

Chemical Science

Accepted Manuscript

Published on 29 September 2026

Licensed under CC-BY-NC 4.0



This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.

Dynamic Charge Distribution as a Key Driver of Catalytic Reactivity in an Artificial Metalloenzyme

David W. Kastner^{1,2}, Linda Leone⁴, Clorice R. Reinhardt¹, Husain Adamji¹, Melissa T. Manetsch,¹ Alessandra Esposito⁴, Flavia Nastro⁴, Yuriy Román-Leshkov^{1,3}, Angela Lombardi⁴, and Heather J. Kulik^{1,3*}

¹*Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

²*Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

³*Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

⁴*Department of Chemical Sciences, University of Naples Federico II, 80126 Naples, Italy*

ABSTRACT: Miniature artificial metalloenzymes provide a simplified platform to extract design principles for engineering catalysts that outperform natural enzymes. Among them, mimochromes (MCs) emerged as excellent candidates for replicating the activity of peroxidases in a compact scaffold composed of a heme center embedded between two helical peptides. However, design optimizations of MCs have largely focused on structural properties, leaving the impact of the electronic environment less explored. To investigate how the electronic environment influences reactivity, we carry out both classical and *ab initio* molecular dynamics (MD) simulations of a series of mimochromes, MC6, MC6*, and MC6*a, with known differences in reactivity. To uncover patterns in electronic structure, we train supervised machine learning (ML) models and carry out statistical analysis. The *ab initio* MD simulations uncover key charge coupling interactions between Arg10' and Glu2' that are not identified in classical dynamics. These residues are determined to be essential in orienting Arg10' in the active site. Furthermore, using our ML classifiers trained to distinguish mimochromes based on partial charge dynamics alone, we identify critical residues such as Arg10', Lys11, Glu2', and Ser6' that distinguish the mimochrome analogues. Disruption of the computationally-identified Glu2'...Lys11 salt bridge in a new mimochrome analogue, MC6*a-LysBoc, that we synthesize and characterize, experimentally supports this prediction by resulting in a seven-fold decrease in the total turnover number. Quantum mechanical calculations show that Arg10' stabilizes the transition state for the formation of the reactive oxo from hydrogen peroxide. These results indicate that partial charge distribution dynamics are important in the reactivity series in mimochromes, which could provide a complementary method to augment static electric fields

and long-range correlated motions for the engineering of a broad array of next-generation metalloenzymes.

1. Introduction.

Enzymes are versatile catalysts that enable the biosynthesis of complex chemical compounds with high chemo-, regio-, and stereospecificity under mild conditions.¹ The growing need for green catalysts has sparked interest in applying computational approaches and directed evolution to engineer non-native enzymes with enhanced reactivity and the capability to perform novel chemical transformations.²⁻⁸ Heme proteins, which contain a protoporphyrin IX cofactor typically with a coordinated iron cation, are of particular interest as they catalyze a remarkable array of challenging reactions, including hydroxylation,^{9,10} halogenation,^{11,12} peroxidation,^{13,14} sulfoxidation,^{15,16} and C–C bond cleavage¹⁷, and can be tuned by modifying the protein scaffold and the coordinating environment of the reactive high-valent metal center.¹⁸ The tunability of the protein and metal environments has served as inspiration for an array of *de novo* designed enzymes that display equivalent or superior reactivities to their natural counterparts and can carry out reactions not found in nature.¹⁹⁻²⁴ However, despite successes, the complexity of natural enzymes has made it difficult to extract design principles from structure–function relationships and quantify the core requirements for catalytic activity.²⁵ To facilitate the investigation of the essential aspects of the protein environment in enzymatic heme catalysis, several *de novo* and miniaturized heme-based enzyme systems have been developed.²⁶⁻³⁴

Exemplary of a miniaturized heme-enzyme system, of mimochromes (MCs) consist of a iron-deuteroporphyrin IX that is covalently bound to two short helical peptides.³¹ Due to the simplicity of the protein scaffold, both the structure and reactivity of the mimochrome family has been successfully tuned and optimized over several generations through principles of rational design.³¹ In particular, MC6³⁵ is the first reported catalytically active analogue, composed of a

proximal tetradecapeptide (TD) chain providing an axial histidine, and a distal decapeptide (D) chain, mimicking the five-coordinated heme site of peroxidases. Modifications of MC6³⁵ have led to MC6*³⁶ and MC6*a³⁷, displaying improved catalytic properties compared to the parent analogue. In this work MC6, MC6*, and MC6*a refer to the iron-containing complexes, although other metal-complexes of MC6*a have been reported.³²⁻³⁴ Residues on the distal D chain are designated with a prime symbol (') to distinguish them from the proximal TD chain sequence (Figure 1).³⁸

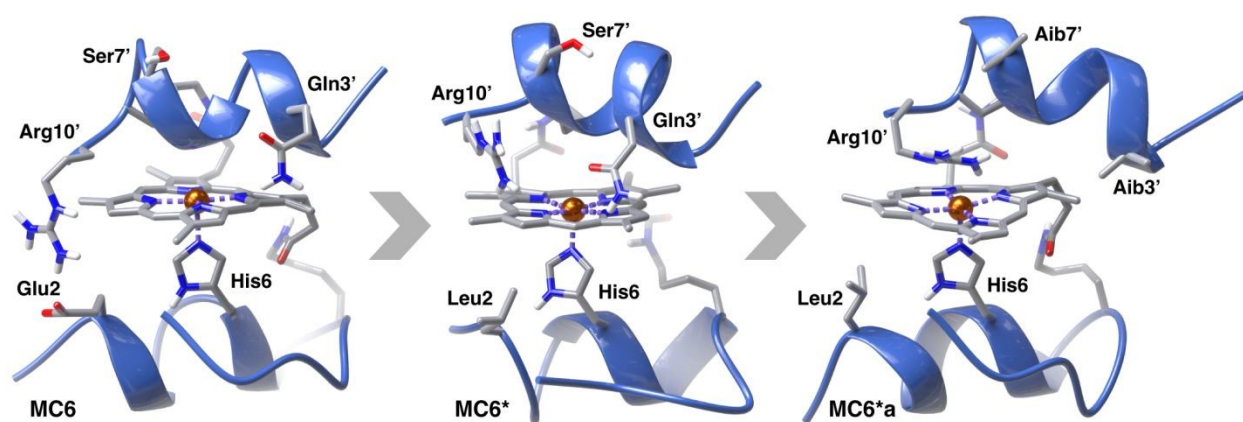


Figure 1. Representative structures of (left) MC6, (middle) MC6*, and (right) MC6*a obtained from the centroids of our MD simulations. Substitution sites and important residues are labeled on the structures. Dark blue cartoon representations are used for the secondary structures and atoms are colored as follows: carbon in gray, nitrogen in blue, oxygen in red, hydrogen in white, and iron in brown. Coordinating bonds are shown as purple dashed lines. The proteins are arranged in the series by which their mutations were carried out, starting from MC6, progressing to MC6*, and finally to MC6*a, and this progression is represented by the gray arrows between the structures.

In the first modification, MC6 was converted to MC6* through the change of an acidic residue (Glu2) on the TD chain³⁶ to a hydrophobic residue (Leu2), which eliminated a salt bridge (Glu2 $\cdots$ Arg10') (Figure 1). This mutation was hypothesized to allow Arg10' to interact with the active site.^{39,40} The resulting flexibility of Arg10' is mechanistically critical, as the analogous distal arginine in natural peroxidases acts as a general acid catalyst that polarizes the O–O bond to promote cleavage and the formation of the reactive ferryl-oxo species (Compound I).⁴¹⁻⁴⁶

Indeed, MC6* displayed a 2-fold higher k_{cat} in the oxidation of 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid (ABTS), together with a ~3-fold improved catalytic efficiency toward both substrates, ABTS and H₂O₂ ($k_{cat}/K_m^{H_2O_2}$ and k_{cat}/K_m^{ABTS}) compared to MC6. In the second iteration, design efforts were focused on improving catalyst robustness. To this aim, two polar residues, Gln3' and Ser7', were both changed to a non-canonical amino acid, aminoisobutyric acid (Aib), which was intended to promote stronger interactions between the D chain and the porphyrin (Figure 1).³⁷ Nuclear magnetic resonance (NMR) chemical shifts revealed below-average values for the Aib residues in MC6*a, suggesting increased protein-heme contact between the porphyrin ring and the Aib residues.³⁷ Analysis of the circular dichroism (CD) spectra in the Soret region further confirms this finding.³⁷ In the context of catalysis, MC6*a displayed a much higher turnover number (TON = 14 000) compared to MC6* (TON = 5 900), demonstrating an increased resistance against oxidative damage. MC6*a shows nearly unchanged $k_{cat}/K_m^{H_2O_2}$ and k_{cat}/K_m^{ABTS} with respect to MC6*, resulting from a doubled k_{cat} value and increased K_m for both substrates, which is compatible with a narrowed and more hydrophobic active site.³⁷ While experimental evidence has provided valuable insights into each iterative mimochrome improvement in activity or structural shifts, the precise mechanisms by which individual substitutions dictate reactivity differences between MC6, MC6*, and MC6*a have not been investigated so far.

As described above, the current explanations for differences in mimochrome reactivity focus on variations of the interactions established by the D chain with the active site, but have not explored potential electronic effects. While geometric and electronic effects are interconnected, even small differences in the electronic environment have shown to have large effects on heme catalysis through changes in the electrostatic potential,⁴⁷ electric fields,⁴⁷⁻⁵⁰ and

local dielectric constants.⁵¹ Of particular interest are the effects of dynamic partial charge coupling, which has been shown computationally to be important in small proteins including to distinguish wild type and mutant species⁵²⁻⁵⁴ and provides an electrostatic corollary to geometric allosteric coupling.⁵⁵ Dynamic motion⁵⁶⁻⁵⁸ and substrate positioning⁵⁹⁻⁶² by the greater enzyme environment is also widely recognized as critical to understanding reaction outcomes of substrates interacting with high-valent Fe(IV)=O species in closely related α -ketoglutarate-dependent, non-heme iron enzymes. Molecular dynamics studies have identified long-range correlated motions connecting distal regions of these enzymes to the active site, with additional implications for substrate positioning, catalytic reactivity, and reaction selectivity.⁵⁶⁻⁵⁸ Unlike structural differences, interrogating differences in the electronic environment is typically difficult to probe experimentally for enzyme systems and often must rely extensively on computational studies.⁶³⁻⁶⁷ While computational studies have been performed on MC6,⁶⁸ no computational analysis of the changes in reactivity across the mimochrome series has been carried out.

In this work, we perform extensive computational analysis across three catalytically active analogues of the mimochrome family (i.e., MC6, MC6*, and MC6*a). Our analysis follows a workflow that progresses from molecular dynamics (MD) simulations and charge coupling to experimental validation, which in turn informs machine learning (ML) models for feature identification and, ultimately, quantum mechanical (QM) reactivity studies (Figure 2). To investigate the role of the electronic environment, we carry out extensive classical and *ab initio* MD simulations of the three mimochromes and observe stark differences in their charge dynamics such as the progressively strengthening Glu2'...Arg10 inter-chain salt bridge and an emergent inter-chain Glu2'...Lys11 interaction in the most robust analogue, MC6*a. Experimental investigation of the previously unreported Glu2'...Lys11 interaction shows that

disrupting this interaction in a new derivative (MC6*a-LysBoc) leads to a seven-fold reduction in TON. These results are further corroborated by supervised ML models, which identify Arg10' as a key feature that distinguishes the mimochromes. Based on the results of the ML models, we demonstrate that Arg10' is critical for lowering the activation free energy of oxo formation. Finally, we show that the electric field projected along the Fe–oxo bond strengthens with reactivity, highlighting opportunities for engineering the electronic environment when designing *de novo* artificial metalloenzymes.

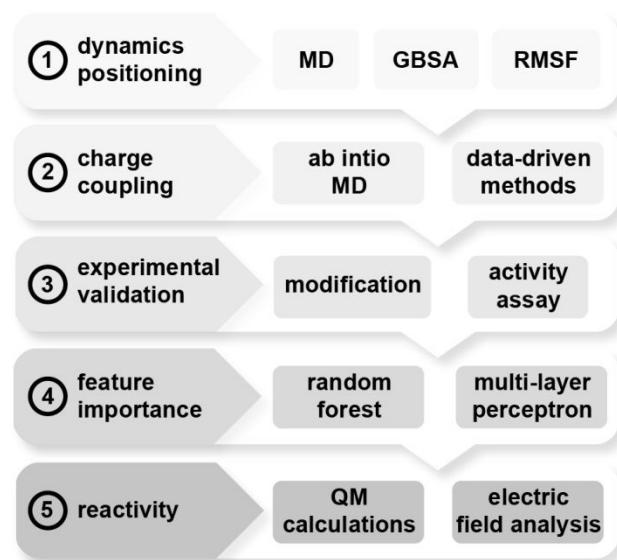


Figure 2. Schematic of the multi-step workflow progressing from (1) molecular dynamics simulations, to (2) ab initio MD for charge coupling analysis, (3) experimental validation, (4) machine learning for feature importance, and concluding with (5) QM calculations to assess reactivity.

2. Computational Details and Experimental Methods.

2a. Classical Molecular Dynamics Simulations and Analysis.

A full flowchart of all calculations and analysis carried out in this work is provided in Supporting Information Figure S1. We performed MD simulations for MC6, MC6*, and MC6*a using AMBER18 and the GPU-accelerated particle mesh Ewald (PME)⁶⁹ molecular dynamics (PMEMD) code.^{70,71} Eight independent MD simulations were run for the resting state and twelve

were run for the oxo-bound state for MC6, MC6*, MC6*a (Supporting Information Text S1). Residue protonation states were assigned based on their default values (i.e., Arg and Lys are positively charged, Glu and Asp are negatively charged, and all other residues are neutral), and the metal center parameter builder (MCPB) was used to derive force field parameters for the metal center in both resting and oxo-bound states (Supporting Information Tables S1–S3 and Figures S2–S3). The equilibration protocol was as follows: i) 1,000 steps of hydrogen atom minimization, 2,000 steps of sidechain minimization with a fixed backbone, and 2,000 steps of unrestrained protein minimization; ii) controlled NVT heating from 0 to 300 K over 10 ps using the Langevin thermostat and a collision frequency of 5.0 ps⁻¹; and iii) a 1 ns NPT simulation with the Berendsen barostat and a 2 ps relaxation time. Following equilibration, we collected 250 ns of production dynamics for each mimochrome in the resting and oxo-bound states, employing 2 fs time steps, and the SHAKE algorithm.⁷² Electrostatics were treated with the PME method with a real-space cutoff of 10 Å.⁶⁹ We subsequently analyzed the structural stability and key non-covalent interactions for each trajectory (Supporting Information Text S2).

2b. Ab Initio Molecular Dynamics Simulations.

Representative frames were selected using the DBSCAN clustering algorithm in CPPTRAJ to obtain clusters and their most representative frames (i.e., cluster centroids).⁷³ The centroid of the largest cluster from each independent simulation was then used as input for semi-empirical quantum mechanical (SQM) MD simulations (Supporting Information Tables S4–S6 and Figures S4–S6).^{59,73} The SQM MD simulations used GFN2-xTB⁷⁴, which incorporates D3 dispersion⁷⁵, as implemented in the TeraChem v1.9 developer version.^{76,77} We employed the conductor-like polarizable continuum model (C-PCM)^{78,79} using a dielectric constant of 80 for water and a cavity derived from 1.2 times Bondi's van der Waals radii⁸⁰ for all atoms, as

implemented in TeraChem.^{81,82} We modeled MC6, MC6*, and MC6*a in the closed-shell singlet (i.e., singlet multiplicity, $2S+1 = 1$) for computational efficiency in sampling protein dynamics, assuming the metal center is Fe(II). For subsequent catalytic cycle modeling, we used revised spin state assignments using Fe(III) as a doublet, as both Fe(II) and Fe(III) are observed experimentally (see Sec. 2c).^{83,84} SQM MD simulations were performed in the NVT ensemble using Langevin dynamics over 20 ps for each of the 8 independent trajectories per mimochrome (Supporting Information Text S3).

2c. Quantum Mechanical Calculations.

A subset of 400 evenly-spaced structures (i.e., every 0.2 ps) were extracted from the SQM MD runs for evaluation of properties with range-separated hybrid DFT. For each snapshot, DFT partial charges were collected at the ω PBEh/LACVP* level of theory using implicit C-PCM solvation, and Mulliken partial charges were summed over each residue^{53,85,86} to mitigate the limitations of partial charge schemes.⁸⁷ We note that one limitation of this study is that an implicit solvation model will not allow for direct solvent-protein hydrogen bonding interactions nor account for potential participation of water in reactions at the active site of the enzyme. Nevertheless, choosing to treat the solvent implicitly was necessary to limit the already high computational cost of the first-principles dynamics and avoids sampling challenges of needing to sample solvent degrees of freedom. In future work, use of machine learning interatomic potentials that predict partial charges could be used to allow for an ML/MM approach that accounts for explicit water molecules. Next, the charge covariance and mutual information (MI) were computed using sklearn and numpy with the nearest neighbors set to 10 for MI using an in-house script following previous work.⁵³ To calculate the energetics of oxo formation from hydrogen peroxide with and without the presence of Arg10', QM calculations were performed on

minimal cluster models consisting of the iron porphyrin, axial histidine ligand, hydrogen peroxide, three water molecules to facilitate proton transfer, and with and without Arg10'. Models with Arg10' included a truncated arginine sidechain (from delta carbon to guanidine group) to maintain computational tractability. All reaction mechanism QM calculations were performed using ORCA version 6.0.0 as TeraChem does not support analytical Hessians for the transition state calculations.⁸⁸ For the reaction mechanism, initial geometry optimizations were carried out at the ω PBEh-D4/LACVP* level of theory,⁸⁹ which corresponds to the LANL2DZ effective core potential applied to iron and 6-31G* applied to the other atoms, for the hydrogen peroxide (reactant) and oxo (product) bound states.⁹⁰⁻⁹² All QM calculations were performed in implicit solvent using C-PCM, as implemented in ORCA⁸⁸, using a cavity that was $1.2 \times$ Bondi radii with a dielectric constant (ϵ) of 4, chosen to approximate the electrostatic environment of the mimochrome helix-porphyrin interface.^{78,93} Transition-state analysis, including reaction-path optimization with nudged elastic band (NEB) calculations and energy refinement performed at the ω PBEh/def2-TZVPP and ω PBEh/def2-QZVPP levels of theory and extrapolated to the complete basis set (CBS) limit with Gibbs free energy corrections, were then performed (Supporting Information Text S4).

2d. Interpretable Machine Learning Models.

We formulated two distinct temporal feature sets: the first containing residue-summed Mulliken charges (23 features), and the second comprising all pairwise distances (351 features). The features corresponding to substituted amino acids (i.e., those that varied between mimics) were omitted from both sets as including the charges for the mutated residues would lead to trivial identification of each protein by the model. We next employed random forest (RF)⁹⁴ and multi-layer perceptron (MLP)⁹⁵ multi-class classifier models with a temporal dataset split,

allocating six independent simulations for training, one for validation, and one for testing. This approach was chosen over a random split to avoid data leakage of the temporally correlated MD data. The RF models were built and trained using scikit-learn v1.3.1.⁹⁶ Hyperparameter searching was performed using a grid search and 3-fold cross-validation (Supporting Information Table S7). The multi-class MLP classifier model was trained with PyTorch version 2.0,⁹⁷ with the ReLU activation function⁹⁸ applied between layers, and utilizing the Adam optimizer (Supporting Information Table S8).⁹⁹ Systematic hyperparameter searches, ROC-based evaluation, and feature-importance analyses were also employed (Supporting Information Text S5).

2e. Experimental Methods.

MC6*a-LysBoc was synthesized starting from MC6*a upon reaction with N-(tert-Butoxycarbonyloxy)succinimide (Boc-OSu). The reaction progress was monitored by analytical RP-HPLC, and the final product was purified by preparative RP-HPLC and its identity confirmed by ESI-MS analysis (Supporting Information Text S6 and Figures S7–S9). The peroxidase activity of MC6*a-LysBoc was evaluated in the oxidation of ABTS with H₂O₂. Kinetic studies were performed using a stopped-flow instrument in 50 mM phosphate buffer pH 6.5/TFE 1:1 (v/v), monitoring the formation of oxidized ABTS^{•+} at $\lambda = 660$ nm. To determine catalytic parameters, initial reaction rates (v_0) were measured while varying the concentration of one substrate and keeping the other fixed, and the experimental data were fitted to a two-substrate Michaelis–Menten equation. The total turnover number (TON) was also measured to evaluate catalyst robustness (Supporting Information Text S7 and Figure S10).

3. Results and Discussion.

3a. Differences in Classical Dynamics Across Closely Related Mimochromes.

To investigate the structural basis for the reactivity differences between MC6, MC6*, and MC6*a, we performed classical MD simulations for each analogue in the resting and oxo-bound states (see Sec. 2a). Our goal was to identify the factors that account for the progressively higher k_{cat} values and the improved robustness (and thus TON) across the series. We first evaluated global secondary structure and backbone motions. Analysis of the secondary structure using DSSP revealed comparable α -helical content for the D and TD chains across all three mimochromes, consistent with previous experimental findings (Supporting Information Figure S11).³⁷ We examined geometric backbone fluctuations by computing pairwise linear correlations between α -carbon ($C\alpha$) motions, which showed nearly indistinguishable patterns for MC6, MC6*, and MC6*a (Supporting Information Figure S12). These findings indicate that geometric $C\alpha$ motions are not sufficient to explain the observed differences in reactivity.

While global dynamics were similar across analogues, residue-wise root-mean-squared fluctuations (RMSF) highlighted localized differences in flexibility, particularly for the residue Arg10'. We found that across all the MD simulations for each mimochrome, the RMSF of Arg10' increases with the experimentally observed trend of k_{cat} increase, with average RMSF values of 3.7 Å for MC6, 4.5 Å for MC6*, and 4.8 Å for MC6*a (Figure 3 and Supporting Information Figure S13). Although the flexibility of Arg10' is enhanced in mimochromes displaying higher k_{cat} , we also observed instances of Arg10' accessing the active site in MC6 due to frequent loss of the Glu2...Arg10' salt bridge during the MD simulations (Supporting Information Figure S13). This occurs despite MC6 not containing the Glu2Leu substitutions. This observation is corroborated by a previous computational study that identified the Glu2...Arg10' interaction as the weakest ion pair in MC6.⁶⁸ While Arg10' displayed high RMSF values, the C-terminus of the

TD chain exhibited the highest degree of flexibility for all mimochromes, usually appearing disordered (Figure 3 and Supporting Information Figure S13).

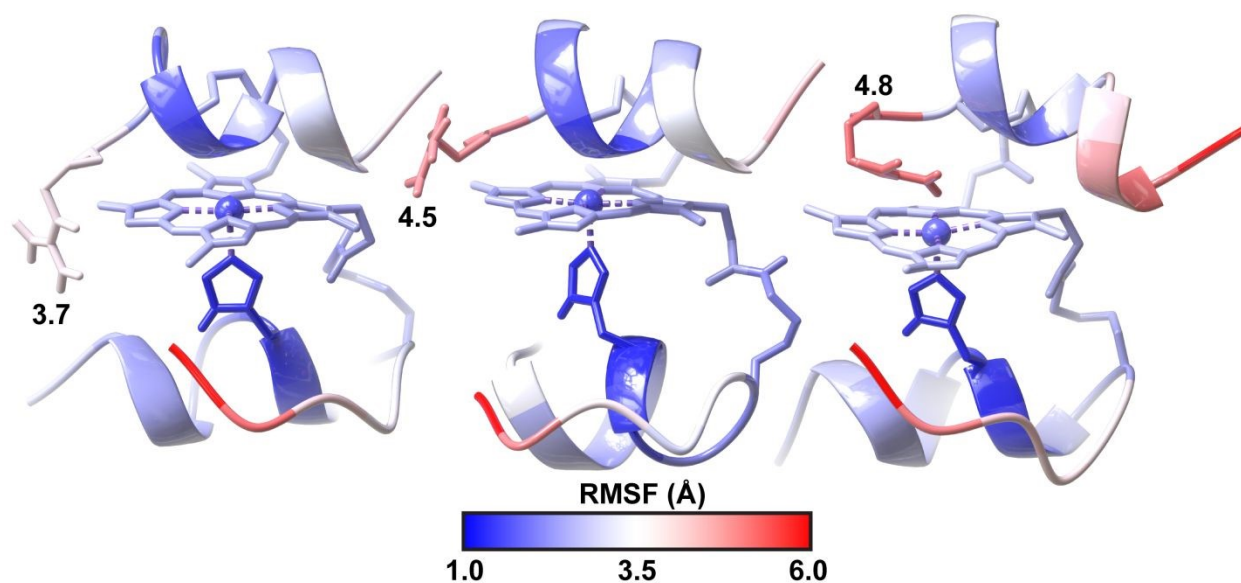


Figure 3. The structures of MC6 (left), MC6* (middle), and MC6*a (right) with their residue-wise root mean squared fluctuation (RMSF). The coloring is the RMSD averaged across all classical MD trajectories for each mimochrome but shown on a single representative structure. The RMSF of the hangman residue, Arg10', is labeled on the structures in angstroms (Å). The color bar ranges from a RMSF low of 1.0 Å to a high of 6.0 Å.

We next quantified the interactions of key residues with the porphyrin ring, motivated by prior NMR studies suggesting increased protein–heme contact in MC6*a.³⁷ We conducted an energy decomposition analysis using the generalized Born model (GBSA) on both the resting and oxo-bound trajectories.^{100,101} In the resting state, Arg10' exhibited the strongest interaction with the porphyrin ring for all three analogues, with average interaction strengths as favorable as -8.7 kcal/mol (Supporting Information Figures S14–S16). In all three cases, the second most strongly interacting residue was Leu5, which is adjacent to the coordinating histidine on the TD chain. These results indicate that Leu5 contributes to stabilizing the positioning of the TD chain relative to the porphyrin. For the oxo-bound state, GBSA analysis revealed a consistent interaction between Arg10' and the oxo with modest average interaction strengths (Supporting

Information Figures S17–S19). Hydrogen-bond analysis further indicated that Arg10'-oxo hydrogen bonds are sampled intermittently in the oxo-bound classical MD simulations, with occupancies of $14.8 \pm 2.9\%$, $19.8 \pm 4.1\%$, and $18.7 \pm 5.0\%$ for MC6, MC6*, and MC6*a, respectively (Supporting Information Table S9). While the total interaction energies are modest, their trend follows the experimentally observed series of k_{cat} (MC6 < MC6* < MC6*a). The interaction strengths result from a balance between strong, favorable electrostatic contributions (MC6: -9.8, MC6*: -9.6, and MC6*a: -16.4 kcal/mol) and large, unfavorable polar-solvation energies that arise from the solvent-exposed nature of the active site (Supporting Information Figures S17–S19).

Overall, the classical MD simulations reveal a high degree of structural similarity across the mimochrome series, as evidenced by comparable secondary structure content and collective backbone motions. Despite this overall similarity, both RMSF and GBSA analyses consistently highlight the central role of Arg10' in all three mimochromes, identifying it as a flexible residue with energetic contributions in both the resting and oxo-bound states. The flexibility of Arg10' subtly increases across the series (MC6 < MC6* < MC6*a), and GBSA analysis indicates a progressively more favorable interaction with the oxo intermediate (Figure 3 and Supporting Information Figures S17–S19). The enhanced flexibility of the distal Arg could enable dynamic stabilization of different transition states along the catalytic cycle, facilitating both peroxide bond cleavage and the subsequent reaction of the Fe–oxo species with the reducing substrate, leading to an overall increase in k_{cat} . Although these trends correlate with reactivity, the inability of other descriptions derived from classical trajectories such as C_α correlated motions and DSSP to account for the observed reactivity differences motivates investigation into the electrostatic properties of the system.

3b. Charge-Coupling Interactions Distinguish Mimochromes.

In our investigation of the geometric differences between MC6, MC6*, and MC6*a, we identified only limited variations, primarily the increased flexibility of Arg10' in MC6*a. We next focused on charge-coupling interactions between residue pairs to determine if electrostatic differences are more noticeable between the mimochromes analogues, as even small changes may produce strong electrostatic differences given their compact structure and multiple charged side chains. To quantify charge-coupling interactions, we utilized linear cross-correlation (CC)¹⁰² and mutual information (MI)^{55,103,104} analyses from information theory of the by-residue-summed Mulliken partial charges on each residue, as introduced previously for charge-coupling in proteins.^{53,54} Cross-correlation captures the linear dependence between charge fluctuations in residue pairs, while mutual information captures both linear and nonlinear relationships through the statistical dependence of charges between residues. We computed CC and MI scores for all DFT-computed snapshots on SQM trajectories (see Sec. 2). For each mimochrome, the reported charge-coupling distributions and CC/MI values were calculated from the concatenated charge data from all eight independent SQM trajectories. The resulting CC and MI heatmaps reveal distinct patterns of charge-coupling for the MC6, MC6*, and MC6*a mini-proteins (Figure 4). The charge-coupling analysis reveals a number of charge-coupling interactions unique to each mimochrome analogue, a strikingly different picture from the C_α correlation analysis for MC6, MC6*, and MC6*a that depicted highly similar patterns for all three mimochromes (Supporting Information Table S10 and Figure S12).



To understand which charge coupling interactions are most important in distinguishing the mimochromes, we first investigated the strongest interaction in MC6 and MC6*. In these cases, the single strongest shared interaction was observed between Ser6' and the porphyrin ring indicated by high CC scores of -0.44 and -0.52, respectively, whereas there is no evidence of charge coupling between Ser6' and the porphyrin in MC6*a, with a CC score of -0.06 (Figure 4 and Supporting Information Figure S20). Observation of the electronic interaction caused us to notice that in both the classical and *ab initio* MD of MC6 and MC6*, Ser6' is most often positioned directly above iron and is often found occupying the open coordination site (Supporting Information Figure S20). This interaction leads to very clear time-dependent

coupling where Ser loses charge (i.e., becomes significantly more positively charged), and the porphyrin accumulates charge (i.e., becomes significantly more negatively charged) upon coordination of the metal by Ser6' (Supporting Information Figure S21). While all three mimochromes share the Ser6' residue in common, the disruption of the Ser6'...porphyrin interaction in MC6*a is important for maintaining an open coordination site for hydrogen peroxide binding. In addition to the charge-coupling interaction between Ser6' and the porphyrin, we also see a charge coupling interaction between Ser6' and Arg10' for MC6 and MC6* with CC scores of -0.59 and -0.30, respectively. However, we do not observe charge coupling between Ser6' and Arg10' in MC6*a with a low cross-correlation score of -0.12 (Supporting Information Figures S22 and S23). The lack of charge coupling between Arg10' and Ser6' in the most reactive MC6*a (i.e. showing the highest k_{cat}) may be important in preventing Ser6' from competing with hydrogen peroxide for hydrogen bonding with Arg10', a residue thought to stabilize the transition state for oxo formation.³⁶

We next considered the intra-chain charge coupling interaction between Glu2' and Arg10' on the D chain, distinct from the inter-chain Glu2'...Arg10 salt bridge. While coupling in MC6 (CC = -0.11) and MC6* (CC = -0.09) was weak and comparable, MC6*a exhibited significantly increased charge coupling (CC = -0.43; Supporting Information Figure S24). A strong inverse correlation in MC6*a charge fluctuations between Glu2' and Arg10' indicates that the two residues are capable of interacting in MC6*a despite being at opposite ends of the D chain with their close interaction occurring above the porphyrin during portions of dynamics (Supporting Information Figure S25). The colocalization of arginine and carboxylate residues has been observed in many peroxidases and plays an important mechanistic role with arginine stabilizing

the O–O bond cleavage and the carboxylate acting to transfer a proton from the metal-proximal oxygen to the distal oxygen,¹⁰⁵⁻¹⁰⁸ suggesting a similar role in MC6*a.

We next analyze the two strongest charge-coupling interactions observed for MC6*a, which correspond to the inter-chain interactions between Glu2' and either Arg10 or Lys11 (Figure 4). Notably, the Glu2'...Arg10 interaction exhibits progressive increase in charge-coupling across the series, with CC values of -0.32, -0.58, and -0.76 for MC6, MC6*, and MC6*a, respectively (Supporting Information Figure S26). The increased coupling between Glu2' and Arg10 likely stabilizes the positioning of the D chain and the overall helix-porphyrin-helix fold through a salt bridge linking the distal (D) and proximal (TD) chains, in agreement with prior observations in the literature.^{31,35,36,68} This interaction is not only strong over the whole trajectory but can be attributed to a clear inverse correlation in the time-dependence of the charges on Glu2' and Arg10, consistent with a persistent electrostatic interaction due to the presence of a salt bridge (Supporting Information Figure S27). The second strongest charge coupling interaction for MC6*a was observed between Glu2' and Lys11, an interaction which has not been previously reported in the literature (Figure 5). Over all SQM trajectories, the CC between Glu2' and Lys11 was the strongest in MC6*a with a value of -0.50 compared to 0.00 in MC6 and 0.13 in MC6* (Figure 5). Much like the Glu2'...Arg10 pair, Glu2'...Lys11 also bridges the D and TD chains, suggesting a stabilizing salt-bridge function. Analysis of charge fluctuations over the MC6*a trajectory reveals sustained, highly anti-correlated charge coupling between Glu2' and Lys11, consistent with persistent electrostatic compensation indicative of a stabilizing salt bridge (Figure 6). Collectively, these results indicate that the increased reactivity of MC6*a is associated with strengthened electrostatic interactions involving Glu2'...Arg10, Glu2'...Lys11, and Glu2'...Arg10', and a concurrent reduction in coupling between Ser6' and

both the porphyrin and Arg10'. In particular, the inter-chain interactions likely contribute to the improved resistance of MC6*a compared to the previous analogues, providing an overall greater stabilization of the helix-heme-helix structure. Concurrently, the reduced coupling of Ser6' with the porphyrin and the intra-chain Glu2'...Arg10' interaction may favor H₂O₂ activation, thereby accounting for the increased k_{cat} of MC6*a. Finally, we note that in these SQM MD simulations we did not explicitly account for water molecules to reduce computational cost and avoid the need to sample solvent degrees of freedom (see Sec. 2c). Nevertheless, from classical MD we note that 3.67-4.93 water molecules come within 6.0 Å of the heme Fe during MD (Supporting Information Table S11). To test whether the implicit-solvent charge analysis was sensitive to nearby solvent electrostatics, we performed paired GFN2-xTB/COSMO calculations on approximately 100 evenly spaced resting-state MD snapshots from each replicate, with either a 10 Å shell of explicit MM water point charges around the heme or without explicit water point charges. The residue-summed Mulliken charges from the two protocols were strongly correlated for MC6, MC6*, and MC6*a ($r = 0.98, 0.96,$ and 0.95), indicating that nearby explicit water point charges do not qualitatively alter the residue-level charge distributions (Supporting Information Figure S28). While these specific interactions do not be strongly modified by the presence of explicit waters, this is unlikely to be the case for all possible charge-charge coupling interactions. Thus, future work could also account for explicit protein-solvent charge coupling, especially when aided by alternative approaches such as emerging machine learning interatomic potentials (MLIPs) that could directly predict partial charges. MLIPs can also address the timescale limitations of the present work, which may be particularly important for couplings that involve larger-scale rearrangements that cannot be captured at short timescales.

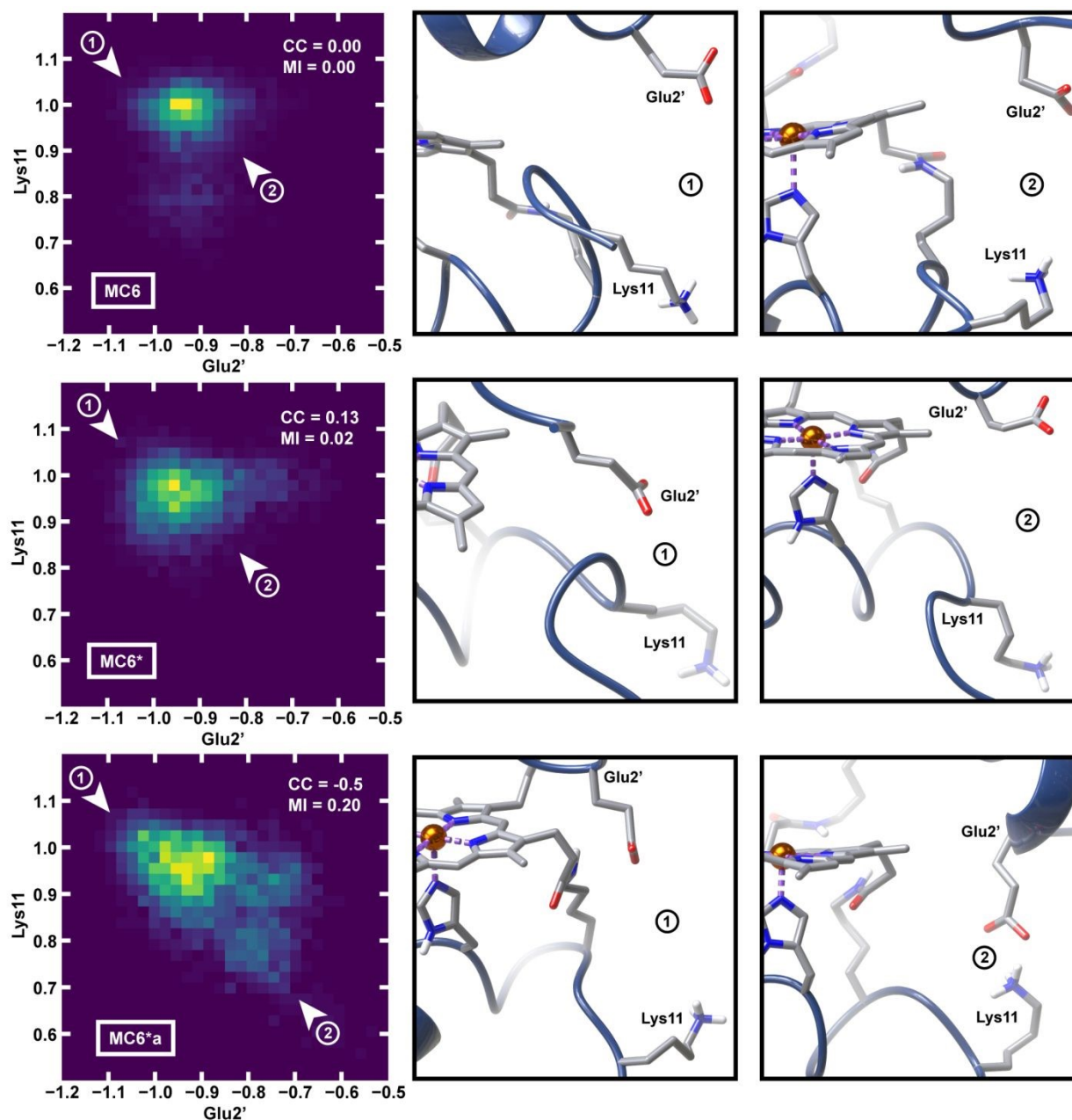


Figure 5. Two-dimensional distributions of charges showing the charge-coupling interaction between Lys11 and Glu2' with representative structures from the extremes of the distribution. (left) Charge coupling distribution for Lys11 and Glu2'. Distributions include all QM snapshots from all trajectories. (middle) Representative structure for Lys11 with a relatively positive charge. (right) Representative structure for Lys11 with a relatively negative charge. Charge coupling plots and structures are shown for (top) MC6, (middle) MC6*, and (bottom) MC6*a. The cross-correlation (CC) and mutual information (MI) scores are shown in inset. Regions with high populations are indicated with yellow and regions with low populations are indicated with purple in the distributions. Dark blue cartoon representations are used for the secondary structures with purple dashed lines for the coordinating bonds. The atoms are colored as follows: carbon in gray, nitrogen in blue, oxygen in red, hydrogen in white, and iron in brown.

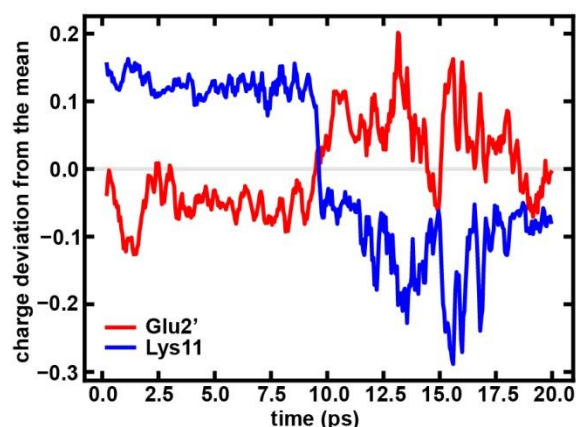


Figure 6. Time-dependent charge deviation from the mean for Lys11 (blue) and Glu2' (red) over time for MC6*a. The time-dependent charge deviation was calculated for 400 DFT-computed snapshots spaced 0.2 ps apart for MC6*a. The depicted snapshots were taken from SQM MD trajectory 7. The y-axis represents the deviation of the charge from its average value, while the x-axis indicates the elapsed time in picoseconds (ps).

3c. Experimental Validation of the Glu2'...Lys11 Interaction.

Based on the results of the MI and CC analysis of charge coupling interactions, we next experimentally investigated the role of the Lys11...Glu2' inter-chain salt bridge on reactivity of MC6*a, as it exhibited the strongest charge coupling signal that has not been described previously. For this purpose, we introduced a tert-butyloxycarbonyl (Boc) group to the side chain of Lys11, a modification that prevents salt bridge formation by neutralizing the ϵ -ammonium group through carbamate formation while introducing steric bulk but not altering the dihedral preferences of the amino acid.¹⁰⁹⁻¹¹¹ Further, the Boc group is removed in strongly acidic conditions, thus stable in the conditions required for product purification and catalytic screening. The derivative of MC6*a with Boc-protected Lys11 sidechain (referred to as MC6*a-LysBoc) was synthesized by reaction with N-(tert-Butoxycarbonyloxy)succinimide (Boc-OSu), which enables selective functionalization of primary amines (Supporting Information Text S6 and Figures S7–S9). To test whether the functionalization affected catalysis, the oxidation of ABTS in the presence of H₂O₂ was assayed.

Catalytic parameters were measured by stopped-flow technique, in the same experimental conditions as previously reported for MC6*a (Supporting Information Text S7).³⁷ Kinetic experiments were performed by keeping fixed the concentration of one substrate (ABTS or H₂O₂) while varying the concentration of the other substrate. The initial reaction rates (v_0) were then plotted against the concentration of the varying substrate, and both datasets were fitted to a 2-substrate Michaelis–Menten kinetic model (Figure 7 and Supporting Information Text S7). The kinetic parameters of MC6*a-LysBoc are reported, together with those of MC6*a as comparison (Table 1).³⁷

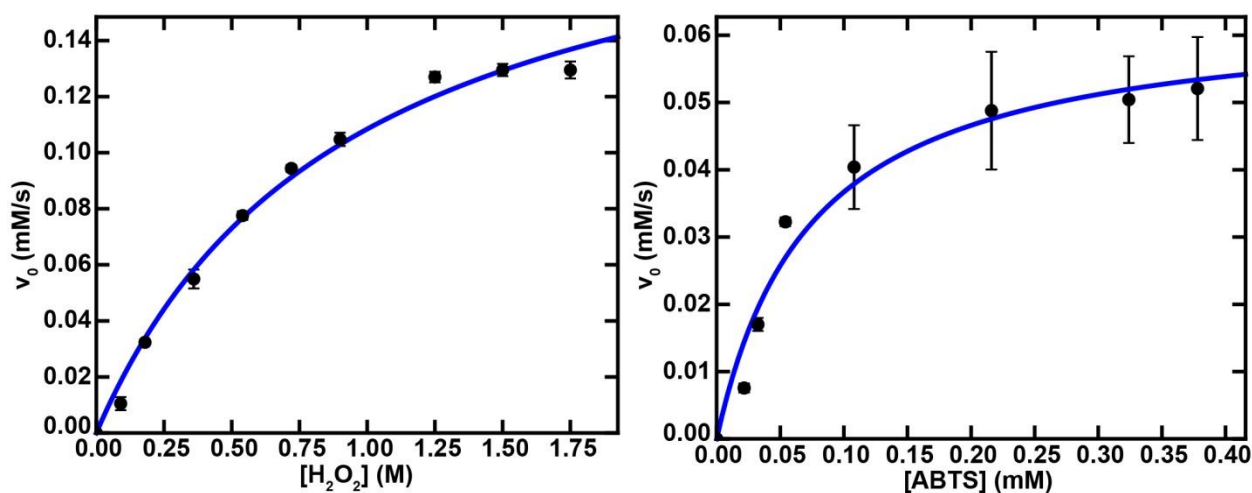


Figure 7. Peroxidase activity of MC6*a-LysBoc. (left) Initial rate dependence towards H₂O₂ concentration reaction conditions were MC6*a-LysBoc (20 nM), ABTS (2 mM) in 50 mM phosphate buffer pH 6.5/TFE (1:1 v/v). (right) Initial rate dependence towards ABTS concentration: reaction conditions were MC6*a-LysBoc (20 nM), H₂O₂ (450 mM) in 50 mM phosphate buffer pH 6.5/TFE (1:1 v/v).

Table 1. Catalytic parameters for ABTS oxidation, catalyzed by MC6*a-LysBOC. The MC6*a parameters, taken from ref. [29] are reported for comparison.

	K_m^{ABTS} (10^{-2} mM)	$K_m^{H_2O_2}$ (M)	k_{cat} (10^3 s ⁻¹)	k_{cat}/K_m^{ABTS} (10^3 mM ⁻¹ s ⁻¹)	$k_{cat}/K_m^{H_2O_2}$ (10^3 mM ⁻¹ s ⁻¹)	TON (10^3)
MC6*a	9 ± 1	0.44 ± 0.05	5.8 ± 0.3	64 ± 8	0.013 ± 0.002	14
MC6*a-LysBoc	19 ± 4	0.77 ± 0.08	8.6 ± 0.6	45 ± 10	0.011 ± 0.001	2

From the analysis of v_0 dependence on H_2O_2 concentration, we observe that much higher concentrations of hydrogen peroxide are required to reach saturation in MC6*a-LysBoc compared to MC6*a (Figure 7). Indeed, while MC6*a reached a plateau between 0.6 and 0.8 M of H_2O_2 , at least 1.4 M of peroxide are required for MC6*a-LysBoc to reach saturation. This finding results in an almost doubled K_m for H_2O_2 and 1.5-fold increased k_{cat} for MC6*a-LysBoc compared to MC6*a. Such increase in k_{cat} and K_m leads to an almost unchanged catalytic efficiency ($k_{\text{cat}} / K_m^{\text{H}_2\text{O}_2}$). Similarly, the K_m for the ABTS substrate is doubled compared to that observed for MC6*a (Figure 7). The increased K_m for both substrates suggests that a more solvent-exposed active site is present in MC6*a-LysBoc, which is hardly saturated with substrate molecules. This is especially true for H_2O_2 , given its reduced size compared to ABTS. Next, we measured the total turnover number (TON) of MC6*a-LysBoc, to evaluate the effect of the sidechain modification on catalyst robustness. Interestingly, the minimum catalyst concentration required for complete conversion of 70 μM ABTS was 35 nM, yielding to a TON of ~ 2000 (Supporting Information Figure S10). This value is much lower compared to that of MC6*a (14000)³⁷ in the same experimental conditions, highlighting that Lys11 modification negatively affects catalyst robustness. Taken together, these results indicate that the Lys11 side chain plays a key role in stabilizing the overall catalyst structure, an observation not recognized prior to our computational analysis. The observed decrease in TON, along with the increased K_m values, suggests that the active site becomes more exposed, which may contribute to oxidative deactivation of the catalyst during the reaction. Notably, Lys11 has been previously employed to immobilize MC6*a onto nanomaterials upon functionalization with a pegylated dibenzocyclooctyne (PEG₄-DBCO) as a “clickable” group.¹¹²⁻¹¹⁴ When MC6*a is immobilized on gold nanoparticles of various shapes, its catalytic efficiency decreases relative to the freely

diffusing enzyme, primarily due to reduced k_{cat} values.¹¹³ In these cases, the K_{m} for H_2O_2 is actually lowered upon immobilization. In contrast, MC6*a-LysBoc exhibits a markedly different behavior: both the k_{cat} and K_{m} increase compared to the unmodified enzyme. This suggests that the effects observed with nanomaterial immobilization arise from different underlying causes, likely related to steric hindrance and conformational constraints imposed by the nanoparticle surface. The effect observed with MC6*a-LysBoc is likely attributable to hindering of ability of the Lys11 sidechain to form a charge-assisted hydrogen bond.

3d. Charge-derived features distinguish mimochromes.

Motivated by our identification of significant differences in the charge coupling patterns between the mimochrome analogues and the experimental validation of the importance of the identified Glu2'...Lys11 charge coupling interaction in MC6*a reactivity, we sought to investigate whether these differences in charge distribution would be sufficient to distinguish the analogues. While we expect differences in charges for the three substituted residues that vary between MC6, MC6*, and MC6*a, we investigated if any of the shared residues also exhibit differences in charge distribution. The distinct electrostatic patterns revealed by the charge-coupling analysis suggested the potential to classify the mimochromes based solely on their charge distributions. To aid in the identification of qualitative distinguishing differences in charge distribution between the mimochromes, we trained supervised machine learning (ML) classifier models (i.e., random forest, RF, and multi-layer perceptron, MLP, see Sec. 2d) to distinguish patterns in charge differences among the mimochromes from their *ab initio* MD trajectories.¹¹⁵ The models were trained as multi-class classifiers to distinguish *ab initio* MD frames as belonging to MC6, MC6*, or MC6*a based on by-residue summed Mulliken partial charges of the shared residues using a random 60-20-20 train-validation-test split. These RF

models showed good performance with area under the curve (AUC) scores of 0.75–0.95 (Figure 8 and Supporting Information Table S12). The MLP model in particular showed AUC scores of 0.87–0.90 (Supporting Information Figures S29 and S30).

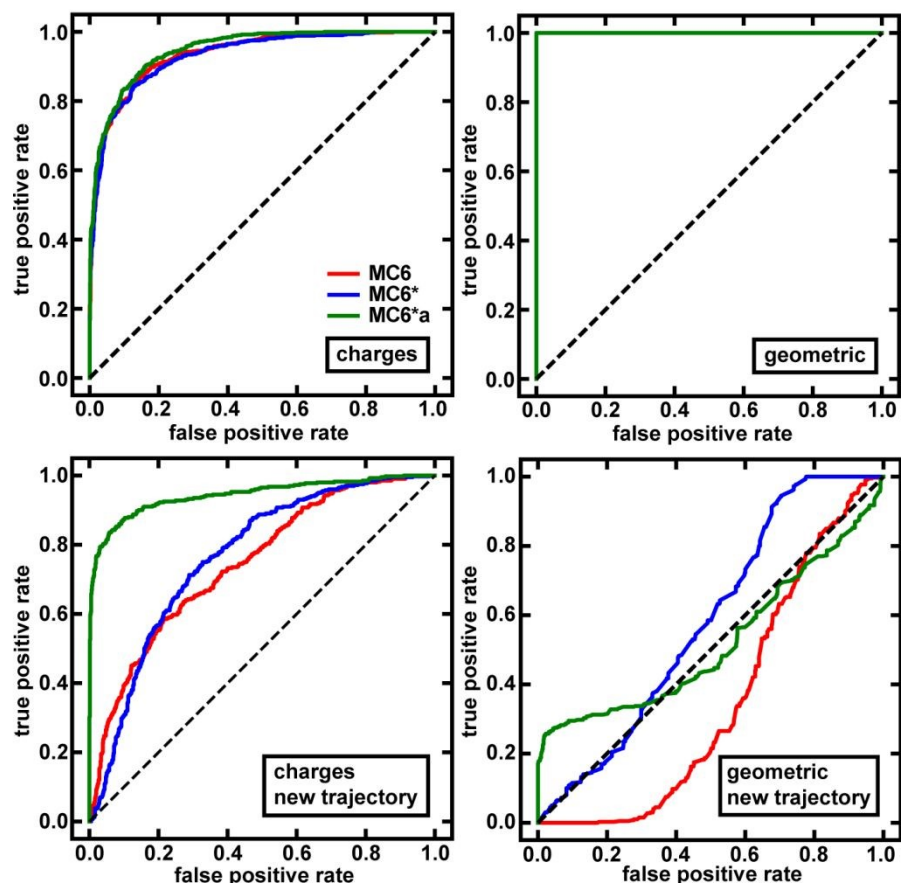


Figure 8. Receiver operating characteristic (ROC) curve for RF models for classification of individual snapshots as belonging to trajectories of one of three mimochromes. The models were trained on (top, left) by-residue summed charges of residues shared by the three mimochromes, (top, right) geometric pair-wise distance features, (bottom, left) by-residue summed charges on an unseen test trajectory, and (bottom, right) pairwise distances for all shared residues on an unseen test trajectory. ROC curves are shown for MC6 (red), MC6* (blue), and MC6*a (green). A diagonal dotted line represents the line of no discrimination.

Given the high computational cost of obtaining key charge-derived features, we repeated our investigation to ascertain if similar patterns are observable in the geometric characteristics of the mimochromes that could have been obtained from a purely classical MD trajectory. To test whether geometric features could distinguish MC6, MC6*, and MC6*a, we trained RF and MLP

models on centroid-to-centroid pairwise distances of all residues shared between the mimochromes from the same DFT-computed snapshots of the *ab initio* MD trajectories (351 features). We found that the models trained on geometric features were also able to almost perfectly classify the mimochromes (i.e., test set AUC scores of 1.0 for all mimochromes, Figure 8 and Supporting Information Figures S31–S32).

Given the temporal correlation in simulation trajectory data, we next tested if the models could generalize to unseen trajectories. We retrained the models with a new train-validation-test split where we used six trajectories for training, one trajectory for validation, and one trajectory for testing. With by-residue summed charges as features, we found that the models were still able to accurately classify the mimochromes for an unseen trajectory with AUC scores of 0.75–0.94 for the RF models and 0.71–0.88 for the MLP (Figure 8 and Supporting Information Figures S33–S34). However, the models trained on the geometric pairwise distances were unable to accurately classify the mimochromes with AUCs of 0.36–0.58 for the RF model and 0.40–0.60 for the MLP (Figure 8 and Supporting Information Figures S35–S36). The fact that the models trained on by-residue summed partial charges were able to generalize to not just randomly split test data but also to held-out, unseen trajectories with only 23 features highlights the importance of charge distribution in understanding mimochrome systems.

Given the strong classification accuracy on unseen trajectories for the models trained on by-residue summed partial charges, we examined the feature rankings to determine which specific residues drive the separation among mimochromes (Figure 9). This ranking of charge-based features contrasts with that of the geometric model, which performed poorly on unseen data and identified an entirely different set of important residue pairs and did not identify the Lys11...Glu2' interaction (Supporting Information Figure S37). From feature importance

analysis on the RF model, we found Arg10' to have the highest feature importance as measured by Gini impurity. The ranking of Arg10' as the residue with the highest feature importance is of interest given its involvement in the hangman effect and its suggested importance in the catalytic improvement between MC6 and MC6*. Lys11 had the second-highest feature importance score and is one of the charged amino acids on the TD chain and is involved in the interchain charge coupling interaction with Glu2' identified in the mutual information and cross correlation analysis (Figure 5). Glu2' had the fourth-highest feature importance score and is involved in prominent charge coupling interactions such as both Glu2'...Arg10 and Glu2'...Arg10' (Supporting Information Figures S20 and S22). While Gln4 had the third-highest feature importance score, we observed no notable charge coupling interactions and it is likely due to its proximity to the Glu2Leu substitution in the TD chain that differentiates MC6 and MC6*. Arg10', Lys11, and Glu2' rank among the highest-scoring features, consistent with their observed charge coupling interactions that distinguish the more reactive mimochromes. Ser6' has a similarly high score, yet its dominant coupling with Arg10' is inversely correlated with activity. Thus, this suggests Ser6' as a rational target for mutagenesis to disrupt this interaction in future mimochrome analogues. Additionally, residues with low feature importance such as Gln3, Ser7, and Gln8 could be considered for future mutagenesis as they were least important in distinguishing the mimochrome analogues. While the charge-coupling analysis suggests key interacting pairs, the feature importance analysis provides a quantitative ranking of the contribution of all shared residues, to the extent that they can be captured by relatively short xTB simulations. Furthermore, the successful classification of unseen trajectories confirms that the overall charge distribution and not pairwise geometric distances, serves as a robust and generalizable feature that distinguishes the analogues.

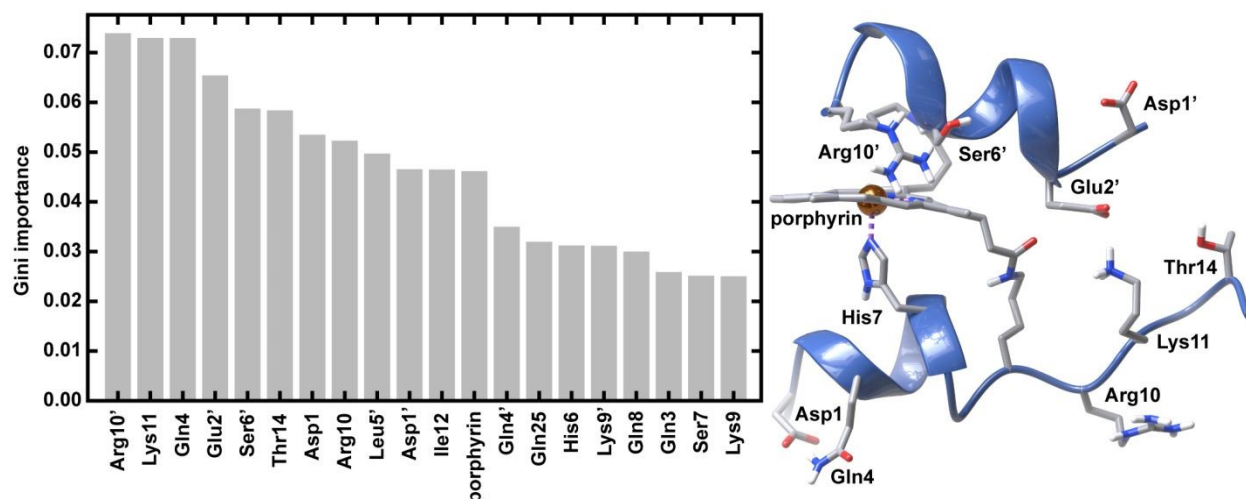


Figure 9. Ranked Gini feature importance for the RF model trained on charge-derived features. (left) The RF model was trained on by-residue summed charges, indicated with their three-letter code and residue index. (right) Representative structure of MC6*a with the residues with high feature importance scores shown in stick representation. Dark blue cartoon representations are used for the secondary structures and atoms are colored as follows: carbon in gray, nitrogen in blue, oxygen in red, hydrogen in white, and iron in brown. Metal coordinating bonds are shown as purple dashed lines.

3e. Reactivity and Mechanistic Insights.

We have identified Arg10' as the residue with the highest feature importance score in the ML models as well as playing a key role in a number of charge-coupling interactions that distinguish MC6*a from former analogues. We next sought to better understand the mechanistic role of Arg10' in the observed mimochrome peroxidase activity. We first performed QM calculations for oxo formation using a minimal model containing only the heme, its distal imidazole, three water molecules and H_2O_2 in the reactant state, with or without Arg10' (Supporting Information Figures S38 and S39). This minimal system was selected to isolate the Arg- H_2O_2 -porphyrin interaction, allowing us to assess a mechanistic feature that should be relevant to other porphyrin-based catalysts that generate high-valent oxo species from hydrogen peroxide. Using these models, we identified transition states for the formation of the reactive oxo intermediate from hydrogen peroxide in the presence and absence of Arg10'. Full

characterization of the transition states was initiated with NEB calculations to identify putative transition states at the ω PBEh/LACVP* level of theory (Supporting Information Figures S40 and S41). Transition states were then optimized, and single-point energy refinement was performed with ω PBEh/def2-TZVPP and ω PBEh/def2-QZVPP extrapolated to the complete basis set (CBS) limit with Gibbs free energy corrections from the lower ω PBEh/LACVP* level of theory (see Sec. 2c and Supporting Information Text S4).

The reaction was first modeled in the absence of Arg10' to isolate its specific contribution to catalysis. Optimization of the identified transition state revealed a barrier with a $\Delta G^\ddagger$ value of 25.2 kcal/mol, which corresponds to a significant reduction in the free energy barrier upon inclusion of entropic effects and extrapolation to the CBS limit (Figure 10 and Supporting Information Tables S13 and S14). The identified transition state corresponds to the expected heterolytic O–O cleavage following protonation of the distal oxygen of the ferric hydroperoxo intermediate (Supporting Information Figure S42). For the oxo-formation reaction with the inclusion of Arg10', we identified a similar transition state for the O–O cleavage of hydrogen peroxide (Supporting Information Figure S43). Full characterization of the transition state reveals a significant reduction in the barrier height following the inclusion of Arg10' with a $\Delta G^\ddagger$ value of 17.1 (Figure 10 and Supporting Information Tables S15 and S16). We also observed a small barrier (TS_0) for the proton transfer; however, it does not produce a free energy at both the lower and higher levels of theory. Overall, Arg10' lowers the free energy activation barrier for O–O bond cleavage from 25.2 kcal/mol to 17.1 kcal/mol through the stabilization of the proton transfer and polarization of the O–O bond. In contrast, the reaction without Arg10' relies entirely on the water molecules to stabilize the transition state, resulting in a significantly higher barrier.

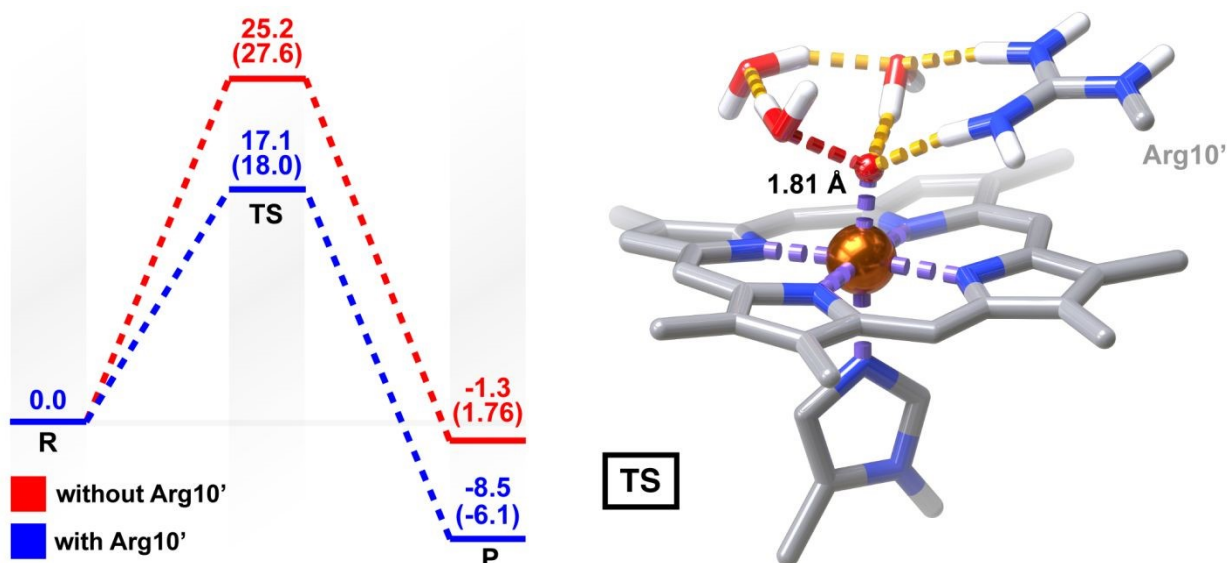


Figure 10. (left) Reaction energy diagram for generation of the oxo-bound intermediate from hydrogen peroxide with (red) and without (blue) the presence of Arg10'. (right) Structure of the transition state for the reaction with Arg10'. Energies were computed at the ω PBEh/def2-TZVPP and ω PBEh/def2-QZVPP levels of theory and extrapolated to the complete basis set (CBS) limit and are reported in kcal/mol above each reaction step, with free energies (ΔG) indicated above and electronic energies (ΔE) indicated below in parentheses. The bond breaking in the transition state is shown as a red dashed line. Atoms are colored as follows: carbon in gray, nitrogen in blue, oxygen in red, hydrogen in white, and iron in brown. Metal-coordinating bonds are shown as purple dashed lines, hydrogen bonds are shown as yellow dashes, and the dissociating O-O bond is shown in red.

The stabilization of the oxo-formation reaction by Arg10' highlights the critical role of electrostatic interactions in the catalytic mechanism. To further probe the collective influence of the protein scaffold's electrostatic environment, we calculated the electric field projected along the Fe-oxo bond, as its strength has been correlated with peroxidase activity in heme enzymes, using snapshots of the full protein from oxo-bound MD simulations (see Sec. 2a and Supporting Information Text S8).^{47,50} We used classical MD to generate the structures because we had not previously generated oxo-bound SQM trajectories, but we evaluate the electric field fully quantum mechanically from the partial charges of the entire system. That is, the projected electric field across the Fe-oxo bond was calculated for each structure from the sum of electric

field contributions from each computed atomic partial charge, atomic dipole, and atomic multipole moment in the extended protein environment (Supporting Information Text S8). The results reveal an increase in the average strength of the projected electric field along the Fe-oxo bond that correlates with the observed k_{cat} : MC6 (-75.1 MV/cm) < MC6* (-87.1 MV/cm) < MC6*a (-95.6 MV/cm) (Figure 11). In particular, a stronger electric field along the Fe-oxo bond reflects increased electrophilicity of the oxo moiety, which enhances Compound I reactivity and is consistent with the experimentally observed increase in k_{cat} across the MC6 series. Upon inspecting the structures with the strongest and weakest projected electric fields, a stronger electric field was observed when the D chain formed the expected α -helical fold and a weaker electric field when it was more disordered (Figure 11 and Supporting Information Figure S44). When the D chain was more structured (i.e., had greater helical content), Arg10' was found in close proximity to the oxo with the two negatively charged residues on the D chain, Asp1' and Glu2' (i.e., charges with strong coupling with a number of residues identified previously) positioned closer to the active site, leading to stronger calculated local electric fields. These findings indicate that the folding and rigidity of the D chain is crucial in affecting both the k_{cat} and robustness (i.e., TON) of mimochromes. Indeed, proper helical folding of the D chain enables the formation of charge coupling interactions that contribute to the global stabilization of the structure (such as Glu2'...Arg10 and Glu2'...Lys11), while also increasing the electric field strength at the active site. Additionally, differences in the macrodipole of the D-chain helix itself between the different mimochromes could contribute to the electric field projected along the Fe-oxo bond. This latter feature may contribute to the progressive increase of k_{cat} from MC6 to MC6*a, suggesting a potential avenue to further improve mimochrome reactivity in future generations.

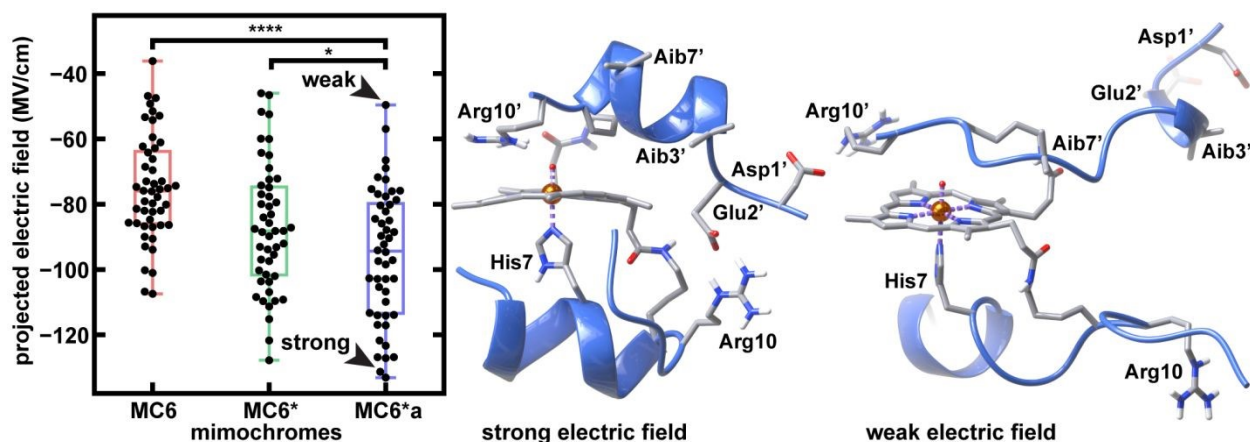


Figure 11. (left) Calculated projected electric fields projected across the Fe-oxo in MV/cm for MC6 (red), MC6* (green), and MC6*a (blue). Significance levels, calculated using two-tailed t-tests, are denoted by stars: **** for $p \leq 0.0001$ (MC6*a vs. MC6) and * for $0.01 < p \leq 0.05$ (MC6*a vs. MC6*). (middle) The MC6*a structure with the strongest electric field and (right) weakest electric fields are depicted. Dark blue cartoon representations are used for the secondary structures and atoms are colored as follows: carbon in gray, nitrogen in blue, oxygen in red, hydrogen in white, and iron in brown. Coordinating bonds are shown as purple dashed lines.

4. Conclusions.

Our integrated computational and experimental analysis of the mimochrome series (MC6, MC6*, and MC6*a) reveals that dynamic electronic properties, rather than global structural dynamics, dictate their different catalytic properties, particularly k_{cat} and TON. While classical MD simulations showed nearly indistinguishable backbone motions across the analogues, *ab initio* MD simulations and data-driven analyses uncovered distinct charge-coupling networks that correlate with experimentally observed trends in peroxidase reactivity. We identified key charge-coupling interactions, such as the inter- and intra-chain interactions Arg10'...Glu2' and Glu2'...Arg10, respectively, and the previously unreported Glu2'...Lys11 salt bridge, that distinguish the fastest (highest k_{cat}) and most robust (highest TON) mimochrome. The functional importance of this interaction was confirmed experimentally. The disruption of the Glu2'...Lys11 interaction in our newly reported MC6*a-LysBoc derivative resulted in a seven-fold decrease in TON, underscoring its role in structural stability. Supervised machine learning

models trained on the partial charge dynamics obtained from the *ab initio* MD simulations successfully classified the mimochromes on unseen trajectories and identified Arg10', Lys11, Glu2', and Ser6' as the most important distinguishing residues.

QM calculations confirmed the mechanistic importance of Arg10', showing it lowers the free energy barrier of the reaction from 25.2 kcal/mol to 17.1 kcal/mol. This electrostatic stabilization is reflected across the full protein, as the calculated electric field projected along the Fe–oxo bond strengthens, in line with the improved reactivity of Compound I. This work demonstrates that peripheral substitutions, such as the Aib residues in MC6*a, can reshape charge-coupling networks to enhance catalysis by strengthening productive long-range interactions while diminishing unproductive ones, such as the Ser6' interaction that blocks the active site in less active analogues. These findings establish a key principle for designing artificial enzymes: reactivity can be tuned through indirect modulation of dynamic charge-coupling networks. This represents a complementary tool to more established methods for understanding static electric fields and long-range, geometric correlated motions. Future studies, perhaps aided by rapidly maturing machine learning foundation models, should build upon the dynamic observations to carry out studies of catalytic mechanisms with larger QM models that incorporate catalytically relevant residues such as Glu2 and Lys11.

Focusing on mutations that could further enhance reactivity beyond the frontier MC6*a based on our overall findings, we note that such mutations would have to be validated experimentally in future work. Subsequent mimochrome generations could be improved by targeting residues identified as important but detrimental, such as Ser6', could be mutated to an Ala to avoid deleterious interactions. We separately propose that single mutations of residues that are non-critical for distinguishing reactivity, such as Gln3, Ser7, and Gln8, to positively

charged residues (i.e., Lys or Arg) could further optimize the electrostatic environment. Interactions between the heme and the D chain could be strengthened by mutating Glu2' to Asp to strengthen the interactions of Arg10' with the heme. Alternatively, a stronger salt bridge could be obtained by mutating Lys11 to an Arg. Nevertheless, we note that these computational predictions of mutants could have unforeseen consequences on protein secondary structure and would require further validation.

The general approach of exploring charge coupling networks through ab initio MD or with emerging neural network potentials that predict charge is anticipated to be a broadly useful tool for understanding catalytic systems that is complementary to more established techniques such as evaluating static electric fields or long-range, correlated motions. The analysis of charge-coupling networks is a likely useful strategy for understanding and engineering catalytic systems because rationally-tuned charge dynamics can produce significant functional effects with minimal structural changes.

AUTHOR INFORMATION

Corresponding Author

*email:hjkulik@mit.edu

Notes

The authors declare no competing financial interest.

Data Availability Statement.

All data in this manuscript is available in the Supporting Information or on Zenodo at the DOI: 10.5281/zenodo.16934208.

Supporting Information.

Additional details for the MC6, MC6*, and MC6*a mimochrome systems, including residue numbering convention, protonation states, and parameters; clustering statistics for MD trajectories and centroids; DSSP, RMSF, water shell, GBSA analyses; electrostatic potential residue correlation plots; random forest and multi-layer perceptron model hyperparameters; model performance and feature importance scores; mutual information heatmaps; charge coupling heatmaps and time-dependent charge deviation plots; oxo-generation reaction details and energetics; electric field calculation details and structural comparisons (PDF).

ACKNOWLEDGMENT

The electric field component of this work was supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing, Office of Basic Energy Sciences, via the Scientific Discovery through Advanced Computing (SciDAC) program (to D.W.K., C.R., and H.J.K.) and the machine learning work was supported by the National Science Foundation under grant number CBET-1846426 (to H.A. and H.J.K.). Finally, mechanistic investigations were supported by the National Institutes of Health under NIGMS award R35GM152027 (to D.W.K., C.R., and H.J.K.). Partial support was provided by National Science Foundation Graduate Research Fellowships under Grant #1745302 for D.W.K. and an Arnold O. Beckman Postdoctoral fellowship for C.R. Y.R.-L. acknowledges support from the US Department of Energy, Office of Basic Energy Sciences, under award no. Part of the work was supported by the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.3, theme 2.a “Green Energies for the Future”, funded by the European Union – NextGenerationEU – Project Title “NEST – Network 4 Energy Sustainable Transition” – CUP E63C22002160007

(to L.L., F.N. and A.L.). A.E. is grateful to the European Research Council (ERC) for being granted a research fellowship in the frame of the ERC-StG_851356_GLYCOSWITCH project – CUP E68D190017100006. This work was also supported by the U.S. Department of Energy under grant number DE-SC0016214 (to H.A. and Y.R.-L.). The authors thank Ilia Kevlishvili and Adam H. Steeves for providing a critical reading of the manuscript.

REFERENCES

- (1) Winkler, C. K.; Schrittwieser, J. H.; Kroutil, W. Power of Biocatalysis for Organic Synthesis. *ACS Central Science* **2021**, *7*, 55-71.
- (2) Kiss, G.; Çelebi-Ölçüm, N.; Moretti, R.; Baker, D.; Houk, K. N. Computational Enzyme Design. *Angewandte Chemie International Edition* **2013**, *52*, 5700-5725.
- (3) Huang, P.-S.; Boyken, S. E.; Baker, D. The Coming of Age of De Novo Protein Design. *Nature* **2016**, *537*, 320-327.
- (4) Miller, D. C.; Athavale, S. V.; Arnold, F. H. Combining Chemistry and Protein Engineering for New-to-Nature Biocatalysis. *Nature Synthesis* **2022**, *1*, 18-23.
- (5) Hassan, I. S.; Fuller, J. T.; Dippon, Vanessa N.; Ta, A. N.; Danneman, M. W.; McNaughton, B. R.; Alexandrova, A. N.; Rovis, T. Tuning through-Space Interactions *Via* the Secondary Coordination Sphere of an Artificial Metalloenzyme Leads to Enhanced Rh(III)-Catalysis. *Chemical Science* **2022**, *13*, 9220-9224.
- (6) Chaturvedi, S. S.; Bím, D.; Christov, C. Z.; Alexandrova, A. N. From Random to Rational: Improving Enzyme Design through Electric Fields, Second Coordination Sphere Interactions, and Conformational Dynamics. *Chemical Science* **2023**, *14*, 10997-11011.
- (7) Kastner, D. W.; Nandy, A.; Mehmood, R.; Kulik, H. J. Mechanistic Insights into Substrate Positioning That Distinguish Non-Heme Fe(II)/ α -Ketoglutarate-Dependent Halogenases and Hydroxylases. *ACS Catalysis* **2023**, *13*, 2489-2501.
- (8) Lu, Y.; Berry, S. M.; Pfister, T. D. Engineering Novel Metalloproteins: Design of Metal-Binding Sites into Native Protein Scaffolds. *Chemical Reviews* **2001**, *101*, 3047-3080.
- (9) Ortiz de Montellano, P. R. Hydrocarbon Hydroxylation by Cytochrome P450 Enzymes. *Chemical Reviews* **2010**, *110*, 932-948.
- (10) Derat, E.; Kumar, D.; Hirao, H.; Shaik, S. Gauging the Relative Oxidative Powers of Compound I, Ferric-Hydroperoxide, and the Ferric-Hydrogen Peroxide Species of Cytochrome P450 toward C-H Hydroxylation of a Radical Clock Substrate. *Journal of the American Chemical Society* **2006**, *128*, 473-484.

- (11) Vaillancourt, F. H.; Yeh, E.; Vosburg, D. A.; Garneau-Tsodikova, S.; Walsh, C. T. Nature's Inventory of Halogenation Catalysts: Oxidative Strategies Predominate. *Chemical Reviews* **2006**, *106*, 3364-3378.
- (12) Bhandari, Y.; Sajwan, H.; Pandita, P.; Koteswara Rao, V. Chloroperoxidase Applications in Chemical Synthesis of Industrial Relevance. *Biocatalysis and Biotransformation* **2023**, *41*, 403-420.
- (13) Shaik, S.; de Visser, S. P.; Ogliaro, F.; Schwarz, H.; Schröder, D. Two-State Reactivity Mechanisms of Hydroxylation and Epoxidation by Cytochrome P-450 Revealed by Theory. *Current Opinion in Chemical Biology* **2002**, *6*, 556-567.
- (14) Huang, Y.; Sha, J.; Zhang, J.; Zhang, W. Challenges and Perspectives in Using Unspecific Peroxygenases for Organic Synthesis. *Frontiers in Catalysis* **2024**, *Volume 4 - 2024*.
- (15) Song, Y.; Wang, Y.; Pan, B.; Cheng, X.; Yang, J.; Chen, Y. Recent Advances in Biocatalytic Preparation of Chiral Sulfoxides. *Asian Journal of Organic Chemistry* **2025**, *14*, e202400708.
- (16) Meunier, B.; de Visser, S. P.; Shaik, S. Mechanism of Oxidation Reactions Catalyzed by Cytochrome P450 Enzymes. *Chemical Reviews* **2004**, *104*, 3947-3980.
- (17) Guengerich, F. P.; Yoshimoto, F. K. Formation and Cleavage of C-C Bonds by Enzymatic Oxidation-Reduction Reactions. *Chemical Reviews* **2018**, *118*, 6573-6655.
- (18) Arnold, F. H. Directed Evolution: Bringing New Chemistry to Life. *Angewandte Chemie International Edition* **2018**, *57*, 4143-4148.
- (19) Privett, H. K.; Kiss, G.; Lee, T. M.; Blomberg, R.; Chica, R. A.; Thomas, L. M.; Hilvert, D.; Houk, K. N.; Mayo, S. L. Iterative Approach to Computational Enzyme Design. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109*, 3790-3795.
- (20) Jeschek, M.; Reuter, R.; Heinisch, T.; Trindler, C.; Klehr, J.; Panke, S.; Ward, T. R. Directed Evolution of Artificial Metalloenzymes for in Vivo Metathesis. *Nature* **2016**, *537*, 661-665.
- (21) Kan, S. B. J.; Lewis, R. D.; Chen, K.; Arnold, F. H. Directed Evolution of Cytochrome C for Carbon-Silicon Bond Formation: Bringing Silicon to Life. *Science* **2016**, *354*, 1048-1051.
- (22) Ma, N.; Fang, W.; Liu, C.; Qin, X.; Wang, X.; Jin, L.; Wang, B.; Cong, Z. Switching an Artificial P450 Peroxygenase into Peroxidase Via Mechanism-Guided Protein Engineering. *ACS Catalysis* **2021**, *11*, 8449-8455.
- (23) Hou, K.; Huang, W.; Qi, M.; Tugwell, T. H.; Alturaifi, T. M.; Chen, Y.; Zhang, X.; Lu, L.; Mann, S. I.; Liu, P.; Yang, Y.; DeGrado, W. F. De Novo Design of Porphyrin-Containing Proteins as Efficient and Stereoselective Catalysts. *Science* **2025**, *388*, 665-670.
- (24) Stenner, R.; Steventon, J. W.; Seddon, A.; Anderson, J. L. R. A De Novo Peroxidase Is Also a Promiscuous yet Stereoselective Carbene Transferase. *Proceedings of the National Academy of Sciences* **2020**, *117*, 1419-1428.

- (25) Chalkley, M. J.; Mann, S. I.; DeGrado, W. F. De Novo Metalloprotein Design. *Nature Reviews Chemistry* **2022**, *6*, 31-50.
- (26) Rosenblatt, M. M.; Wang, J.; Suslick, K. S. De Novo Designed Cyclic-Peptide Heme Complexes. *Proceedings of the National Academy of Sciences of the United States of America* **2003**, *100*, 13140-13145.
- (27) Polizzi, N. F.; Wu, Y.; Lemmin, T.; Maxwell, A. M.; Zhang, S.-Q.; Rawson, J.; Beratan, D. N.; Therien, M. J.; DeGrado, W. F. De Novo Design of a Hyperstable Non-Natural Protein–Ligand Complex with Sub-Å Accuracy. *Nature Chemistry* **2017**, *9*, 1157-1164.
- (28) Kalvet, I.; Ortmayer, M.; Zhao, J.; Crawshaw, R.; Ennist, N. M.; Levy, C.; Roy, A.; Green, A. P.; Baker, D. Design of Heme Enzymes with a Tunable Substrate Binding Pocket Adjacent to an Open Metal Coordination Site. *Journal of the American Chemical Society* **2023**, *145*, 14307-14315.
- (29) Polanco, E. A.; Opdam, L. V.; Passerini, L.; Huber, M.; Bonnet, S.; Pandit, A. An Artificial Metalloenzyme That Can Oxidize Water Photocatalytically: Design, Synthesis, and Characterization. *Chemical Science* **2024**, *15*, 3596-3609.
- (30) Nastri, F.; D'Alonzo, D.; Leone, L.; Zambrano, G.; Pavone, V.; Lombardi, A. Engineering Metalloprotein Functions in Designed and Native Scaffolds. *Trends in Biochemical Sciences* **2019**, *44*, 1022-1040.
- (31) Leone, L.; Chino, M.; Nastri, F.; Maglio, O.; Pavone, V.; Lombardi, A. Mimochrome, a Metalloporphyrin-Based Catalytic Swiss Knife†. *Biotechnology and Applied Biochemistry* **2020**, *67*, 495-515.
- (32) Leone, L.; D'Alonzo, D.; Maglio, O.; Pavone, V.; Nastri, F.; Lombardi, A. Highly Selective Indole Oxidation Catalyzed by a Mn-Containing Artificial Mini-Enzyme. *ACS Catalysis* **2021**, *11*, 9407-9417.
- (33) Leone, L.; Muñoz-García, A. B.; D'Alonzo, D.; Pavone, V.; Nastri, F.; Lombardi, A. Peptide-Based Metalloporphyrin Catalysts: Unveiling the Role of the Metal Ion in Indole Oxidation. *Journal of Inorganic Biochemistry* **2023**, *246*, 112298.
- (34) Salamatian, A. A.; Alvarez-Hernandez, J. L.; Ramesh, K. B.; Leone, L.; Lombardi, A.; Bren, K. L. Electrocatalytic Co₂ Reduction by a Cobalt Porphyrin Mini-Enzyme. *Chemical Science* **2025**, *16*, 5707-5716.
- (35) Nastri, F.; Lista, L.; Ringhieri, P.; Vitale, R.; Faiella, M.; Andreozzi, C.; Travascio, P.; Maglio, O.; Lombardi, A.; Pavone, V. A Heme–Peptide Metalloenzyme Mimetic with Natural Peroxidase-Like Activity. *Chemistry – A European Journal* **2011**, *17*, 4444-4453.
- (36) Vitale, R.; Lista, L.; Cerrone, C.; Caserta, G.; Chino, M.; Maglio, O.; Nastri, F.; Pavone, V.; Lombardi, A. An Artificial Heme-Enzyme with Enhanced Catalytic Activity: Evolution, Functional Screening and Structural Characterization. *Organic & Biomolecular Chemistry* **2015**, *13*, 4859-4868.
- (37) Caserta, G.; Chino, M.; Firpo, V.; Zambrano, G.; Leone, L.; D'Alonzo, D.; Nastri, F.; Maglio, O.; Pavone, V.; Lombardi, A. Enhancement of Peroxidase Activity in Artificial Mimochrome Vi Catalysts through Rational Design. *ChemBioChem* **2018**, *19*, 1823-1826.

- (38) Chino, M.; Leone, L.; Zambrano, G.; Pirro, F.; D'Alonzo, D.; Firpo, V.; Aref, D.; Lista, L.; Maglio, O.; Natri, F.; Lombardi, A. Oxidation Catalysis by Iron and Manganese Porphyrins within Enzyme-Like Cages. *Biopolymers* **2018**, *109*, e23107.
- (39) Lee, C. H.; Dogutan, D. K.; Nocera, D. G. Hydrogen Generation by Hangman Metalloporphyrins. *Journal of the American Chemical Society* **2011**, *133*, 8775-8777.
- (40) Bediako, D. K.; Solis, B. H.; Dogutan, D. K.; Roubelakis, M. M.; Maher, A. G.; Lee, C. H.; Chambers, M. B.; Hammes-Schiffer, S.; Nocera, D. G. Role of Pendant Proton Relays and Proton-Coupled Electron Transfer on the Hydrogen Evolution Reaction by Nickel Hangman Porphyrins. *Proceedings of the National Academy of Sciences of the United States of America* **2014**, *111*, 15001-15006.
- (41) Berglund, G. I.; Carlsson, G. H.; Smith, A. T.; Szöke, H.; Henriksen, A.; Hajdu, J. The Catalytic Pathway of Horseradish Peroxidase at High Resolution. *Nature* **2002**, *417*, 463-468.
- (42) Rodríguez-López, J. N.; Lowe, D. J.; Hernández-Ruiz, J.; Hiner, A. N. P.; García-Cánovas, F.; Thorneley, R. N. F. Mechanism of Reaction of Hydrogen Peroxide with Horseradish Peroxidase: Identification of Intermediates in the Catalytic Cycle. *Journal of the American Chemical Society* **2001**, *123*, 11838-11847.
- (43) Rodriguez-Lopez, J. N.; Smith, A. T.; Thorneley, R. N. F. Role of Arginine 38 in Horseradish Peroxidase: A Critical Residue for Substrate Binding and Catalysis. *Journal of Biological Chemistry* **1996**, *271*, 4023-4030.
- (44) Rodriguez-Lopez, J. N.; Smith, A. T.; Thorneley, R. N. F. Effect of Distal Cavity Mutations on the Binding and Activation of Oxygen by Ferrous Horseradish Peroxidase. *Journal of Biological Chemistry* **1997**, *272*, 389-395.
- (45) Veitch, N. C. Horseradish Peroxidase: A Modern View of a Classic Enzyme. *Phytochemistry* **2004**, *65*, 249-259.
- (46) Hiner, A. N. P.; Raven, E. L.; Thorneley, R. N. F.; García-Cánovas, F.; Rodríguez-López, J. N. Mechanisms of Compound I Formation in Heme Peroxidases. *Journal of Inorganic Biochemistry* **2002**, *91*, 27-34.
- (47) Bím, D.; Alexandrova, A. N. Local Electric Fields as a Natural Switch of Heme-Iron Protein Reactivity. *ACS Catalysis* **2021**, *11*, 6534-6546.
- (48) Shaik, S.; de Visser, S. P.; Kumar, D. External Electric Field Will Control the Selectivity of Enzymatic-Like Bond Activations. *Journal of the American Chemical Society* **2004**, *126*, 11746-11749.
- (49) Shaik, S.; Mandal, D.; Ramanan, R. Oriented Electric Fields as Future Smart Reagents in Chemistry. *Nature Chemistry* **2016**, *8*, 1091-1098.
- (50) Peng, W.; Yan, S.; Zhang, X.; Liao, L.; Zhang, J.; Shaik, S.; Wang, B. How Do Preorganized Electric Fields Function in Catalytic Cycles? The Case of the Enzyme Tyrosine Hydroxylase. *Journal of the American Chemical Society* **2022**, *144*, 20484-20494.
- (51) Song, Y.; Mao, J.; Gunner, M. R. Electrostatic Environment of Hemes in Proteins: Pk_as of Hydroxyl Ligands. *Biochemistry* **2006**, *45*, 7949-7958.

- (52) Kulik, H. J. Large-Scale Qm/Mm Free Energy Simulations of Enzyme Catalysis Reveal the Influence of Charge Transfer. *Physical Chemistry Chemical Physics* **2018**, *20*, 20650-20660.
- (53) Yang, Z.; Hajlasz, N.; Steeves, A. H.; Kulik, H. J. Quantifying the Long-Range Coupling of Electronic Properties in Proteins with Ab Initio Molecular Dynamics. *Chemistry–Methods* **2021**, *1*, 362-373.
- (54) Steeves, A. H.; Kulik, H. J. Insights into the Stability of Engineered Mini-Proteins from Their Dynamic Electronic Properties. *Electronic Structure* **2022**, *4*, 034005.
- (55) Cortina, G. A.; Kasson, P. M. Excess Positional Mutual Information Predicts Both Local and Allosteric Mutations Affecting Beta Lactamase Drug Resistance. *Bioinformatics* **2016**, *32*, 3420-3427.
- (56) Palacios, P. M.; Li, X.; Jaber Sathik Rifayee, S. B.; Tang, H.; Karabencheva-Christova, T.; Christov, C.; Chang, W.-c.; Guo, Y. Regio- and Stereoselective Halogenation by an Iron(Ii)- and 2-Oxoglutarate-Dependent Halogenase in the Biosynthesis of Halogenated Nucleosides. *Journal of the American Chemical Society* **2025**, *147*, 48166-48179.
- (57) Thomas, M. G.; Jaber Sathik Rifayee, S. B.; Christov, C. Z. How Do Variants of Residues in the First Coordination Sphere, Second Coordination Sphere, and Remote Areas Influence the Catalytic Mechanism of Non-Heme Fe(Ii)/2-Oxoglutarate Dependent Ethylene-Forming Enzyme? *ACS Catalysis* **2024**, *14*, 18550-18569.
- (58) Waheed, S. O.; Ramanan, R.; Chaturvedi, S. S.; Lehnert, N.; Schofield, C. J.; Christov, C. Z.; Karabencheva-Christova, T. G. Role of Structural Dynamics in Selectivity and Mechanism of Non-Heme Fe(Ii) and 2-Oxoglutarate-Dependent Oxygenases Involved in DNA Repair. *ACS Central Science* **2020**, *6*, 795-814.
- (59) Mehmood, R.; Qi, H. W.; Steeves, A. H.; Kulik, H. J. The Protein's Role in Substrate Positioning and Reactivity for Biosynthetic Enzyme Complexes: The Case of Syrb2/Syrb1. *ACS Catalysis* **2019**, *9*, 4930-4943.
- (60) Matthews, M. L.; Neumann, C. S.; Miles, L. A.; Grove, T. L.; Booker, S. J.; Krebs, C.; Walsh, C. T.; Bollinger, J. M. Substrate Positioning Controls the Partition between Halogenation and Hydroxylation in the Aliphatic Halogenase, Syrb2. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 17723-17728.
- (61) Martinie, R. J.; Livada, J.; Chang, W.-c.; Green, M. T.; Krebs, C.; Bollinger, J. M.; Silakov, A. Experimental Correlation of Substrate Position with Reaction Outcome in the Aliphatic Halogenase, Syrb2. *J. Am. Chem. Soc.* **2015**, *137*, 6912-6919.
- (62) Kastner, D. W.; Nandy, A.; Mehmood, R.; Kulik, H. J. Mechanistic Insights into Substrate Positioning across Non-Heme Fe(Ii)/ α -Ketoglutarate-Dependent Halogenases and Hydroxylases. *ACS Catalysis* **2023**, *13*, 2489-2501.
- (63) Nandy, A.; Adamji, H.; Kastner, D. W.; Vennelakanti, V.; Nazemi, A.; Liu, M.; Kulik, H. J. Using Computational Chemistry to Reveal Nature's Blueprints for Single-Site Catalysis of C–H Activation. *ACS Catalysis* **2022**, *12*, 9281-9306.

- (64) Jurich, C.; Yang, Z. J. High-Throughput Computational Investigation of Protein Electrostatics and Cavity for Sam-Dependent Methyltransferases. *Protein Science* **2023**, *32*, e4690.
- (65) Eberhart, M. E.; Wilson, T. R.; Johnston, N. W.; Alexandrova, A. N. Geometry of Charge Density as a Reporter on the Role of the Protein Scaffold in Enzymatic Catalysis: Electrostatic Preorganization and Beyond. *Journal of Chemical Theory and Computation* **2023**, *19*, 694-704.
- (66) Yan, S.; Ji, X.; Peng, W.; Wang, B. Evaluating the Transition State Stabilization/Destabilization Effects of the Electric Fields from Scaffold Residues by a Qm/Mm Approach. *Journal of Physical Chemistry B* **2023**, *127*, 4245-4253.
- (67) Nazemi, A.; Steeves, A. H.; Kastner, D. W.; Kulik, H. J. Influence of the Greater Protein Environment on the Electrostatic Potential in Metalloenzyme Active Sites: The Case of Formate Dehydrogenase. *Journal of Physical Chemistry B* **2022**, *126*, 4069-4079.
- (68) Perrella, F.; Raucci, U.; Chiariello, M. G.; Chino, M.; Maglio, O.; Lombardi, A.; Rega, N. Unveiling the Structure of a Novel Artificial Heme-Enzyme with Peroxidase-Like Activity: A Theoretical Investigation. *Biopolymers* **2018**, *109*, e23225.
- (69) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·Log(N) Method for Ewald Sums in Large Systems. *Journal of Chemical Physics* **1993**, *98*, 10089-10092.
- (70) Gotz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with Amber on Gpus. 1. Generalized Born. *Journal of Chemical Theory and Computation* **2012**, *8*, 1542-1555.
- (71) Salomon-Ferrer, R.; Gotz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with Amber on Gpus. 2. Explicit Solvent Particle Mesh Ewald. *Journal of Chemical Theory and Computation* **2013**, *9*, 3878-3888.
- (72) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of N-Alkanes. *Journal of Computational Physics* **1977**, *23*, 327-341.
- (73) Roe, D. R.; Cheatham, T. E. Ptraj and Cpptraj: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and Computation* **2013**, *9*, 3084-3095.
- (74) Bannwarth, C.; Ehlert, S.; Grimme, S. Gfn2-Xtb—an Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *Journal of Chemical Theory and Computation* **2019**, *15*, 1652-1671.
- (75) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (Dft-D) for the 94 Elements H-Pu. *Journal of Chemical Physics* **2010**, *132*, 154104.
- (76) Ufimtsev, I. S.; Martinez, T. J. Quantum Chemistry on Graphical Processing Units. 3. Analytical Energy Gradients, Geometry Optimization, and First Principles Molecular Dynamics. *Journal of Chemical Theory and Computation* **2009**, *5*, 2619-2628.

- (77) Seritan, S.; Bannwarth, C.; Fales, B. S.; Hohenstein, E. G.; Isborn, C. M.; Kokkila-Schumacher, S. I. L.; Li, X.; Liu, F.; Luehr, N.; Snyder, J. W.; Song, C. C.; Titov, A. V.; Ufimtsev, I. S.; Wang, L. P.; Martinez, T. J. Terachem: A Graphical Processing Unit-Accelerated Electronic Structure Package for Large-Scale Ab Initio Molecular Dynamics. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2021**, *11*, e1494.
- (78) Lange, A. W.; Herbert, J. M. A Smooth, Nonsingular, and Faithful Discretization Scheme for Polarizable Continuum Models: The Switching/Gaussian Approach. *Journal of Chemical Physics* **2010**, *133*.
- (79) York, D. M.; Karplus, M. A Smooth Solvation Potential Based on the Conductor-Like Screening Model. *Journal of Physical Chemistry A* **1999**, *103*, 11060-11079.
- (80) Bondi, A. Van Der Waals Volumes and Radii. *Journal of Physical Chemistry* **1964**, *68*, 441-451.
- (81) Liu, F.; Luehr, N.; Kulik, H. J.; Martínez, T. J. Quantum Chemistry for Solvated Molecules on Graphical Processing Units Using Polarizable Continuum Models. *J. Chem. Theory Comput.* **2015**, *11*, 3131-3144.
- (82) Liu, F.; Sanchez, D. M.; Kulik, H. J.; Martinez, T. J. Exploiting Graphical Processing Units to Enable Quantum Chemistry Calculation of Large Solvated Molecules with Conductor-Like Polarizable Continuum Models. *International Journal of Quantum Chemistry* **2019**, *119*, e25760.
- (83) Ranieri, A.; Monari, S.; Sola, M.; Borsari, M.; Battistuzzi, G.; Ringhieri, P.; Natri, F.; Pavone, V.; Lombardi, A. Redox and Electrocatalytic Properties of Mimochrome Vi, a Synthetic Heme Peptide Adsorbed on Gold. *Langmuir* **2010**, *26*, 17831-17835.
- (84) Vitale, R.; Lista, L.; Lau-Truong, S.; Tucker, R. T.; Brett, M. J.; Limoges, B.; Pavone, V.; Lombardi, A.; Balland, V. Spectroelectrochemistry of Feiii- and Coiii-Mimochrome Vi Artificial Enzymes Immobilized on Mesoporous Ito Electrodes. *Chemical Communications* **2014**, *50*, 1894-1896.
- (85) Karelina, M.; Kulik, H. J. Systematic Quantum Mechanical Region Determination in Qm/Mm Simulation. *Journal of Chemical Theory and Computation* **2017**, *13*, 563-576.
- (86) Kulik, H. J.; Zhang, J.; Klinman, J. P.; Martínez, T. J. How Large Should the Qm Region Be in Qm/Mm Calculations? The Case of Catechol O-Methyltransferase. *Journal of Physical Chemistry B* **2016**, *120*, 11381-11394.
- (87) Fonseca Guerra, C.; Handgraaf, J.-W.; Baerends, E. J.; Bickelhaupt, F. M. Voronoi Deformation Density (Vdd) Charges: Assessment of the Mulliken, Bader, Hirshfeld, Weinhold, and Vdd Methods for Charge Analysis. *Journal of Computational Chemistry* **2004**, *25*, 189-210.
- (88) Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The Orca Quantum Chemistry Program Package. *Journal of Chemical Physics* **2020**, *152*.
- (89) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (Dft-D) for the 94 Elements H-Pu. *J Chem Phys* **2010**, *132*.

- (90) Lee, C. T.; Yang, W. T.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron-Density. *Physical Review B* **1988**, *37*, 785-789.
- (91) Becke, A. D. Density-Functional Thermochemistry. Iii. The Role of Exact Exchange. *Journal of Chemical Physics* **1993**, *98*, 5648-5652.
- (92) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *Journal of Physical Chemistry* **1994**, *98*, 11623-11627.
- (93) Li, L.; Li, C.; Zhang, Z.; Alexov, E. On the Dielectric “Constant” of Proteins: Smooth Dielectric Function for Macromolecular Modeling and Its Implementation in Delphi. *Journal of Chemical Theory and Computation* **2013**, *9*, 2126-2136.
- (94) Tin Kam, H. Proceedings of 3rd International Conference on Document Analysis and Recognition, 1995; p 278-282 vol.1.
- (95) Rumelhart, D. E.; Hinton, G. E.; Williams, R. J. Learning Representations by Back-Propagating Errors. *Nature* **1986**, *323*, 533-536.
- (96) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, É. Scikit-Learn: Machine Learning in Python. *Journal of Machine Learning Research* **2011**, *12*, 2825–2830.
- (97) Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; Desmaison, A.; Köpf, A.; Yang, E.; DeVito, Z.; Raison, M.; Tejani, A.; Chilamkurthy, S.; Steiner, B.; Fang, L.; Bai, J.; Chintala, S. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems*; Curran Associates Inc., 2019, Article 721
- (98) Krizhevsky, A.; Sutskever, I.; Hinton, G. E. Imagenet Classification with Deep Convolutional Neural Networks. *Communications of the ACM* **2017**, *60*, 84–90.
- (99) Kingma, D. P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2015**.
- (100) Miller, B. R.; McGee, T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. Mmpbsa.Py: An Efficient Program for End-State Free Energy Calculations. *Journal of Chemical Theory and Computation* **2012**, *8*, 3314-3321.
- (101) Tsui, V.; Case, D. A. Theory and Applications of the Generalized Born Solvation Model in Macromolecular Simulations. *Biopolymers* **2001**, *56*, 275-291.
- (102) Ichiye, T.; Karplus, M. Collective Motions in Proteins: A Covariance Analysis of Atomic Fluctuations in Molecular Dynamics and Normal Mode Simulations. *Proteins* **1991**, *11*, 205-217.
- (103) McClendon, C. L.; Friedland, G.; Mobley, D. L.; Amirkhani, H.; Jacobson, M. P. Quantifying Correlations between Allosteric Sites in Thermodynamic Ensembles. *Journal of Chemical Theory and Computation* **2009**, *5*, 2486-2502.
- (104) Guo, J.; Zhou, H.-X. Protein Allostery and Conformational Dynamics. *Chemical Reviews* **2016**, *116*, 6503-6515.

- (105) Pfanzagl, V.; Nys, K.; Bellei, M.; Michlits, H.; Mlynek, G.; Battistuzzi, G.; Djinovic-Carugo, K.; Van Doorslaer, S.; Furtmüller, P. G.; Hofbauer, S.; Obinger, C. Roles of Distal Aspartate and Arginine of B-Class Dye-Decolorizing Peroxidase in Heterolytic Hydrogen Peroxide Cleavage. *Journal of Biological Chemistry* **2018**, *293*, 14823-14838.
- (106) Lučić, M.; Wilson, M. T.; Svistunenko, D. A.; Owen, R. L.; Hough, M. A.; Worrall, J. A. R. Aspartate or Arginine? Validated Redox State X-Ray Structures Elucidate Mechanistic Subtleties of Fe^{IV}=O Formation in Bacterial Dye-Decolorizing Peroxidases. *JBIC Journal of Biological Inorganic Chemistry* **2021**, *26*, 743-761.
- (107) Lučić, M.; Svistunenko, D. A.; Wilson, M. T.; Chaplin, A. K.; Davy, B.; Ebrahim, A.; Axford, D.; Tosha, T.; Sugimoto, H.; Owada, S.; Dworkowski, F. S. N.; Tews, I.; Owen, R. L.; Hough, M. A.; Worrall, J. A. R. Serial Femtosecond Zero Dose Crystallography Captures a Water-Free Distal Heme Site in a Dye-Decolorising Peroxidase to Reveal a Catalytic Role for an Arginine in Fe^{IV}=O Formation. *Angewandte Chemie International Edition* **2020**, *59*, 21656-21662.
- (108) Deemagarn, T.; Wiseman, B.; Carpena, X.; Ivancich, A.; Fita, I.; Loewen, P. C. Two Alternative Substrate Paths for Compound I Formation and Reduction in Catalase–Peroxidase Katg from *Burkholderia pseudomallei*. *Proteins* **2007**, *66*, 219-228.
- (109) Rinkovec, T.; Tarana, S.; Modrušan, M.; Vidović, N.; Tomić, A.; Cindro, N.; Speranza, G.; Tomišić, V.; Horvat, G. Anion Complexes of Small Boc-Protected Lysine Cyclopeptides in Solution. *Journal of Molecular Liquids* **2025**, *433*, 127837.
- (110) FitzSimons, T. M.; Duti, I. J.; Conrad, N.; Agrawal, P.; Niemoeller, A.; Guinn, E.; Wang, J.; Vasiukhina-Martin, A.; Rosales, A. M. Gelation Behavior of Short Protected Peptides in Organic Medium. *Soft Matter* **2025**, *21*, 4751-4760.
- (111) Zong, Q.; Geng, H.; Ye, L.; Zhang, A.; Shao, Z.; Feng, Z. Synthesis and Chelation Capability of Fmoc and Boc Mono-Substituted Cyclo(L-Lys-L-Lys)S. *Chemical Research in Chinese Universities* **2016**, *32*, 484-492.
- (112) Zambrano, G.; Chino, M.; Renzi, E.; Di Girolamo, R.; Maglio, O.; Pavone, V.; Lombardi, A.; Natri, F. Clickable Artificial Heme-Peroxidases for the Development of Functional Nanomaterials. *Biotechnology and Applied Biochemistry* **2020**, *67*, 549-562.
- (113) Renzi, E.; Esposito, A.; Leone, L.; Chávez, M.; Pineda, T.; Lombardi, A.; Natri, F. Biohybrid Materials Comprising an Artificial Peroxidase and Differently Shaped Gold Nanoparticles. *Nanoscale Advances* **2024**, *6*, 3533-3542.
- (114) Esposito, A.; Leone, L.; De Simone, A.; Fusco, G.; Natri, F.; Lombardi, A. Catalytic Nanomaterials by Conjugation of an Artificial Heme-Peroxidase to Amyloid Fibrils. *ACS Applied Materials & Interfaces* **2024**, *16*, 45371-45382.
- (115) Fleetwood, O.; Kasimova, M. A.; Westerlund, A. M.; Delemotte, L. Molecular Insights from Conformational Ensembles Via Machine Learning. *Biophysical Journal* **2020**, *118*, 765-780.

Data Availability Statement.

All data in this manuscript is available in the Supporting Information or on Zenodo at the DOI: 10.5281/zenodo.16934208.