene. Encouraged by the results above, we sought to apply the evolved $Mb_{BTIC-C2}$ biocatalyst to the cyclopropanation of C3-functionalized benzothiophene (**3a**), as we envisioned this reaction would provide access to an alternative tetracyclic scaffold (**Figure 1c**). Surprisingly, no detectable amount of the desired product **4a** was obtained, clearly indicating the need for very different active site configurations for enabling the cyclization of **3a** vs. its regioisomer **1a** (**Figure 2d**). To develop a biocatalyst for this reaction, we re-screened our in-house library of Mb variants against substrate **3a**, the large majority of which showed no activity (**Table S2**). Unlike for **1a**, V68G (and other single mutations at position 68) had no beneficial effect toward transformation of **3a**, further highlighting the divergent catalyst requirement for the two reactions (**Figure 2d**). In contrast, Mb(H64V,V68A) yielded a viable biocatalyst for the intramolecular cyclopropanation of **3a**, producing **4a** in low yield (23% AY, **Figure 2d**) but with excellent enantioselectivity (>99% *ee*) toward formation of the 3a*S*,3b*S*,8b*S* enantiomer as determined by X-ray difraction (**Figure 1c**). Building upon these results, a protein engineering campaign was undertaken that resulted in the identification of Mb(L29F,H64V,V68A,I107L), named $Mb_{BTIC-C3}$, as an improved biocatalyst for this reaction, furnishing 3-fold improved yield of **4a** (23%→60% AY) while retaining excellent enantioselectivity (>99% *ee*) (**Figure 2d**). After optimization of the $Mb_{BTIC-C3}$-catalyzed reaction (i.e., 4°C, slow addition of **3a**, **Table S3**), the cyclopropanation product **4a** could be obtained in quantitative yields and enantiopure form (>99% *ee*) (**Figure 2d**). Under catalyst limited conditions, $Mb_{BTIC-C3}$ was found to support up to 440 TON, producing **4** in 87% yield using only 0.2 mol% catalyst (**Table S2**). Moreover, both $Mb_{BTIC-C2}$ and $Mb_{BTIC-C3}$ could be applied in whole cells (OD_{600}:20 and OD_{600}:60, respectively) for the stereoselective synthesis of **4a** in 62% AY and of **2a** in 65% AY, respectively, demonstrating the compatibility of these enzymatic reactions with whole-cell biotransformations (**Table S1 and S2)**.

Side-by-side comparison of the mutations in $Mb_{BTIC-C3}$ vs. $Mb_{BTIC-C2}$ revealed interesting similarities and differences between the two variants. Similarly to $Mb_{BTIC-C2}$, $Mb_{BTIC-C3}$ incorporated space-creating mutations at positions 68 and 107 (**Figure 2b**), although the nature of the optimal residue differs in each case (V68: Gly vs. Ala; I107: Ala vs. Leu). In contrast to $Mb_{BTIC-C2}$, however, mutations at positions Phe43 and Phe46 were detrimental for the cyclization of **3a**. In addition, for the latter reaction, an aromatic residue (Phe) was beneficial in place of leucine at position 29, whereas mutation of this position provided no benefit for the cyclization of **1a**. These differences, along with the lack of reactivity of $Mb_{BTIC-C3}$ toward **1a**, further evidenced the orthogonal active site requirements for enabling the Mb-catalyzed cyclization of the two regioisomeric substrates.

Substrate Scope of the evolved biocatalysts. Next, we investigated the substrate scope of the $Mb_{BTIC-C2}$ biocatalyst using variously substituted benzothienyl molecules (**Figure 3a**). These experiments showed that the enzyme can tolerate a range of electron-withdrawing and electron-donating groups on position C5 delivering the desired cyclopropyl-lactone products (**2b-f**) in 36-99% yields and with excellent enantioselectivity (>99% *ee*) (**Figure 3a**). Similar results were obtained for a series of 6-substituted substrates indicating a significant tolerance of the biocatalyst to variation at the C6 position to furnish **2g-j** in 25-44% yields and high enantiopurity

(97-99% *ee*, **Figure 3a**). Improved yields for the synthesis of **2g** and **2j,** respectively, could be achieved using whole cells expressing the $Mb_{BTIC-C2}$ variant (51-55% vs. 35-39%). The 7-fluoro-substituted substrate **1k** was also processed by $Mb_{BTIC-C2}$ to give **2k** in good yield and high enantiopurity (>99% *ee*). In contrast, substitution at the C4 site (e.g., -Br) were not tolerated, resulting in unreacted starting material.

Using a similar approach, we investigated the scope of the $Mb_{BTIC-C3}$-catalyzed reaction using a series of C3-functionalized benzothiophenes bearing substitutions in positions C5, C6, and C7 (**Figure 3b**). Compared to $Mb_{BTIC-C2}$, these experiments revealed a significantly different trend for $Mb_{BTIC-C3}$ in terms of the positional effect of substitutions on the enzyme activity. Specifically, substitutions on position C7 were well tolerated by $Mb_{BTIC-C3}$ producing **4g-i** with high efficiency (75%-85% yields; **Figure 3b**). For **3g** and **3i,** nearly quantitative yields (96%) could be achieved by using a higher catalyst loading (1.6 vs. 0.8 mol%). In contrast, only limited substitutions (e.g., -F, $-CH_3$) at position C5 and C6 of the substrate were accepted by the enzyme, with the higher yield of **4f** vs. **4b** (62% vs. 12%) suggesting a larger tolerance toward C6 vs. C5 substitution on the benzothiophene ring. Despite these limitations, all of the $Mb_{BTIC-C3}$-catalyzed reactions were found to proceed with excellent enantiocontrol, furnishing the desired cyclopropanation products in 97-99% *ee* (**Figure 3b**). To further investigate the scope of this reaction, the enzyme was then challenged with substrates **3j** and **3k**, resulting in the successful synthesis of **4j** and **4k** in quantitative yields and 50-99% *ee* (**Figure 3b**). These results demonstrated that the biocatalytic transformation can be readily extended to thiophene-based rings and other heterocycles, respectively. Importantly, the new molecular scaffolds made accessible through these $Mb_{BTIC-C2}$- and $Mb_{BTIC-C3}$-catalyzed reactions combine several highly sought-after features for FBDD,[15,16] namely 3D shape, stereochemical complexity (3 stereogenic centers), and Ro3-compliant physico-chemical properties (**Figure 1c**).

Mechanistic Studies. Hemoprotein-catalyzed cyclopropanation reactions using diazo compounds involve the formation of a reactive iron porphyrin carbene (IPC) intermediate, which can engage the olefin to yield the cyclopropanation product via concerted or stepwise-diradical pathways.[33,42] To gain insight into the intrinsic mechanism of the reactions investigated here, these transformations were studied using density functional theory (DFT) with a truncated model to facilitate computations. The reaction first involves the loss of N_2 and formation of a Fe-carbene IPC intermediate $\mathbf{Im1_{C2}}$ (open-shell singlet, ΔG = 3.6 kcal/mol) or $\mathbf{Im1_{C3}}$ (triplet, ΔG = 0.8 kcal/mol) from **1a** and **3a** respectively (**Figure 4**). The lowest Gibbs activation energy barriers for this step involve an open-shell singlet (OSS) transition state, $\mathbf{TS1_{C2}}$ (26.9 kcal/mol)

and **TS1$_{C3}$** (27.7 kcal/mol) (**Figure 4**). In the second step of the reaction, the reactive Fe-carbene (**Im1$_{C2}$** or **Im1$_{C3}$**) attacks C2, alpha to the sulfur atom, via **TS2$_{C2}$** (quintet, $\Delta G^{‡}$ = 14.7 kcal/mol), or C3 via **TS2$_{C3}$** (open-shell singlet, $\Delta G^{‡}$ = 16.2 kcal/mol), forming a five-membered spirocycle intermediate, **Im2$_{C2}$**, or a six membered ring, **Im2$_{C3}$**, respectively, both with a radical localized and stabilized at the benzylic position (**Figure 4**). These important geometric differences between optimized intermediates **Im2$_{C2}$** and **Im2$_{C3}$** also imply dramatically different geometric requirements for the benzothiophene ring when approaching to the carbene center during the cyclization **TS2$_{C2}$** and **TS2$_{C3}$** transition states (**Figure 4**). This is in line with the experimentally observed divergent catalyst requirements for C2- vs. C3-linked substrates. Both intermediates (**Im2$_{C2}$** and **Im2$_{C3}$**, lowest in energy triplet state) undergo a very fast second C–C bond formation to generate the benzothienyl cyclopropyl products **2a** or **4a** with barriers of less than 2-4 kcal/mol (**TS3$_{C2}$**, triplet $\Delta G^{‡}$= 3.4 kcal/mol; **TS3$_{C3}$**, triplet $\Delta G^{‡}$= 1.4 kcal/mol). These studies imply the occurrence of spin-crossing events along the radical reaction pathway, leading to a multi-state mechanism.[43]

The calculated activation barriers for the transition states, **TS1$_{C3}$**, of C5-, C6-, and C7-functionalized benzothiophenes are similar ($\Delta G^{‡}$ ~ 27 kcal/mol) with the Br substituent. Slightly higher activation barriers ($\Delta G^{‡}$ ~ 30 kcal/mol) for Me and Cl substituents are predicted (**Table S4**). For all substituents at C5-, C6-, and C7-, the calculated activation barriers are within 1.6 kcal/mol, indicating the intrinsic substrate reactivity is independent of the position of substituents.

To probe the formation of radical intermediate species along these reaction pathway, the Mb$_{BTIC-C3}$-catalyzed cyclization of **3a** was carried out in the presence and absence of the radical trapping reagent DMPO.[33,42] These experiments showed a 45% reduction in the yield of the cyclopropanation product in the presence of DMPO compared to the same reaction performed in the absence of it (**Figure S1**). Furthermore, a DMPO-adduct in the former reaction was observed via GC-MS (**Figure S1**). Similar results were observed for the Mb$_{BTIC-C2}$-catalyzed cyclization of **1a**, which showed a 37% reduction in cyclopropanation activity in the presence of DMPO (**Figure S1**). Computations demonstrate a triplet radical character of the carbenes, **Im1$_{C2}$** and **Im1$_{C3}$** (**Figure S2**). The radical intermediates, **Im2**, are too short lived to be trapped by a bimolecular mechanism (**Figure 4**), and hence it is proposed that DMPO reacts with **Im1** intermediates instead (**Figure S2**). This mechanism differs from the concerted mechanism—mediated by a closed-shell singlet IPC intermediate—previously established for the Mb-catalyzed *intermolecular* cyclopropanation of vinylarenes with diazo esters.[42] This difference can be attributed to the electronic properties of the thiophene and aromatic rings that stabilize radical formation at the benzylic position. This behavior, along with the calculated DFT pathways, highlight the multi-state reactivity of the hemoprotein in carbene transfer reactions.[33]

Crystallographic Studies. To gain insights into the structures of the evolved enzymes Mb$_{BTIC-C2}$ and Mb$_{BTIC-C3}$, these proteins were crystallized in their ferric aquo-complex state and their X-ray structure solved to 1.3 Å resolution. Structural alignment of these structures with that of wild type Mb (pdb

1JW8)[44] yielded root-mean square deviation (RMSD) values for the protein backbone of 0.22 Å and 0.21 Å, respectively, indicating these proteins share a very similar fold. Further inspection of the crystal structures revealed however that the two variants exhibit a significantly different active site configuration as shaped by the respective mutations (**Figure 5a,b**). In particular, $Mb_{BTIC\text{-}C2}$ features an enlarged heme pocket (volume: 335 $Å^2$ vs. 125 $Å^2$ for wild-type Mb) as result multiple space-creating mutations, in particular above the inner pyrrole rings A and B of the heme cofactor (**Figure 5a**). The Phe residue replacing the distal His64 is found in two possible orientations of similar occupancy (60:40), acting as a gating residue at the interface between the heme cavity and the solvent (**Figure S3**). In stark contrast, the active site of MbC3 is characterized by a much smaller volume compared to both $Mb_{BTIC\text{-}C2}$ and Mb (237 $Å^2$; **Figure 5b**), as dictated by the presence of Phe29 (from the L29F mutation) and its packing against the side chain of Leu107 (from the I197L mutation). Together with the H64V mutation, this active site configuration completely obstructs and reduces the space available in the 'inner side' of the pocket (i.e., above heme ring B and A, respectively), while it creates a cavity above the solvent exposed rim of the porphyrin cofactor (**Figure 5b**).

Origins of Enantiocontrol in the Enzymatic Reactions. Using the new solved x-ray structures, quantum mechanical DFT studies using the cluster model approach[45,46] were carried out to understand the basis for the high stereoselectivity exhibited by $Mb_{BTIC\text{-}C2}$ and $Mb_{BTIC\text{-}C3}$ in the respective reactions. Starting from $Mb_{BTIC\text{-}C2}$, DFT calculations were performed to reoptimize $\mathbf{TS2_{C2}}$ (**Figure 5c**), i.e. the TS leading to the experimentally generated enantiomer, as well as its enantiomeric transition state ($\mathbf{TS2\text{-}enan_{C2}}$), in the presence of nearby amino acid residues in the protein active site, building theoretical enzyme ("*theozyme*") models previously used by our group and others.[47,48] The final MD snapshot from 500 ns MD simulations were used as starting points for DFT cluster model optimizations (Figure 5c). In the optimized $\mathbf{TS2_{C2}}$/$Mb_{BTIC\text{-}C2}$ complex, the benzothiophene ring is placed within the heme pocket, with the heteroaryl ring oriented away from the Phe64 and Val107 residues and occupying the cavity above heme ring B created by the V68G mutation (**Figure 5c**). In contrast, in the optimized complex with $\mathbf{TS2\text{-}enan_{C2}}$, which would lead to the opposite enantiomer, the benzothiophene ring is situated over ring A in close proximity to residues Phe64 and Ile43, resulting in steric clashes with these residues (**Figure 5c**). These unfavorable interactions result in a significantly higher energy for $\mathbf{TS2\text{-}enan_{C2}}$ vs. $\mathbf{TS2_{C2}}$ ($\Delta\Delta G^{‡}$ = +6.9 kcal/mol), which can explain the high enantioselectivity of the enzyme observed experimentally.

We also investigated the origins for the high enantiocontrol exerted by $Mb_{BTIC\text{-}C3}$ in the cyclization of **3a**. In the DFT optimized $Mb_{BTIC\text{-}C3}$ cluster model structure of $\mathbf{TS2_{C3}}$, the benzothiophene ring is oriented above ring D of the heme and sits below Val64 (as in the final MD snapshot), pointing toward the solvent-exposed side of the cofactor (**Figure 5d**). This configuration is largely dictated by the presence of Phe29 residue, which occupies the inner side of the heme pocket (**Figure 5b**). In the optimized structure of $\mathbf{TS2\text{-}enan_{C3}}$, the benzothiophene ring sterically clashes with Val64 and displaces the Phe29 ring, resulting in a ΔG value of 10.7 kcal/mol higher than that of $\mathbf{TS2_{C3}}$. This large energy difference can thus

rationalize the excellent enantiomeric excess (>99% *ee*) observed for the 3a*S*,3b*S*,8b*S*-configured product generated by the $Mb_{BTIC-C3}$ variant.

Molecular Dynamics Simulations on $Mb_{BTIC-C2}$ system. We performed molecular dynamics (MD) simulations to investigate how the protein scaffold and additional amino acid residues of $Mb_{BTIC-C2}$ contribute to stabilization of the **$TS2_{C2}$** transition state (**Figure 6a**). For these studies, the lowest energy spin-state **$TS2_{C2}$** transition state optimized from DFT calculations was docked into the full protein structure in the presence of explicit water as solvent. Two independent 500ns MD simulations starting from these optimized $TS2_{C2}$ docked near the enzyme "gate" resulted in the exploration of binding poses with the aromatic end of the substrate to deep inside the active site pocket along the MD trajectories. Steric clashes with residues Phe64, Ile43, Val107, and Leu29 induce a rotation of the benzothiophene ring away from the solvent exposed active site "gate" in the hemoprotein. Such conformational flexibility is consistent with the different orientations found for this residue in the crystal structure and further reveals its role as a "gatekeeper" residue between the heme pocket and the solvent.

To rationalize the structure-activity trends observed for the $Mb_{BTIC-C2}$ biocatalyst (**Figure 3a**), equivalent **TS2**-docked MD simulations were performed using 5-methoxy- (5OMe; **2f**) and 6-methyl-benzothiophene (6Me; **2j**) derivatives as the substrates. While the overall structural features of the resulting **$TS2_{C2}$** complexes were similar to those obtained with **1a**, important differences were also evident. With the 5OMe substrate (**2f**), rotation of the benzothienyl-ester moiety about the Fe-porphyrin ring occurs much faster along MD trajectories than observed with substrate **1a** (**Figure 6b**). In addition, during the MD trajectory the molecule reorients such that the thiophene ring lies above heme ring C (*vs.* ring B for **1a**) and the polar oxygen atom of the methoxy group is pointed toward the solvent (**Figure 6c**). In contrast, MD simulations with the 6Me-substituted substrate (**2j**) showed the benzothiophene ring as embedded deeply in the enzyme pocket and shielded away from the solvent (**Figure 6c**), adopting a conformation that differs from those observed with substrate **1a** and the 5-OMe substrate. Experimentally, the enzyme activity toward these substrates follows the order **2f** (5-OMe) > **1a** > **2j** (6-Me) (**Figure 3a**). Since this trend correlates with the degree of exposure of the benzothiophene ring to the solvent in the corresponding TS as determined by MD simulations, we hypothesize that substrates capable of adopting more solvent exposed orientations during the cyclopropanation step may be more efficiently processed by the enzyme.

Modelling of MbC3-catalyzed reaction via docking and MD simulations. To better understand the reactivity of $Mb_{BTIC-C3}$ and its lineage, substrate **3a** as well as the **3a**-derived carbene were docked into the crystal structure of $Mb_{BTIC-C3}$. In both cases, the lowest-energy binding pose from docking predictions shows that the benzothiophene ring is exposed to the solvent ('BT out' conformation) and binds to a crevice on the protein surface created by the H64V mutation (**Figure 6d**). Interestingly, similar docking studies on Mb(H64V,H64V), the earliest intermediate in the $Mb_{BTIC-C3}$ evolutionary lineage (**Figure 2c**), predict an alternate conformation ('BT in' conformation) in which the benzothiophene ring is buried inside the protein (**Figure S7)**. In $Mb_{BTIC-C3}$, the latter conformation is prevented by steric clashes with the bulky

Phe residue at position 29, introduced through the highly beneficial L29F mutation (**Figure 3b**). Thus, in addition to dictating high enantioselectivity in the cyclopropanation step (**Figure 5d**), the L29F mutation contributes to favor the 'BT out' conformation, which is associated with more efficient catalysis as described earlier.

MD simulations were then performed using the lowest in energy **TS2$_{C3}$** transition state docked into the Mb$_{BTIC-C3}$ active site as starting point. Interestingly, these simulations showed how, starting from the 'BT out' conformation, the heme-bound substrate briefly samples a slightly different conformation during the MD trajectory, in which the benzothiophene ring pushes away the nearby residue Phe43 (**Figure 6f**). However, the substrate quickly reorients itself to a final orientation that is nearly identical to the starting point, i.e., with the benzothiophenyl ring lying over ring D of the heme (**Figure 6e,f**). As made evident from these studies, both the accessible conformations and the preferred orientation of the heme-bound **TS2$_{C3}$** in Mb$_{BTIC-C3}$ (**Figure 6e,f**) differs completely from those found for the heme-bound **TS2$_{C2}$** in Mb$_{BTIC-C2}$ (**Figure 6a,b**), where the benzothiophene ring is pushed above ring B of the heme (vs. ring D for **3a**) and toward the inner side of the heme pocket by the concerted action of residues Leu29 and Phe64. These features can thus explain the orthogonal reactivity of these biocatalysts toward the cyclopropanation of the C2- vs. C3-functionalized benzothiophene substrates due to different active site requirements in order to stabilize the different key transition states **TS2$_{C2}$** and **TS2$_{C3}$**.

Rational design of improved biocatalyst Mb$_{BTIC-C3}$$^{+}$. Leveraging the mechanistic insights from the studies above, we sought to understand and potentially overcome the substrate scope limitations of Mb$_{BTIC-C3}$, which show very limited tolerance toward substitutions at the C5 position of the benzothiophene ring (**Figure 3b**). MD simulations with **TS2**-bound in Mb$_{BTIC-C3}$ for the inactive 5-bromo-substituted substrate (C5-Br) showed that the ester group of the substrate lies at a close distance (despite variations of up to 7.8 Å, distances < 3.5 Å are observed consistently throughout the entire 500 ns MD simulation) from Leu107, suggesting unfavorable steric clashes with it that may prevent this productive binding mode. Based on these analyses, we designed a new variant, called Mb$_{BTIC-C3}$$^{+}$, in which a L107V mutation is introduced which reduces such steric clashes facilitating the productive binding mode for **TS2** to take place, as confirmed by MD simulations with the C5-Br/Mb$_{BTIC-C3}$$^{+}$ system (**Figure 6g**). Gratifyingly, Mb$_{BTIC-C3}$$^{+}$ was determined to exhibit up to 5-fold improved activity over Mb$_{BTIC-C3}$ for the synthesis of C5-substituted substrates **4b** and **4e** (**Figure 3b**). Furthermore, albeit in low to moderate yields (7-20%), this biocatalyst is now able to produce the 5-chloro- and 5-bromo-substituted products **4c** and **4d**, which were unattainable using Mb$_{BTIC-C3}$ (**Figure 3b**). While expanding the substrate scope of the present method, these results highlight the value of combining evolutionary and rational strategies toward the development of these carbene transferases.

Conclusion

We report the development of a novel biocatalytic strategy for the construction of sp^3-rich tetracyclic compounds via the intramolecular cyclopropanation of benzothiophenes. These molecular scaffolds, which are unexplored in medicinal chemistry, embody several highly-sought after features for use as 'fragments' in fragment-based drug discovery, owing to their 3D and stereochemical complexity combined with Ro3-compliant properties (**Figure 1c**). Despite the inherent challenge posed by these transformations using chemical and biological carbene transfer catalysts alike (**Figure 1a; Table S1**), two complementary biocatalysts could be evolved from myoglobin to enable the highly stereoselective cyclization of both C2- and C3-functionalized benzothiophenes and related heterocycles. Our mechanistic and DFT studies demonstrate a diradical mechanism for these intramolecular cyclopropanation reactions, which contrasts with the concerted mechanism previously observed for myoglobin-catalyzed intermolecular cyclopropanation with diazoesters,[42] highlighting the mechanistic plasticity of this hemoprotein in carbene transfer reactions. Crystallographic and computational (MD) studies have provided additional insights into structural features of the two evolved biocatalysts that are responsible for the high level of stereocontrol in the corresponding reactions as well as for their orthogonal (and complementary) reactivity. We also used the computational modelling information to design and expand the substrate scope of the $Mb_{BTIC\text{-}C3}$ enzyme to enable the transformation of a series of previously unreactive substrates. These results highlight the benefits of combining evolutionary and rational strategies for the development of enzymes to catalyze synthetically useful transformations not found in nature.

Declarations

SUPPORTING INFORMATION

Experimental procedures including synthetic methods, compound characterization data, computational methods, atom coordinates of DFT models, additional figures and tables, Nature Research reporting summaries, details of author contributions and competing interests; and statements of data and code availability are available online.

DATA AVAILABILITY STATEMENT

Data that support the findings of this study are included in this published article (and its supplementary information files). Additional datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request. Crystallographic data for small molecules have been deposited in the Cambridge Crystallographic Data Centre (CCDC), while protein crystal structures reported in this manuscript have been deposited in the Protein Data Bank (PDB) as described in the supplementary information files.

CODE AVAILABILITY

No custom computer codes or mathematical algorithms have been used for this research.

COMPETING INTERESTS

The authors have no competing interests.

AUTHOR INFORMATION

Corresponding Authors:

* rfasan@ur.rochester.edu; houk@chem.ucla.edu

ACKNOWLEDGMENTS

This work was supported by the U.S. National Institute of Health grant GM098628 (R.F.). D.V. acknowledges support from the National Science Foundation Graduate Fellowship Program. K. N.H. and A.S. acknowledges support from the National Science Foundation CHE-1764328. M.G.B. acknowledges the support from the Spanish MICINN (Ministerio de Ciencia e Innovación) project PID2019-111300GA-I00, and the Ramón y Cajal program via the RYC 2020-028628-I fellowship. The authors are grateful to Dr. William Brennessel and Dr. Jermaine Jenkins at the University of Rochester for assistance with crystallographic analyses. MS and X-ray instrumentation at the University of Rochester are supported by U.S. National Science Foundation grants CHE-0946653 and CHE-1725028 and the U.S. National Institute of Health grant S10OD030302.

References

1 Bornscheuer, U. T., Huisman, G. W., Kazlauskas, R. J., Lutz, S., Moore, J. C. & Robins, K. Engineering the third wave of biocatalysis. *Nature* **485**, 185-194, (2012).

2 Devine, P. N., Howard, R. M., Kumar, R., Thompson, M. P., Truppo, M. D. & Turner, N. J. Extending the application of biocatalysis to meet the challenges of drug development. *Nat. Rev. Chem.* **2**, 409-421, (2018).

3 Slagman, S. & Fessner, W. D. Biocatalytic routes to anti-viral agents and their synthetic intermediates. *Chem Soc Rev* **50**, 1968-2009, (2021).

4 Savile, C. K., Janey, J. M., Mundorff, E. C., Moore, J. C., Tam, S., Jarvis, W. R., Colbeck, J. C., Krebber, A., Fleitz, F. J., Brands, J., Devine, P. N., Huisman, G. W. & Hughes, G. J. Biocatalytic Asymmetric Synthesis of Chiral Amines from Ketones Applied to Sitagliptin Manufacture. *Science* **329**, 305-309, (2010).

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

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

7 Brandenberg, O. F., Fasan, R. & Arnold, F. H. Exploiting and engineering hemoproteins for abiological carbene and nitrene transfer reactions. *Curr. Opin. Biotech.* **47**, 102-111, (2017).

8 Schwizer, F., Okamoto, Y., Heinisch, T., Gu, Y. F., Pellizzoni, M. M., Lebrun, V., Reuter, R., Kohler, V., Lewis, J. C. & Ward, T. R. Artificial Metalloenzymes: Reaction Scope and Optimization Strategies. *Chem. Rev.* **118**, 142-231, (2018).

9 Ren, X. & Fasan, R. Engineered and Artificial Metalloenzymes for Selective C-H Functionalization. *Curr. Opin. Green Sustain. Chem.* **31**, (2021).

10 Liu, Z. & Arnold, F. H. New-to-nature chemistry from old protein machinery: carbene and nitrene transferases. *Curr. Opin. Biotech.* **69**, 43-51, (2021).

11 Ramsden, J. I., Cosgrove, S. C. & Turner, N. J. Is it time for biocatalysis in fragment-based drug discovery? *Chem. Sci.* **11**, 11104-11112, (2020).

12 Erlanson, D. A., Fesik, S. W., Hubbard, R. E., Jahnke, W. & Jhoti, H. Twenty years on: the impact of fragments on drug discovery. *Nat. Rev. Drug Discov.* **15**, 605-619, (2016).

13 Erlanson, D. A., Davis, B. J. & Jahnke, W. Fragment-Based Drug Discovery: Advancing Fragments in the Absence of Crystal Structures. *Cell Chem. Biol.* **26**, 9-15, (2019).

14 Jhoti, H., Williams, G., Rees, D. C. & Murray, C. W. The 'rule of three' for fragment-based drug discovery: where are we now? *Nat. Rev. Drug Discov.* **12**, 644-+, (2013).

15 Murray, C. W. & Rees, D. C. Opportunity Knocks: Organic Chemistry for Fragment-Based Drug Discovery (FBDD). *Angew. Chem. Int. Ed.* **55**, 488-492, (2016).

16 Morley, A. D., Pugliese, A., Birchall, K., Bower, J., Brennan, P., Brown, N., Chapman, T., Drysdale, M., Gilbert, I. H., Hoelder, S., Jordan, A., Ley, S. V., Merritt, A., Miller, D., Swarbrick, M. E. & Wyatt, P. G. Fragment-based hit identification: thinking in 3D. *Drug Discov Today* **18**, 1221-1227, (2013).

17 Over, B., Wetzel, S., Grutter, C., Nakai, Y., Renner, S., Rauh, D. & Waldmann, H. Natural-product-derived fragments for fragment-based ligand discovery. *Nat. Chem.* **5**, 21-28, (2013).

18 Yu, B., Zheng, Y. C., Shi, X. J., Qi, P. P. & Liu, H. M. Natural Product-Derived Spirooxindole Fragments Serve as Privileged Substructures for Discovery of New Anticancer Agents. *Anti-Cancer Agent Me* **16**, 1315-1324, (2016).

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

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

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

22 Key, H. M., Dydio, P., Liu, Z. N., Rha, J. Y. E., Nazarenko, A., Seyedkazemi, V., Cark, D. S. & Hartwig, J. F. Beyond Iron: Iridium-Containing P450 Enzymes for Selective Cyclopropanations of Structurally Diverse Alkenes. *ACS Centr. Sci.* **3**, 302-308, (2017).

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

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

25 Wittmann, B. J., Knight, A. M., Hofstra, J. L., Reisman, S. E., Kan, S. B. J. & Arnold, F. H. Diversity-Oriented Enzymatic Synthesis of Cyclopropane Building Blocks. *ACS Catal.* **10**, 7112-7116, (2020).

26 Nam, D., Steck, V., Potenzino, R. J. & Fasan, R. A Diverse Library of Chiral Cyclopropane Scaffolds via Chemoenzymatic Assembly and Diversification of Cyclopropyl Ketones. *J Am Chem Soc* **143**, 2221-2231, (2021).

27 Srivastava, P., Yang, H., Ellis-Guardiola, K. & Lewis, J. C. Engineering a dirhodium artificial metalloenzyme for selective olefin cyclopropanation. *Nat. Commun.* **6**, 7789, (2015).

28 Sreenilayam, G., Moore, E. J., Steck, V. & Fasan, R. Metal substitution modulates the reactivity and extends the reaction scope of myoglobin carbene transfer catalysts. *Adv. Synth. Cat.* **359**, 2076–2089, (2017).

29 Dydio, P., Key, H. M., Nazarenko, A., Rha, J. Y., Seyedkazemi, V., Clark, D. S. & Hartwig, J. F. An artificial metalloenzyme with the kinetics of native enzymes. *Science* **354**, 102-106, (2016).

30 Ohora, K., Meichin, H., Zhao, L. M., Wolf, M. W., Nakayama, A., Hasegawa, J., Lehnert, N. & Hayashi, T. Catalytic Cyclopropanation by Myoglobin Reconstituted with Iron Porphycene: Acceleration of Catalysis due to Rapid Formation of the Carbene Species. *J Am Chem Soc* **139**, 17265-17268, (2017).

31 Villarino, L., Splan, K. E., Reddem, E., Alonso-Cotchico, L., de Souza, C. G., Lledos, A., Marechal, J. D., Thunnissen, A. M. W. H. & Roelfes, G. An Artificial Heme Enzyme for Cyclopropanation Reactions. *Angew. Chem. Int. Ed.* **57**, 7785-7789, (2018).

32 Zhao, J. M., Bachmann, D. G., Lenz, M., Gillingham, D. G. & Ward, T. R. An artificial metalloenzyme for carbene transfer based on a biotinylated dirhodium anchored within streptavidin. *Catal. Sci. Technol.* **8**, 2294-2298, (2018).

33 Carminati, D. M. & Fasan, R. Stereoselective Cyclopropanation of Electron-Deficient Olefins with a Cofactor Redesigned Carbene Transferase Featuring Radical Reactivity. *ACS Catal.* **9**, 9683-9697, (2019).

34 Stenner, R., Steventon, J. W., Seddon, A. & Anderson, J. L. R. A de novo peroxidase is also a promiscuous yet stereoselective carbene transferase. *Proc. Natl. Acad. Sci. USA* **117**, 1419-1428, (2020).

35 Chandgude, A. L., Ren, X. & Fasan, R. Stereodivergent Intramolecular Cyclopropanation Enabled by Engineered Carbene Transferases. *J Am Chem Soc* **141**, 9145-9150, (2019).

36 Ren, X. K., Liu, N. Y., Chandgude, A. L. & Fasan, R. An Enzymatic Platform for the Highly Enantioselective and Stereodivergent Construction of Cyclopropyl-delta-lactones. *Angew. Chem. Int. Ed.* **59**, 21634-21639, (2020).

37 Yong, K., Salim, M. & Capretta, A. Intramolecular carbenoid insertions: Reactions of alpha-diazo ketones derived from furanyl-, thienyl-, (benzofuranyl)-, and (benzothienyl)acetic acids with rhodium(II) acetate. *J. Org. Chem.* **63**, 9828-9833, (1998).

38 Padwa, A., Wisnieff, T. J. & Walsh, E. J. Synthesis of Cycloalkenones Via the Intramolecular Cyclopropanation of Furanyl Diazo Ketones. *J. Org. Chem.* **51**, 5036-5038, (1986).

39 Ryan, J., Siauciulis, M., Gomm, A., Macia, B., O'Reilly, E. & Caprio, V. Transaminase Triggered Aza-Michael Approach for the Enantioselective Synthesis of Piperidine Scaffolds. *J Am Chem Soc* **138**, 15798-15800, (2016).

40 Zawodny, W., Montgomery, S. L., Marshall, J. R., Finnigan, J. D., Turner, N. J. & Clayden, J. Chemoenzymatic Synthesis of Substituted Azepanes by Sequential Biocatalytic Reduction and Organolithium-Mediated Rearrangement. *J Am Chem Soc* **140**, 17872-17877, (2018).

41 Zhang, R. K., Chen, K., Huang, X., Wohlschlager, L., Renata, H. & Arnold, F. H. Enzymatic assembly of carbon-carbon bonds via iron-catalysed sp(3) C-H functionalization. *Nature* **565**, 67-72, (2019).

42 Wei, Y., Tinoco, A., Steck, V., Fasan, R. & Zhang, Y. Cyclopropanations via Heme Carbenes: Basic Mechanism and Effects of Carbene Substituent, Protein Axial Ligand, and Porphyrin Substitution. *J Am Chem Soc* **140**, 1649-1662, (2018).

43 de Visser, S. P., Ogliaro, F., Harris, N. & Shaik, S. Multi-state epoxidation of ethene by cytochrome P450: A quantum chemical study. *J Am Chem Soc* **123**, 3037-3047, (2001).

44 Kondrashov, D. A., Zhang, W., Aranda, R. t., Stec, B. & Phillips, G. N., Jr. Sampling of the native conformational ensemble of myoglobin via structures in different crystalline environments. *Proteins* **70**, 353-362, (2008).

45 Blomberg, M. R. A., Borowski, T., Himo, F., Liao, R. Z. & Siegbahn, P. E. M. Quantum Chemical Studies of Mechanisms for Metalloenzymes. *Chem. Rev.* **114**, 3601-3658, (2014).

46 Himo, F. Recent Trends in Quantum Chemical Modeling of Enzymatic Reactions. *J Am Chem Soc* **139**, 6780-6786, (2017).

47 Lind, M. E. S. & Himo, F. Quantum Chemistry as a Tool in Asymmetric Biocatalysis: Limonene Epoxide Hydrolase Test Case. *Angew. Chem. Int. Ed.* **52**, 4563-4567, (2013).

48 Liao, R. Z. & Thiel, W. On the Effect of Varying Constraints in the Quantum Mechanics Only Modeling of Enzymatic Reactions: The Case of Acetylene Hydratase. *J. Phys. Chem. B* **117**, 3954-3961, (2013).

Figures

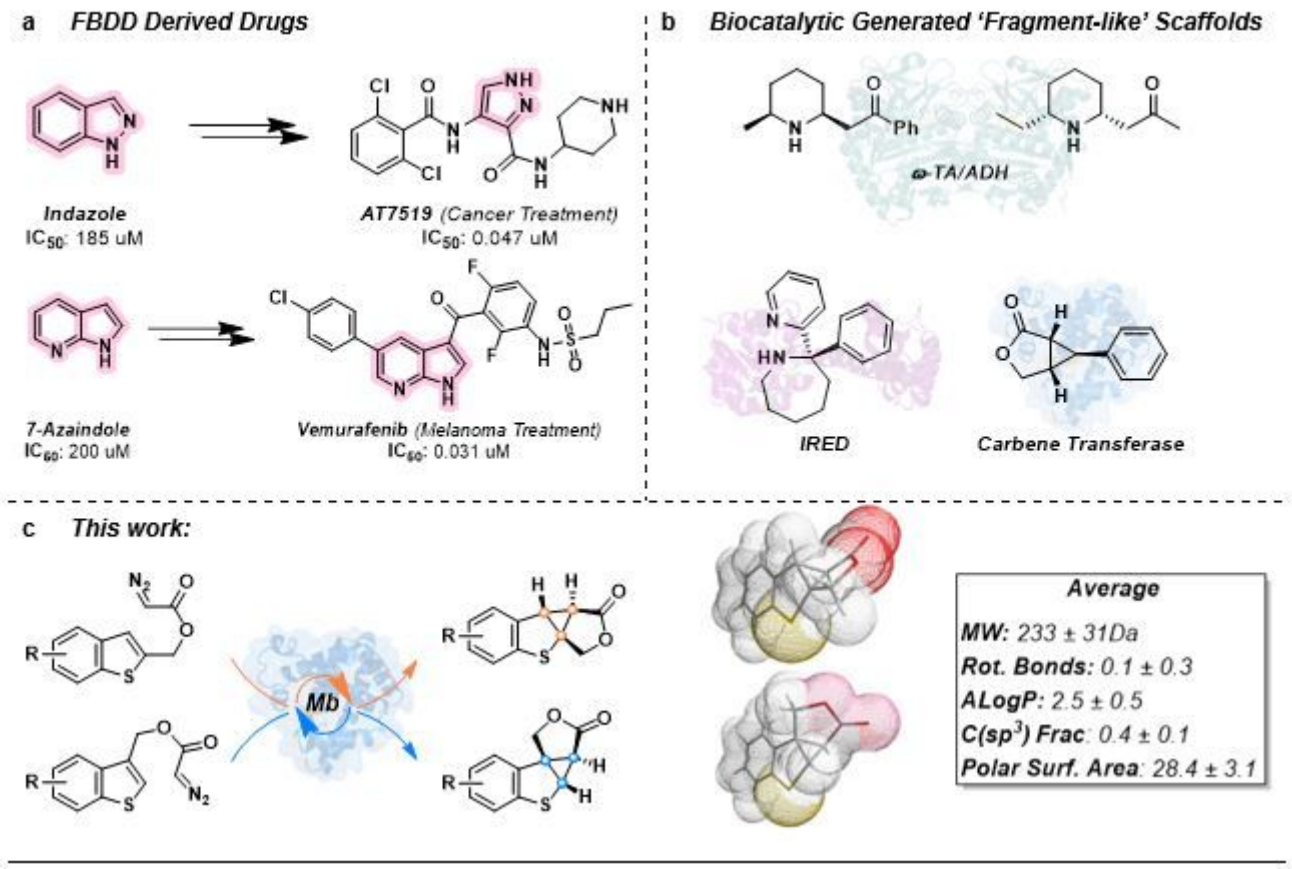

Figure 1

***Sp*^{2} vs. *sp*^{3}-rich molecular 'fragments' for applications in fragment-based drug discovery**. **(a)** Representative drug molecules discovered and developed via fragment-based drug discovery (FBDD). **(b)** 'Fragment-like' scaffolds generated via biocatalysis include chiral piperidines,[39] azepans,[40] and cyclopropyl-g-lactones.[35] **(c)** Biocatalytic intramolecular cyclopropanation of C2- and C3-functionalized benzothienyl substrates to give *sp*^{3}-rich tetracyclic scaffolds with distinct 3D shapes and Ro3-compliant properties.

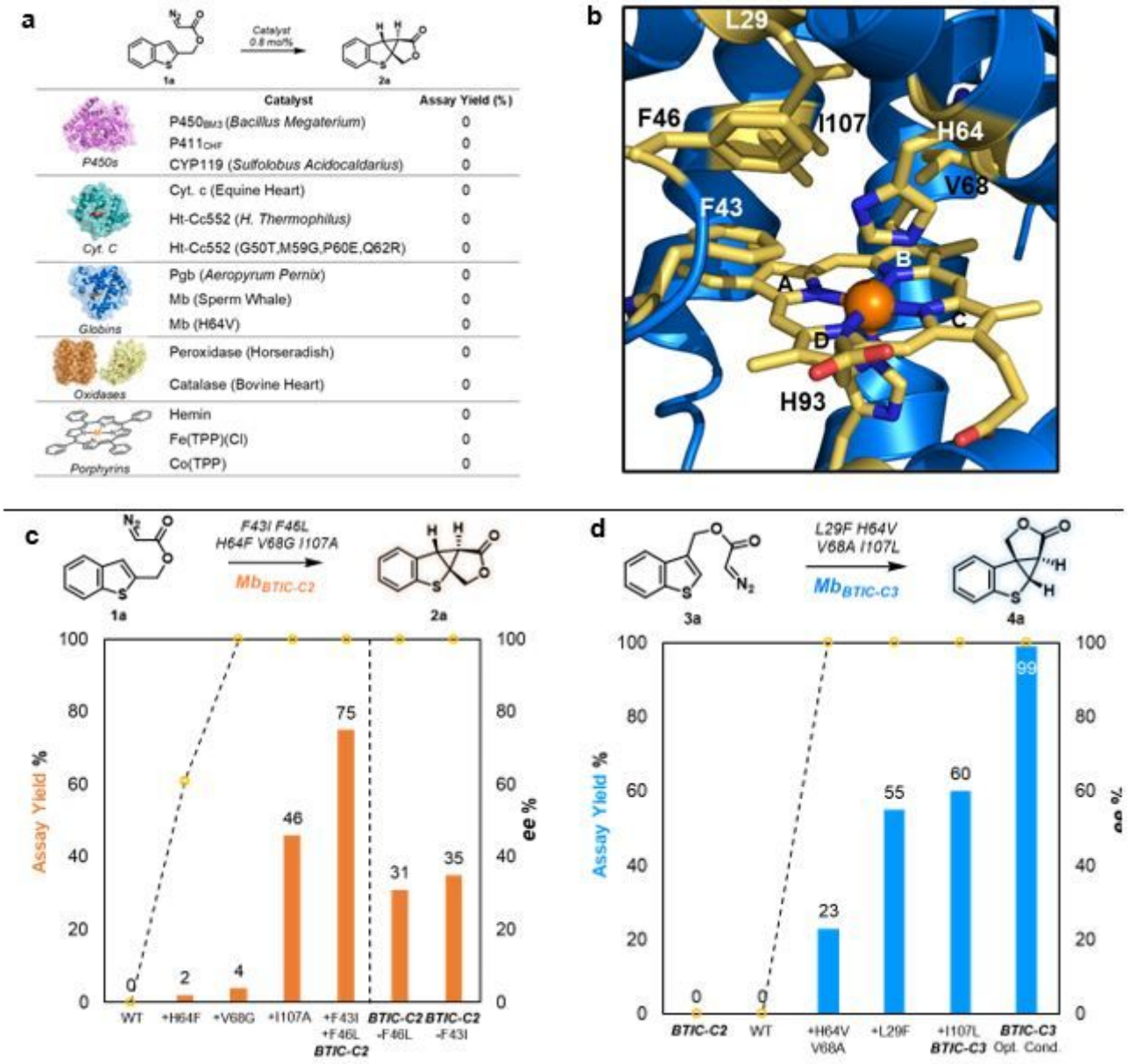

Figure 2

Biocatalytic intramolecular cyclopropanation of benzothiophenyl substrates. a) Activity of a diverse panel of hemoproteins and porphyrin catalysts toward the intramolecular cyclopropanation of **1a**. **b)** Active site of sperm whale myoglobin (Mb; pdb 1JW8) with the heme and surrounding amino acid residues highlighted as sticks in gold. Heme b pyrrole rings labeled according to convention. **c-d)** Activity and enantioselectivity of Mb variants in the intramolecular cyclopropanation of **1a** or **3a** along the evolutionary paths leading to $Mb_{BTIC\text{-}C2}$ (**c**) and $Mb_{BTIC\text{-}C3}$ (**d**). Assay yields (AY) were determined by GC analysis using calibration curves generated with isolated products ($n \geq 2$; SE < 10%).

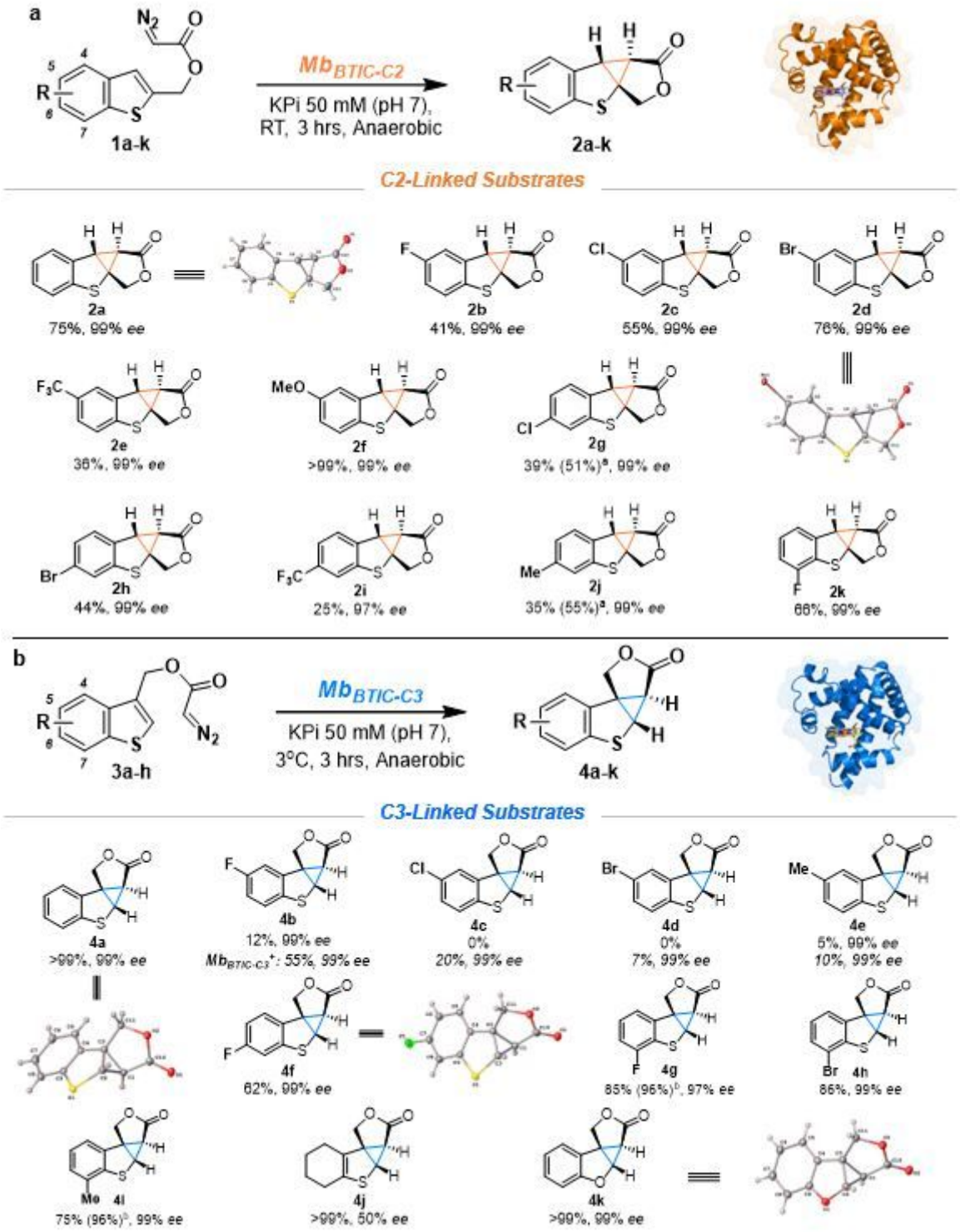

Figure 3

Substrate scope of $Mb_{BTIC-C2}$ and $Mb_{BTIC-C2}$ biocatalysts. Reaction conditions: 20 μM Mb catalyst, 2.5 mM diazo compound, 10 mM $Na_2S_2O_4$, 3 hrs, anaerobic conditions. All reactions with $Mb_{BTIC-C3}$ and $Mb_{BTIC-C3}^+$ were carried out using slow diazo addition at 3°C. [a] Using whole cells (OD_{600} = 40). [b] Using 40 μM Mb catalyst.

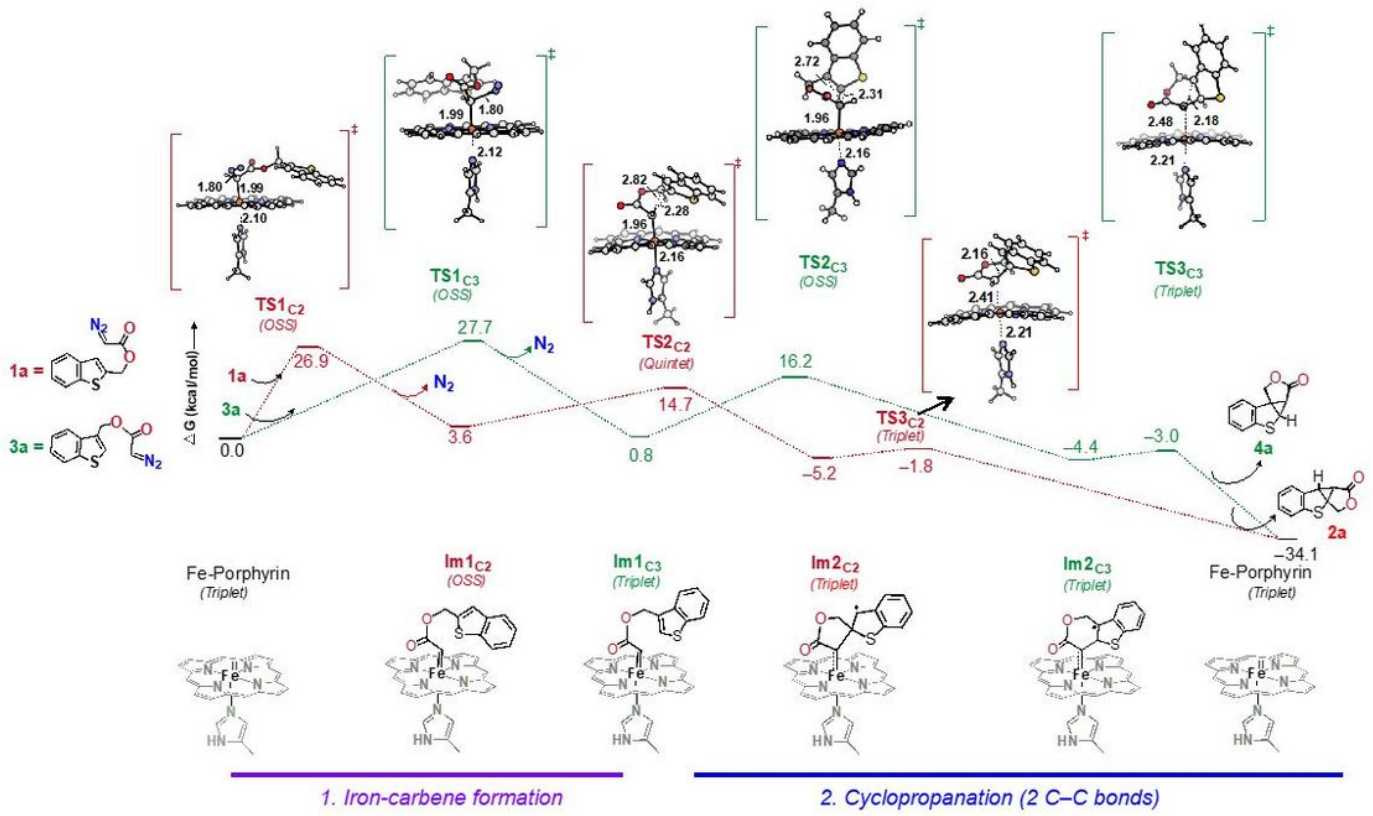

Figure 4

DFT analysis of the reaction mechanism for intramolecular cyclopropanation of C2- and C3-benzothiophenyl-methyl-diazoacetate catalyzed by a truncated iron porphyrin with an axial 4-methylimidazole ligand as a simplified model. The reaction proceeds via a first (1) iron–carbene formation followed by (2) cyclopropanation (2 C–C bonds). ΔG values are calculated at the B3LYP-D3BJ/def2TZVP (SMD, ε=4) // B3LYP-D3BJ/6-31G(d)+SDD (Fe) level. For each stationary point, the Gibbs free energy is provided for its lowest energy spin state. A detailed free energy profiles are provided in **Figure S8** in Supporting Information.

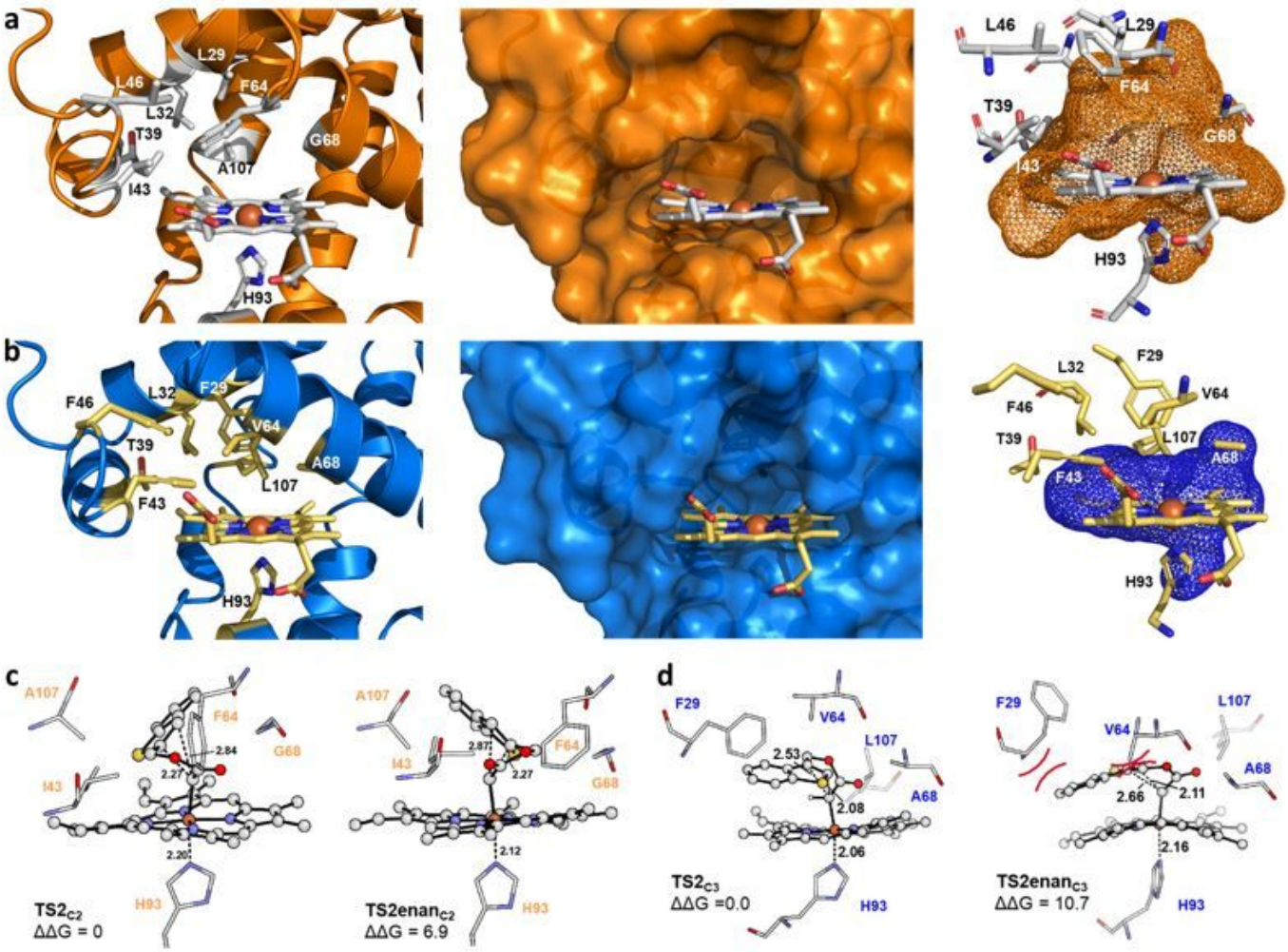

Figure 5

Crystal structures of $Mb_{BTIC\text{-}C2}$ and $Mb_{BTIC\text{-}C2}$ and corresponding stereochemical models. a-b, High-resolution crystal structure of (**a**) $Mb_{BTIC\text{-}C2}$ and (**b**) $Mb_{BTIC\text{-}C3}$, highlighting the amino acid residues surrounding the heme cofactor (*left panel*), molecular surface representation of the heme binding site (*middle panel*), and the active site volume (*right panel*). **c-d,** Optimized geometries of (**a**) **$TS2_{C2}$** and its enantiomeric transition state (**$TS2enan_{C2}$**) in the theozyme model of $Mb_{BTIC\text{-}C2}$, and **(d) $TS2_{C3}$** and its enantiomeric transition state (**$TS2enan_{C3}$**) in the theozyme model of $Mb_{BTIC\text{-}C3}$. Relative Gibbs energies (ΔΔG) are given in kcal·mol^{-1}, and all distances are in angstroms (Å). See computational details in the Supporting Information. See also **Figure S4**.

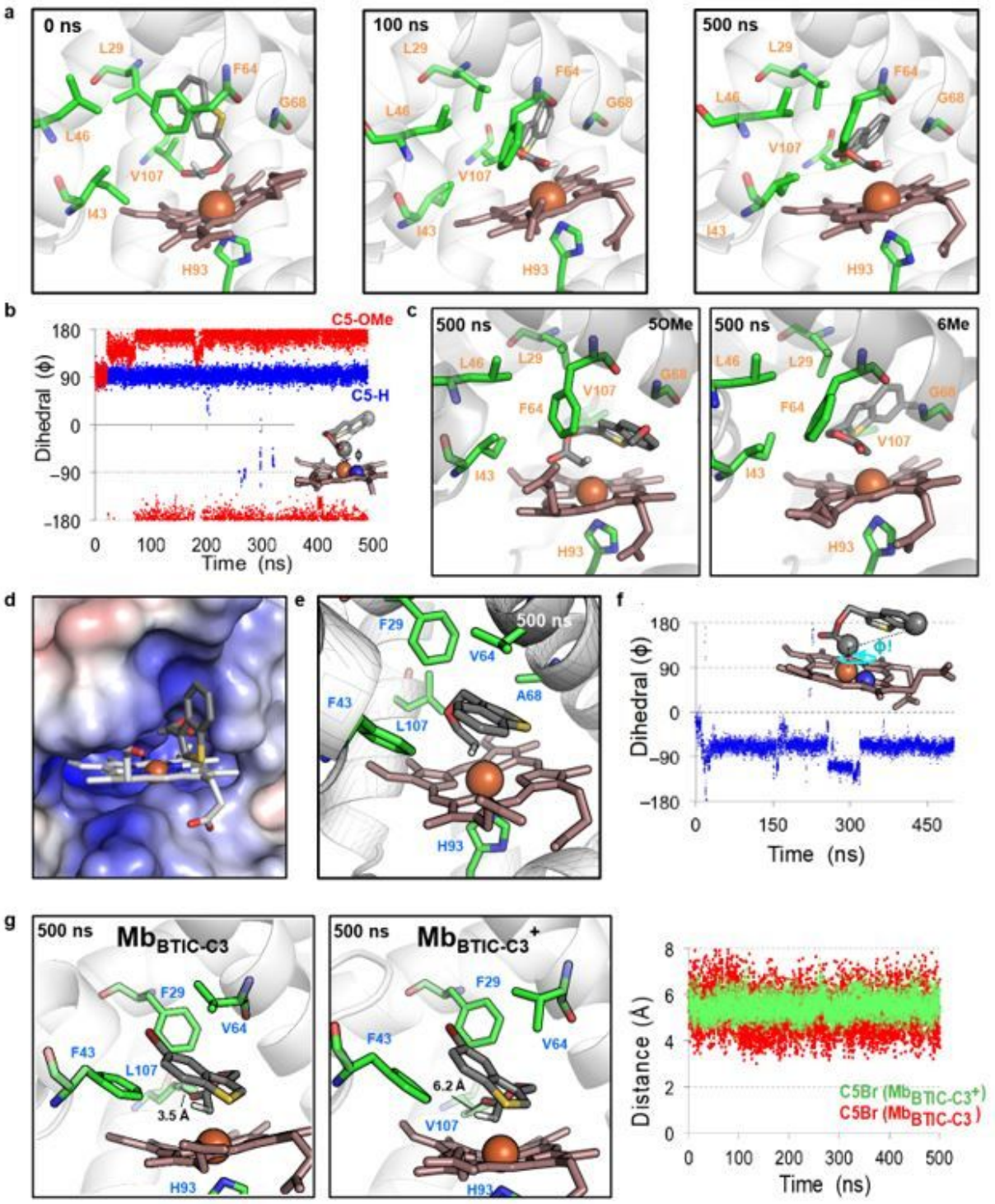

Figure 6

Analysis of the $Mb_{BTIC\text{-}C2}$, $Mb_{BTIC\text{-}C3}$, and $Mb_{BTIC\text{-}C3}^{+}$ catalyzed cyclization reactions via molecular dynamics (MD) simulations. **a,** MD snapshots for the **$TS2_{C2}$** docked in $Mb_{BTIC\text{-}C2}$ protein at 0, 100, and 500ns. **b,** Variation of the pseudo dihedral angle (ϕ) in the MD simulation for the heme-carbene intermediates corresponding to the 5-OMe and 6-Me substrates (i.e., compounds **2f** and **2j**) in $Mb_{BTIC\text{-}C2}$.

The atoms used for measurement of the ϕ angle are displayed as spheres. **c,** End-of-simulation (500 ns) poses for the **TS2$_{C2}$** transition states corresponding to the 5-OMe and 6-Me substrates in Mb$_{BTIC-C2}$. See **Figure S5** for additional time points. **d,** Lowest-energy pose for the diazo substrate **3a** after molecular docking into the structure of Mb$_{BTIC-C3}$, showing the benzothiophene ring in the 'BT out' conformation. **e,** End-of-simulation (500 ns) pose for the **TS2$_{C3}$** complex in Mb$_{BTIC-C3}$. See **Figure S6** for additional time points. **f,** Variation of pseudo dihedral angle (ϕ!) across the different geometries of the heme-bound **TS2$_{C3}$** complex sampled during the 500 ns MD simulation. The reference atoms for the ϕ! angle are displayed as spheres. **g,** MD snapshots at 500 ns of the **TS2$_{C3}$** analogues of the 5-Br substituted substrates (compound **3d**) in Mb$_{BTIC-C3}$ vs. Mb$_{BTIC-C3}^{+}$. The closest distance between the carbonyl (C=O) oxygen atom and the Leu107 residue in these complexes is indicated. The graph compares the variation of these distances for the C5-Br substrates in Mb$_{BTIC-C3}$ vs. Mb$_{BTIC-C3}^{+}$ throughout the 500 ns MD simulation.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- VargasetalSI.pdf
- CrystalStructures.zip