Multiscale Computational Investigation of the C-H Activation Reactivity of Engineered Myoglobin

A dissertation

submitted in the partial fulfilment of the degree

of

MASTER OF SCIENCE

IN

CHEMISTRY

Submitted by

KRITIKA

(Roll No. 302102006)

Under the guidance of

Dr. DEBASISH MANDAL

(Assistant Professor)

School of Chemistry and Biochemistry

Thapar Institute of Engineering and Technology (Deemed to be University)

Patiala-147004, Punjab (India)

www.thapar.edu

(July 2023)

CERTIFICATE

This is to certify that the dissertation entitled "Multiscale Computational Investigation of the C-H Activation Reactivity of Engineered Myoglobin", being submitted by Ms. Kritika in partial fulfilment of the requirement for the award of the degree of Masters of Science in Chemistry and being submitted to the School of Chemistry and Biochemistry, Thapar Institute of Engineering and Technology, Patiala is a bonafide work carried out by her under my supervision. The work has reached the standard necessary for submission, and the contents of this dissertation have not been submitted to any other university or institute for the award of any degree or diploma.

Debasish mandal

Dr. Debasish Mandal

Assistant Professor

School of Chemistry and Biochemistry

Thapar Institute of Engineering and Technology, Patiala - 147004

DECLARATION

I hereby declare that the thesis entitled "Multiscale Computational Investigation of the C-H Activation Reactivity of Engineered Myoglobin" is an accurate representation of the work I completed for the partial fulfilment of the requirements for the award of the degree of Master of Science in Chemistry, at Thapar Institute of Engineering and Technology, Patiala, under the guidance of Dr. Debasish Mandal, Assistant Professor, School of Chemistry & Biochemistry, Thapar Institute of Engineering and Technology, Patiala, during January' 2023 to July' 2023. No portion of the information contained in this report has been provided to another university or institute for the award of any degree.

Kritika

Kritika

Date: 17-07-2023

It is certified that the student's above statement is correct to the best of my knowledge and belief.

Debasish mandal

Dr. Debasish Mandal

Assistant Professor

School of Chemistry & Biochemistry

Thapar Institute of Engineering and Technology, Patiala - 147004

ACKNOWLEDGEMENT

I would like to express my deepest gratitude to my thesis supervisor, Dr. Debasish Mandal, for his guidance, expertise, and unwavering support. His insightful feedback, constructive criticism, and dedication played a pivotal role in shaping this research project.

I am indebted to the Thapar Institute of Engineering and Technology & School of Chemistry and Biochemistry for providing a stimulating academic environment, access to resources, and opportunities for intellectual growth. Their commitment to excellence in education has been instrumental in shaping my academic journey. I extend my deepest thanks to Dr. Satnam Singh, Head SCBC, Thapar Institute and Engineering and Technology, Patiala, for providing me with this opportunity.

I'd like to thank INSPIRE GRANT [DST/INSPIRE/04/2016/001948] for providing computational resources and financial assistance.

Lastly, I want to acknowledge my family and friends for their encouragement and unwavering belief in my abilities. Their constant support and encouragement have been a source of strength throughout my academic pursuits.

Kritika

Kritika

ABSTRACT

The current thesis reported a combined MD and QM/MM investigation for the C-H activation reactivity of wild-type and three reconstituted myoglobin. We have tried to unravel the reactivity of myoglobin reconstituted with Mn porphycene [rMb(MnPC)]. Along with it, we also explored why the reconstituted iron porphycene [rMb(FePC)] is not reactive. From the MD analysis, it was observed that the substrate ethyl-benzene is not stable in the active site of [rMb(FePC)], whereas it is sufficiently stable in the pocket of [rMb(MnPC)]. This may be linked with the reactivity of the [rMb(MnPC)]. Axial histidine ligand (HIS93) might also play an important role in the reactivity here, as in the case of [rMb(MnPC)], it goes far away from the central Mn porphycene. QM/MM investigations show that rMb(MnPc)'s active Mn(IV)=O species extract hydrogen atoms in a similar way to the CYP450's Cpd I species, and the triplet state is more favorable than the singlet state.

CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS AND SYMBOLS

Abbreviation	Full form
nMb	Native Myoglobin
rMb	Reconstituted Myoglobin
MD	Molecular Dynamics
QM/MM	Quantum Mechanics/Molecular Mechanics
CYP	Cytochrome P450
CPO	Chloroperoxidase
CpdI	Compound-I
MCPB	Metal Centre Parameter Builder
EPR	Electron Paramagnetic Resonance
EDA	Ethylenediamine
DFT	Density Functional Theory
PDB	Protein Data Bank
AMBER	Assisted Model Building with Energy Refinement
RESP	Restrained Electrostatic Potential
GAFF	General AMBER force field
RMSD	Root-Mean-Square deviation
CPPTRAJ	C++ Trajectory Analysis
PES	Potential Energy Surface
PRFO	Partitioned Rational Function Optimization
HDLC	Hybrid Delocalized Internal Coordinates
rMb(MnPc)	Reconstituted Myoglobin with Manganese Porphycene
rMb(MnPor)	Reconstituted Myoglobin with Manganese Porphyrin
rMb(FePc)	Reconstituted Myoglobin with Iron Porphycene
nMb(FePor)	Reconstituted Myoglobin with Iron Porphyrin
HAT	Hydrogen Atom Transfer
ZPE	Zero-Point Energy
TS	Transition State
PCET	Proton-Coupled Electron Transfer
SNO	Spin Natural Orbitals
SP	Single Point Energy

Chapter 1: Introduction

Heme proteins are ubiquitous in countless biological processes e.g., oxygen transportation, hydroxylation, peroxide decomposition, electron transfer, oxidation, enzymatic catalysts, cellular signaling, etc[1,2]. These proteins rely on metal cofactors and versatile coordinating features for their unique catalytic activity, and the reactivity of these cofactors is also tightly regulated by the protein matrix[3]. Figure 1 depicts a few instances of reactions catalyzed by compound I of the heme protein.

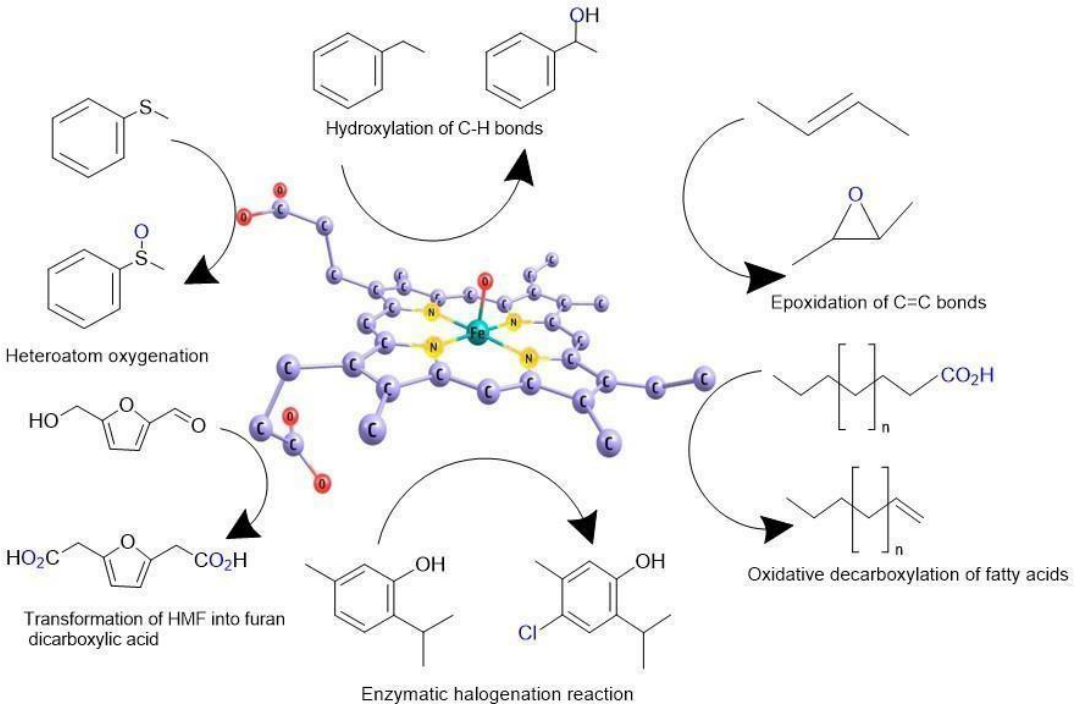


Figure 1: Few instances of the reactions catalyzed by compound I of the heme protein.

Hemoproteins possess a heme group, i.e., an iron ion coupled with a porphyrin ring as a cofactor[1]. The porphyrin macrocycle is highly symmetrical with four pyrrole units connected by the methine bridges and exhibits a dianionic character. The central metal iron can exhibit variable oxidation states, thereby facilitating the protein's involvement in redox reactions.

Among the numerous classes of heme proteins, our focus is mainly concentrated on myoglobin. Myoglobin's primary function is in muscle cells, where it helps regulate the storage and flow of oxygen. The protein chain, also called as globin, shields the heme group from oxo-heme attack, preventing the formation of hematin and the irreversible auto- oxidation of heme. The hydrophobic active site of the protein is composed of a heme b (Fe- porphyrin) cofactor that is non-covalently bound to it. Here Fe is linked to a proximal histidine residue (HIS93), and oxy-myoglobin is stabilized by a hydrogen bonding interaction between O2 and a distal

histidine residue (HIS93), and oxy-myoglobin is stabilized by a hydrogen bonding interaction between O2 and a distal histidine residue (HIS64)[4,5]. Crystal structure of myoglobin is shown is Figure 2.

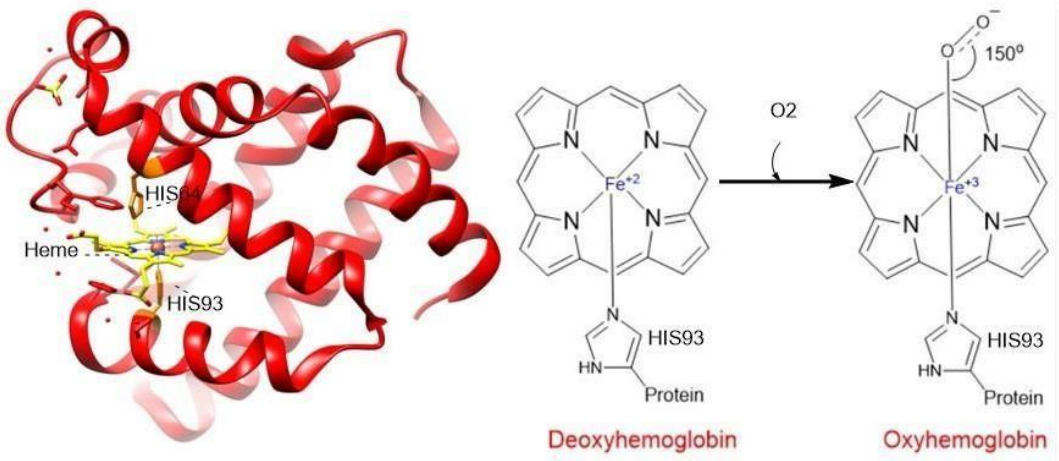


Figure 2: Crystal structure of myoglobin and its action.

However, unlike peroxidases and cytochrome P450s, myoglobin's active site is not ideal for activating small molecules such as H_2O_2 and O_2 and catalyzing alkane hydroxylation via C-H bond activation, which may be due to the lack of substrate-binding site in myoglobin. The fundamental structural differences in their active sites are shown in Figure 3[1,6].

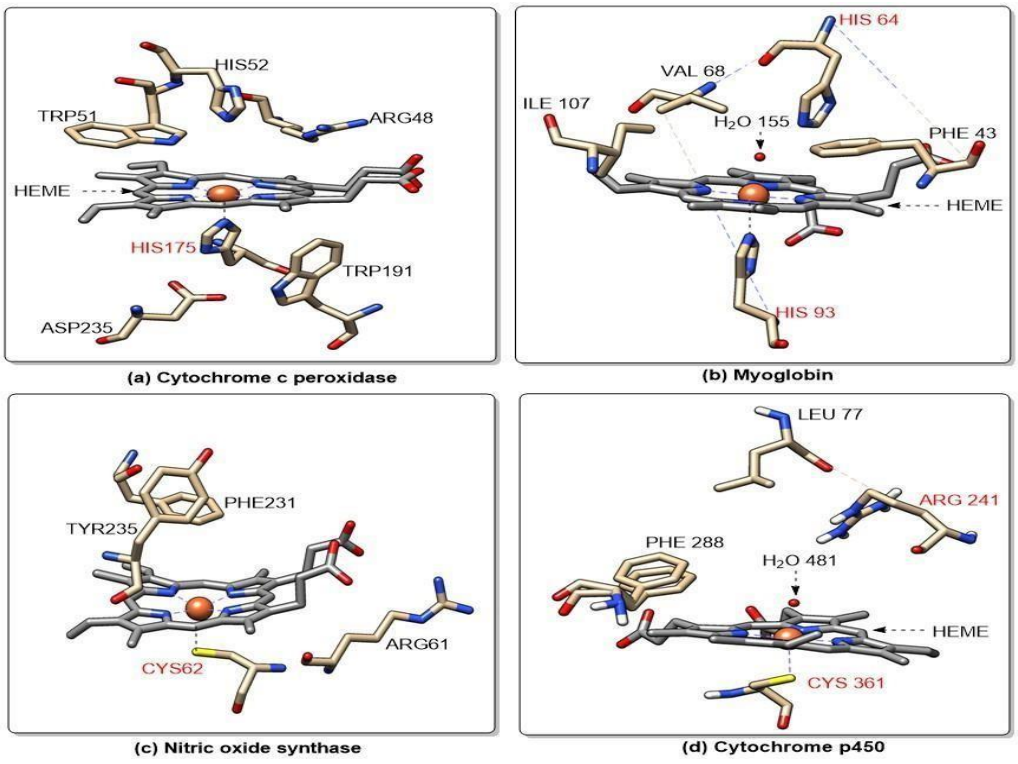


Figure 3: Difference between the active sites of various classes of heme proteins, revealing the different active amino-acid residues and their cofactors (a) cytochrome c peroxidase (b) myoglobin (c) nitric oxide synthase (d) cytochrome p450.

It opened up the possibilities of enzyme engineering towards myoglobin for improving enzymatic functions for C-H activation and other reactions[7,8]. This will enhance understanding of the diverse interactions occurring within the active site's second coordination sphere and the enzyme complex catalytic mechanisms, as each ligand displays different redox properties.

Heme b of myoglobin can easily be removed due to its weak non-covalent interaction with protein. The resultant apo-protein serves as a viable framework for the reaction scaffolding. Engineering of hemoproteins for the creation of new and useful proteins can be done in four main ways as shown in Figure 4.

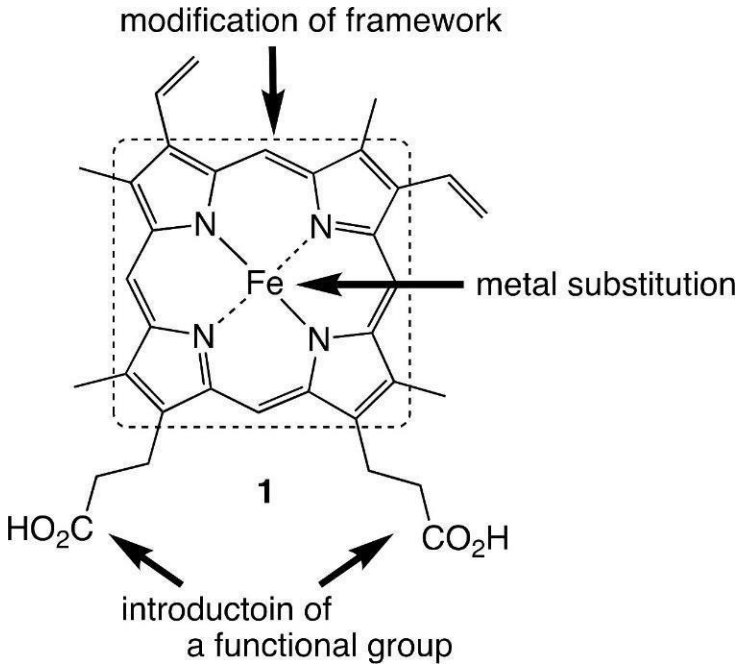


Figure 4: The molecular structure of heme b, along with three distinct heme modifications.

(a) modifications of peripheral propionate side chains with different functional groups[9], (b) substitution of the central metal atom of heme group with other metal atoms like Mn, Co, Ru, (c) the introduction of mutations to the amino acid residues within the active site, as well as (d) substitution of non-porphyrinoid cofactors for the native cofactor (see Figure 5). These cofactors include porphycene, corrole, phthalocyanine, tetradehydrocorrin and corrolazine[10,11]. A schematic image for these replacements has been presented in Figure 5.

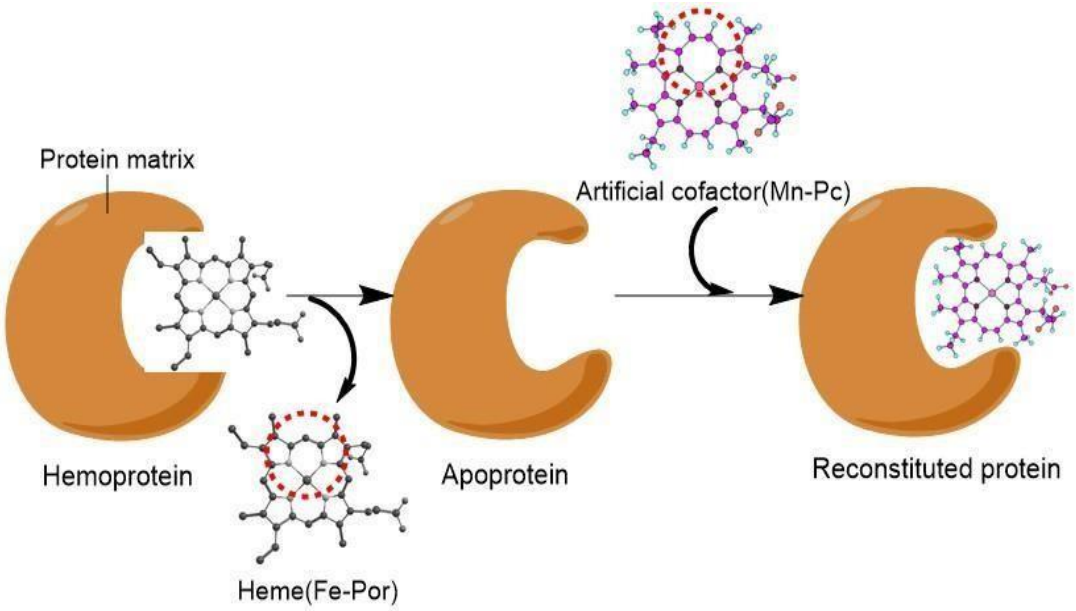


Figure 5: A schematic diagram for preparing protein conjugates based on the heme group reconstitution.

The heme group is easily removed from the native protein under acidic conditions, yielding the apoprotein, which is then reconstituted with the synthetic non-porphyrinoid cofactors.

In recent years, Hayashi et al. conducted a study of the structural characteristics and chemical reactivity of various hemoproteins utilizing the methodology of enzyme engineering. They substituted the native protoporphyrin with porphycene in myoglobin and discovered that metal-porphycene myoglobin complexes can aid in dioxygen binding[4]. Porphycene was used to substitute porphyrin since it is an isomer of porphyrin with completely different structural and electrical features[12]. It has a linearized porphyrin ring system with two neighboring pyrrole units joined by two carbon-carbon bonds and two pyrrole-pyrrole linkages. This carbon-carbon bond changes the electrical distribution of the ring, increasing aromaticity and redox characteristics compared to typical porphyrins. Porphycene linearity enhances conjugation and pi-electron delocalization.

Oohara and co-workers reported that on reconstitution of nMb porphyrin with manganese porphycene macrocycle, peroxidase activity of rMb(MnPc) increases towards hydroxylation of ethylbenzene(C-H bond) with a turnover frequency of 33 h^{-1} and the turnover number of 13. Fe porphycene, manganese porphyrin and iron porphyrin did not show the same reactivity as shown in Figure 6.

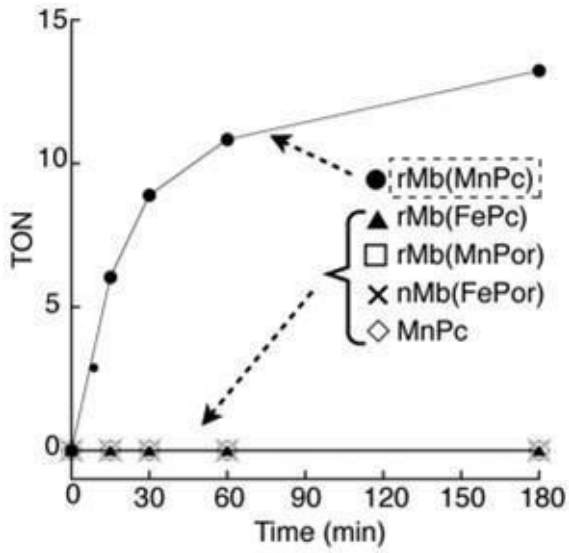


Figure 6: The time course plots of TON for the hydroxylation reaction of ethyl benzene with native and reconstituted myoglobin. The picture has been adopted from ref 4.

Furthermore, it is observed that the manganese porphycene (MnPc) in the absence of a protein matrix does not display any apparent catalytic activity. This observation strongly suggests that the protein matrix plays a crucial role in regulating the hydroxylation reaction. The iron ion present in myoglobin exhibits a high degree of suitability in executing its oxidative chemistry. The replacement of iron with manganese, which lacks the ability to bind dioxygen in air, has led to the investigation of new forms of reactivity that are not observed in the natural form of myoglobin. For the oxygen molecule to bond effectively, unpaired electrons are needed, which manganese doesn't have in its ground state. Manganese porphycene of reconstituted protein rMb(MnPc) was characterized by UV−vis spectroscopy and electrospray ionization mass spectrometry. The high-resolution (2.2 Å) crystal structure of the rMb(MnIIIPc) was reported. It showed the coordination bond between HIS93 and Mn, which is 0.36 Å longer than that of the manganese porphyrin analogue. The distal HIS64 residue exists in two different conformations and stabilizes the oxo complex through hydrogen bonding interaction. This crystal structure predicts that MnPc will provide both six- and five-coordinated complexes in the heme pocket.

C-H bond hydroxylation experimental mechanistic studies:

The investigation involved the utilization of logarithmic representation to analyze the rate constants related to the hydroxylation process of ethyl benzene, toluene, and cyclohexane, catalyzed by rMb. The plotted data was then correlated with bond dissociation energy,

revealing a negative slope and linear relationship. The observed kinetic isotope effects suggested that the activation of C−H bonds serves as the rate-limiting step[13] in the reaction. An active metal oxo species is formed, which is also known as compound-I. rMb(FePc) and rMb(MnPor) do not form compound I-like entities and hence cannot perform HAT(hydrogen atom abstraction) reactions. rMb(FePc) and rMb(MnPor) do not form compound I-like entities and hence cannot perform hydrogen atom abstraction (HAT) reaction.

Experimentally, the oxidation state of Mn in rMb(MnPc) is suggested as +5 on the basis of the absence of typical porphycene pi-cation radical spectra. It also follows OH-rebound mechanism similar to cytochrome p450, as presented in following reaction[14].

$$Mn(V)=O(Pc) + R{-}H \rightarrow Mn(IV)\text{-}OH(Pc) + R\cdot \rightarrow Mn(III)Pc + R{-}OH \qquad (1)$$

The present work focuses on theoretical studies explaining the higher oxidative reactivity of manganese porphycene reconstituted myoglobin compared to MnPor, FePor and FePc of myoglobin. A rigorous molecular dynamics simulation along with the multiscale QM/MM studies have been performed to investigate the ligand effect at the active site and dynamic behavior of myoglobin at the atomic level and to get insights into the energetics, activation barrier and pathways of the reaction.

Chapter2: Literature review

The experimental findings by Jonathan Rittle *et al.* confirmed the presence and catalytic proficiency of compound I, which is an active metal oxo species in cytochrome p450[15]. Based on the experimental observations, it has been determined that CYP119-I exhibits the ability to catalyze the oxidation of unreactive hydrocarbons. The CYP119 intermediate's Mössbauer characteristics and spectra matched those for CPO-I that have been previously reported. This resemblance made it easier to identify the CYP119 intermediate: The P450 Cpd I was best characterized as featuring an iron(IV)oxo moiety in dynamic equilibrium with a ligand-centered radical species. They reported KIEs ≥ 10 for p450 hydroxylation, which accorded with results from studies on intramolecular competition. The extent of this phenomenon suggested that CYP I extract hydrogen from a substrate. An iron(IV)hydroxide species is formed, which promptly interacts with the substrate to yield a hydroxylated product. Groves proposed this rebound mechanism.

The introduction of a manganese ion in place of an iron ion of the heme moiety of myoglobin has yielded a protein exhibiting enhanced reactivity, as evidenced by its proficient capacity to catalyze C-H bond cleavage, as reported by Kari and co-workers[16]. The oxidation of Mn-myoglobin by meta chloro-perbenzoic acid resulted in the formation of a reaction intermediate that exhibited reactivity towards 1,4 cyclohexadiene. Using spectroscopy techniques, the intermediate was identified as a Manganese-oxo species. Groves and workers investigation compared the intermediate's EPR spectra to the fully described manganese- substituted horseradish peroxidase. These results demonstrate that this species in Mn- myoglobin oxidizes chemically inert C-H bonds.

Yuan-Bo et al. created various mutants to study how Mn protoporphyrin IX activates H2O2 in myoglobin[17]. They reported that H_2O_2 activation rates increased 5-fold and 10-fold in the F43H Mn(III)Mb mutant. First, it was noted that the dual distal histidine work together to activate H_2O_2 by Mn porphyrin within a protein scaffold. The reactivity to activate H_2O_2, therefore, changes significantly as a result of adjusting the distal histidine to bind and orient H_2O_2, as confirmed by their UV-vis spectra.

The study conducted by Elnaz et al. looked into the examination of the influence exerted by different axial ligands on the catalytic function of Mn-Por[18]. Their findings indicate that the catalytic activity of manganese porphyrin depends upon the presence and nature of hydrogen bonding interactions between the ligand and the ring's functional groups. Based on the

theoretical findings, it is postulated that the presence of these ligands has the potential to modulate the S/Q and S/T energy gaps, thereby influencing the reactivity. The observed phenomenon of Mn-N bond shortening can be attributed to the occurrence of hydrogen bonding interactions between the carbon-hydrogen (C-H) groups of pyridines and the methoxy moiety of the phenyl ring. The robust sigma bonding interaction between manganese and Npy induces a reduction in the strength of the manganese-oxygen (Mn-O) bond, owing to the influence of the manganese-nitrogen axial Mn-N(axial) bond on the Mn-O bond. Thus, this interaction resulted in higher reactivity of Mn porphyrin.

Reconstituted myoglobin containing iron porphycene is more reactive towards styrene cyclopropanation than wild-type myoglobin, according to Oohara et al. They revealed that rMb accelerates the catalysis by a factor of 26 when compared to nMb[10]. Using stopped-flow techniques, they investigated the transient changes in the absorption spectrum profile of rMb and nMb in response to the EDA. The reaction of reconstituted myoglobin with EDA produced a new species, which was the reactive metallocarbene species. This was the first time an active metallocarbene species in a hemoprotein had been observed using spectroscopy. For native myoglobin, this metallocarbene species was reported as not active. Theoretically, rMb produced a lower kinetic barrier and fewer steps in intersystem crossing than nMb in the DFT calculations. This difference is most likely caused by porphycene's stronger ligand field effect, which stabilizes the triplet of the resting state when compared to porphyrin.

Chapter 3: Theoretical details

3.1 : MD simulations

Molecular dynamics simulations employ computational algorithms to numerically solve the equations of motion derived from Newton's laws of motion. These simulations enable the investigation of the intricate motion and behavior of atoms within a molecular assembly by determining the resultant force and subsequent acceleration experienced by each individual atom[19]. In a conventional MD simulation, atomic locations are updated in real-time using Newton's second law.

$$F_{(o)} = m_{(o)} + a_{(o)}$$

Where F_o is the force acting on an object o and a_o is the acceleration with object's mass m_o.

The equations for potential energy used in Molecular Mechanics to determine energy and their parameters/constants are called Force Fields. Force field energy includes bonded terms (Bond, Angle, Torsion, Improper) for covalently bound atoms and nonbonded terms for van der Waals and long-range electrostatic interactions.

$$E_{(TOTAL)} = E_{(BONDED\ INTERACTIONS)} + E_{(NONBONDED\ INTERCATIONS)}$$

An ensemble links the microscopic motions of particles to the system's macroscopic properties. All MD simulations in this study used the conventional ensemble (NVT) or the isothermal-isobaric ensemble (NPT). After specifying particle positions and velocities, Newton's equations must be numerically integrated. The velocity verlet method calculates locations and velocities simultaneously at each timestep. The selection of an appropriate timestep (Δt) for mathematical integration is of utmost importance in molecular dynamics simulations, as it influences the convergence of the computational calculations. Utilizing smaller timesteps in simulations enhances accuracy, although at the expense of the raised cost of computation. Increasing the timestep results in a greater exploration of the conformational space, yet it also introduces fluctuations in the simulation.

There are three types of boundary conditions (BC) that may be set in MD simulations: (a) Periodic boundary conditions (b), reflecting wall conditions (c), and a vacuum state. The PBC is widely employed approach, where system is replicated infinitely in all directions, creating a periodic lattice that allows for the simulation of an infinite system.

The computational investigations were carried out employing the Amber software package. The Amber software suite encompasses Amber tools for the purpose of analysis and configuration of simulations. Nucleic acids and amino acid residues are included in a set of molecular mechanical force fields. In order to generate parameters for metal containing entities present in the proteins, one approach that makes use of bonded models is the metal center parameter builder (MCPB)[20].

The major coordination sphere of the metal-protein complex was the focus of the ab initio calculations, which were performed using the B3LYP/6-31G* basis set. The resulting force field parameters are more accurate approximations of the AMBER force field (AMBER ff14SB). Charges (qi) and bond (Kr) and angle (K) force constants are among the other found parameters. MCPB.py can represent over 80 M+ and offers compatibility with diverse AMBER force fields. For our work, this methodology was employed to parameterize the iron (Fe) within the heme moiety and manganese (Mn).

3.2 : QM/MM Method

In recent years, hybrid QM/MM techniques have emerged as an essential resource for understanding enzyme reactions[21]. On the one hand, the protein's surroundings are crucial to the reaction mechanism and must be addressed explicitly. Solvated proteins with thousands of atoms require classical force fields because QM methods cannot handle them. On the other hand, MM approaches are limited to describing the dynamics of proteins and will not work for chemical reactions. Therefore, the best way to explore enzymatic reactions is to use hybrid calculations, where the reactive centre is addressed using a QM method, and the protein and solvent molecules are treated with the MM method as depicted by Figure 7.

Due to its advantageous cost/performance ratio, density functional theory (DFT) is the approach that is most frequently utilized in QM/MM computations. DFT offers an effective method for investigating the behavior of electrons within a system by analyzing their electron density distribution. Its fundamental principles revolve around the Hohenberg-Kohn theorems[22], which define that electron density serves as the exclusive determinant of both the energy and ground-state electron density function of the entire system. B3LYP is one of the most often used exchange-correlation functionals because it typically produces satisfactory results for a variety of chemical systems.

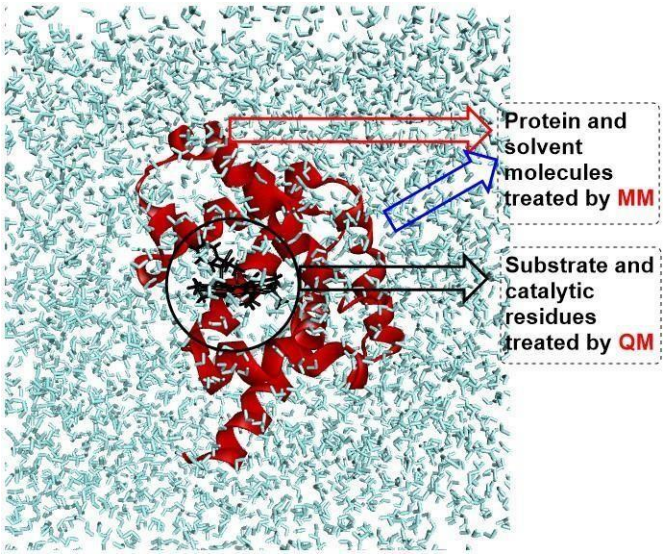


Figure 7: QM/MM model of a protein showing different protein and solvent treated with MM method and active region treated by QM.

Van der Waals dispersive interactions are improperly described by DFT. As a result, empirical corrections that take dispersion into account have been developed (DFT-D technique). The additive approach and the subtractive approach are two common types of QM/MM techniques. The energy of the MM section of the system is calculated using the force field and added to the energy of the QM active area in the additive technique.

$$E_{(total)} = E_{QM(QM)} + E_{MM(MM)} + E_{QM\text{-}MM}$$

The energy of the whole system is calculated using the subtractive approach at the MM level. Then, the QM energy of the active region is added to the MM energy of the entire system. In order to avoid counting the same energy twice, the molecular mechanics energy of the active part is subtracted.

$$E_{(total)} = E_{(MM+total)} + E_{QM(QM)} - E_{MM(QM)}$$

The concept of electrostatic embedding, which considers how the atomic charges in the MM area polarize the QM region, is supported by both the techniques. The van der Waals interactions are often evaluated at the MM level. A Lennard-Jones potential is typically used to describe these interactions. The bond which is truncated at the (QM-MM) interface necessitates the application of a capping strategy. These include the boundary-atom, link- atom and localized-orbital methods. We have used link-atom method, which introduces a non-real atomic centre L (typically a hydrogen atom). It is attached to another MM region atom by a covalent bond and completely saturates its free valency.

Chapter4: Computational Methodology

4.1 Setup of the system:

The crystal structures of horse heart myoglobin reconstituted with manganese porphycene (PDB code 3WI8) and horse heart met manganese myoglobin (PDB code 2O58), with resolutions of 2.2 Å and 1.65 Å, respectively, were used to determine the initial structures of the enzymes rMb(MnPc) and rMb(MnPor). The crystal structure of native myoglobin nMb(FePor) was taken from PDB code IYMB, and for myoglobin containing iron porphycene Mb(FePc), it was 2D6C. The distal water that was present in the rMb(MnPc) at a distance of 2.790 above Mn was changed to an oxygen atom using the chemcraft software. Similarly, distal water that was present 2.521 m above Mn was modified into an oxygen atom in rMb(MnPor). Compound I, the oxo manganese complex, will be formed as a result of this. The ethylbenzene substrate was modelled into the active site of rMb(MnPc) using UCSF Chimera docking analysis. In this study, the active site of rMb(MnPc) was utilized to model the ethylbenzene substrate via UCSF Chimaera docking analysis. The best docking pose was subsequently selected for additional molecular dynamics (MD) simulations. Using the pdb2pqr programme, the pKa values of amino acid residues were obtained in order to determine the protonation status of the residues. PDB2PQR employs the PROPKA algorithm to determine the protonation states of ionizable residues. The forcefield parameterization of oxo manganeseporphyrin and oxo manganese porphycene complexes was carried out utilizing the Metal Centre Parameter Builder (MCPB) tools. Force constants and RESP[23] charges were calculated for the Mn=O complex, the axial histidine, and the substrate using Gaussian 16 at the level of B3LYP/6-31G*.

The ethylbenzene parameters were determined using the GAFF, in the Antechamber module of Amber 20. The partial atomic charges for the substrate were obtained using the AM1-BCC charge technique. The protein surface was neutralized with the counter ions i.e cl^- ion, to lower the total charge of the system. Each enzyme complex was solvated with TIP3P water molecule- filled rectangular box, with a protein system boundary of 10 Å from the box margins. All of the MD simulations for the complex used the Amber ff14SB force field. A similar system setup procedure was followed for the rMb(MnPor), nMb(FePor) and rMb(FePc).

4.2 MD simulations:

Both protein-substrate complexes were simulated using molecular dynamics (MD) using the pmemd.cuda of the Amber20 software and the Amber ff14SB force field. The system's geometries were minimized to eliminate the bad contacts and relax the system after proper parameterization and setup. The entire complex was minimized using the steepest descent algorithm of 5000 steps followed by 15000 steps of the conjugate gradient algorithm. On the solute molecules of the complexes, positional constraints with a weak force constant of 5 kcal mol-1 Å-2 were imposed. The entire system was subsequently heated under control from 0 to 298.15 K for 50 ps at constant volume (NVT ensemble) using a Langevin thermostat and a canonical ensemble with a collision frequency of 1 ps-1. Because the water molecules added by the tleap tool were loose, a brief heating period was chosen to prevent the formation of vacuum bubbles. Since the barostat used to control pressure is unreliable at low temperatures, the isothermal-isobaric ensemble (NPT) was not used in this case. The solute molecules were constrained throughout the heating process by a harmonic potential of 5 kcal mol-1 2. The Langevin thermostat and Berendsen barostat were then used to complete two rounds of equilibration at the target temperatures and pressures of 300 K and 1.0 atm, respectively. The systems were kept running for 1 ns with a weak constraint of 1 kcal mol^{-2}, a collision frequencyof 2 ps, and a pressure relaxation time of 1 ps. The systems were then equilibrated for 5 ns under identical circumstances. The equilibration stage enables a more natural alignment of atoms and molecules with respect to one another. The treatment of the water molecules involved the SHAKE algorithm. After the system had equilibrated, a productive MD was performed for 50 ns with a timestep of 2 fs. The coordinates were saved for the complexes every 0.1 ps. The trajectories were analyzed using the most populated trajectories from the MD simulations. The root means square deviation (RMSD) and distances between residues were analyzed. The HIERAGGLO algorithm, which is implemented in CPPTRAJ, was then used toperform cluster analysis.

4.3 QM/MM calculations:

QM/MM calculation was performed for rMb(MnPc), which according to experimental evidence, showed a higher reactivity. From the clustering of MD trajectories, sample snapshots were taken from the structures with the highest population. Compared to randomly selecting photos, this selection is statistically more rigorous. The ChemShell[24] software package was used to do the QM/MM calculations. Turbomole[25] was utilized to simulate the quantum mechanical (QM) region. On the other hand, DL_POLY[26] was employed to handle the molecular mechanics (MM) portion, employing the AMBER force field. The QM area

included the manganese-porphycene ring and the ethylbenzene substrate, whereas the active region comprised all of the water and protein residues that were 9 Å of the MnPc ring and the substrate. The electronic embedding approach was used to account for the polarizing impact of the protein's surroundings on the QM region. Hydrogen link atoms employing the charge-shift model was employed to cope with the QM/MM border.

The QM calculations were performed using Unrestricted B3LYP functional. The UB3LYP functional has been observed to exhibit enhanced accuracy in comparison to a diverse range of other functionals when applied to an Mn-oxo system within the context of hydrogen abstraction reactions. To optimize the geometry and for the frequency calculations, we employed the def2-SVP basis set. Calculations of single-point energy were performed using the larger all-electron basis set, def2-TZVP. Two-step transition state optimizations were performed. First, a relaxed potential energy scan was performed along the reaction coordinate (distance between oxygen and hydrogen to be abstracted) with a step size of 0.5 Å from reactant to the intermediate. The local maximum energy point on the potential energy surface (PES) was subsequently subjected to unconstrained optimization utilizing the P-RFO optimizer within the HDLC code. Afterwards, all optimized geometries were subjected to frequency calculations.

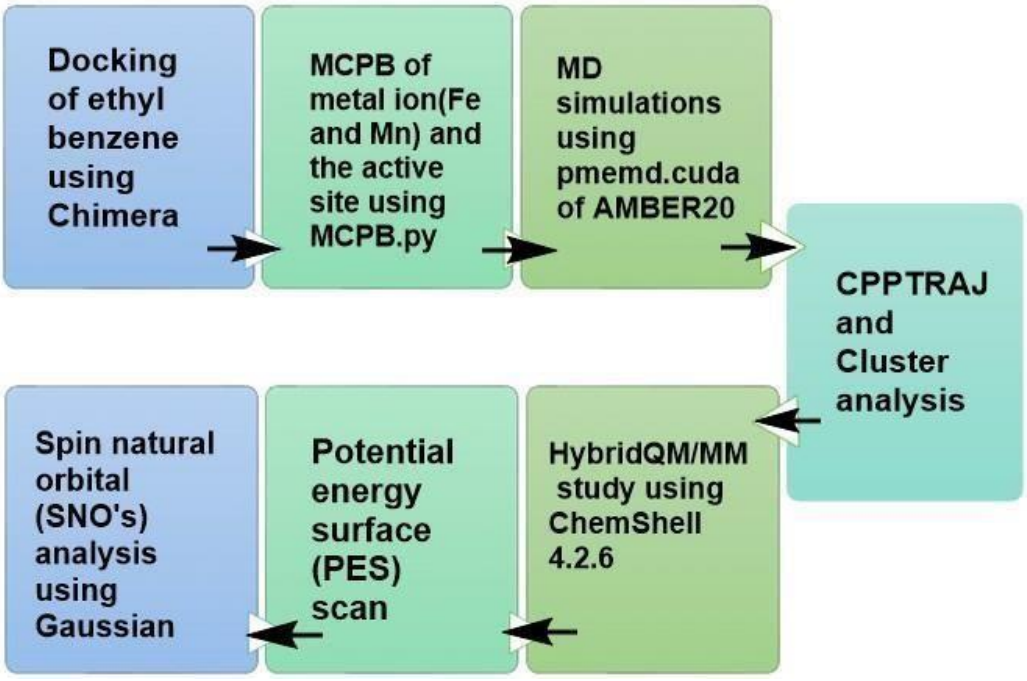


Figure 8: The computational methodology employed in this study is outlined in this flow chart, along with the software packages utilized.

Chapter5: Results and Discussion

5.1: MD simulation results:

An MD simulation was conducted to investigate the stability of substrate at the active binding site for reconstituted myoglobin, i.e., rMb(MnPc) and rMb(FePc). For the rMb(MnPc), it is seen that initially the ethyl benzene position fluctuates between 4 to 5.8 Å from manganese porphycene ring, but after 20 ns the substrate comes nearer to ring with a distance varying from 3.3 to 4.3 Å and become stable. Clustering analysis also confirmed that H7 of ethyl benzene is indeed closer to the Mn=O moiety with most populated snapshot possesses only 2.4 Å distance between H7 and O=Mn. The cluster analysis and most probable structure have been presented in figure 9a and 9b respectively.

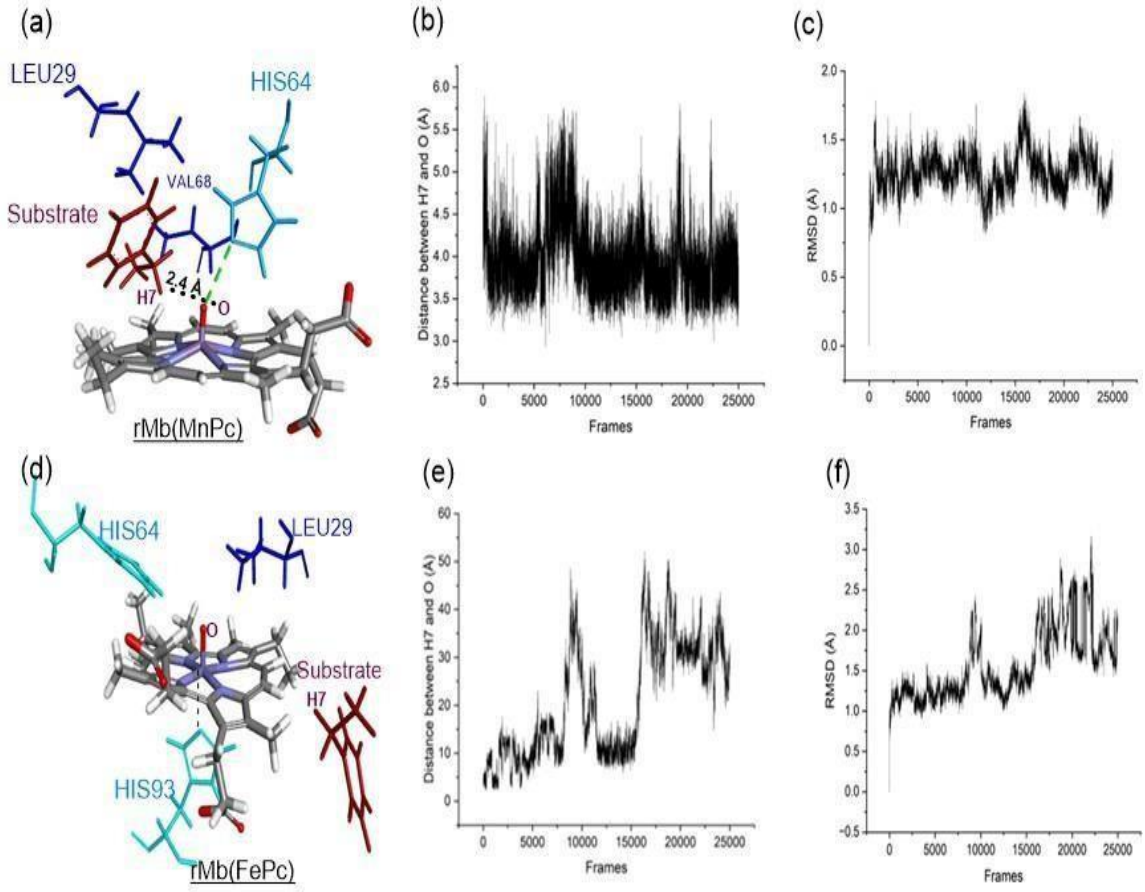


Figure 9: Graph showing the mobility of EB substrate in the active binding site of rMb(MnPc) and rMb(FePc) during the MD simulation. Active site of most populated structures shows key residues interacting with substrate of HIS93 making h-bond with Mn=O moiety of rMb(MnPc). Whereas the substrate was not stable in rMb(FePc), as shown by the RMSD graph and distance fluctuation graph of H7 and O.

This proves that rMb(MnPc) have proper binding site, where substrate is stable and closer to the active CpdI during the entire md simulation of 50 ns. HIS, VAL, LEU interacts with the substrate and HIS64 is seen stabilising the Mn=O moiety by making hydrogen bond with it.

We have done similar investigation for rMb(FePc) where it was found that the substrate goes far away from the protein's active site as shown by the most populated snapshot of the entire trajectory as shown in figure 9d and 9e, respectively. We have also analysed the RMSD to check the stability of the complexes during MD simulation and presented in the figure 9c and 9f for rMb(MnPc) and rMb(FePc), respectively. From the RMSD plot it is clear that EB is quite stable in the pocket of rMb(MnPc) (RMSD is within ~1.5 Å) but it is not so stable in rMb(FePc) (RMSD fluctuates within ~1.5 to 3.2 Å). This proves the experimental observation that rMb(FePc) does not show any reactivity with the ethyl benzene due to its non-suitable active binding site.

It was also experimentally observed that free MnPc is also not reactive towards C-H activation[4]. So, to prove the role of protein matrix we have extended our MD for all four types of reconstituted and wild-type myoglobin as shown in figure 6.

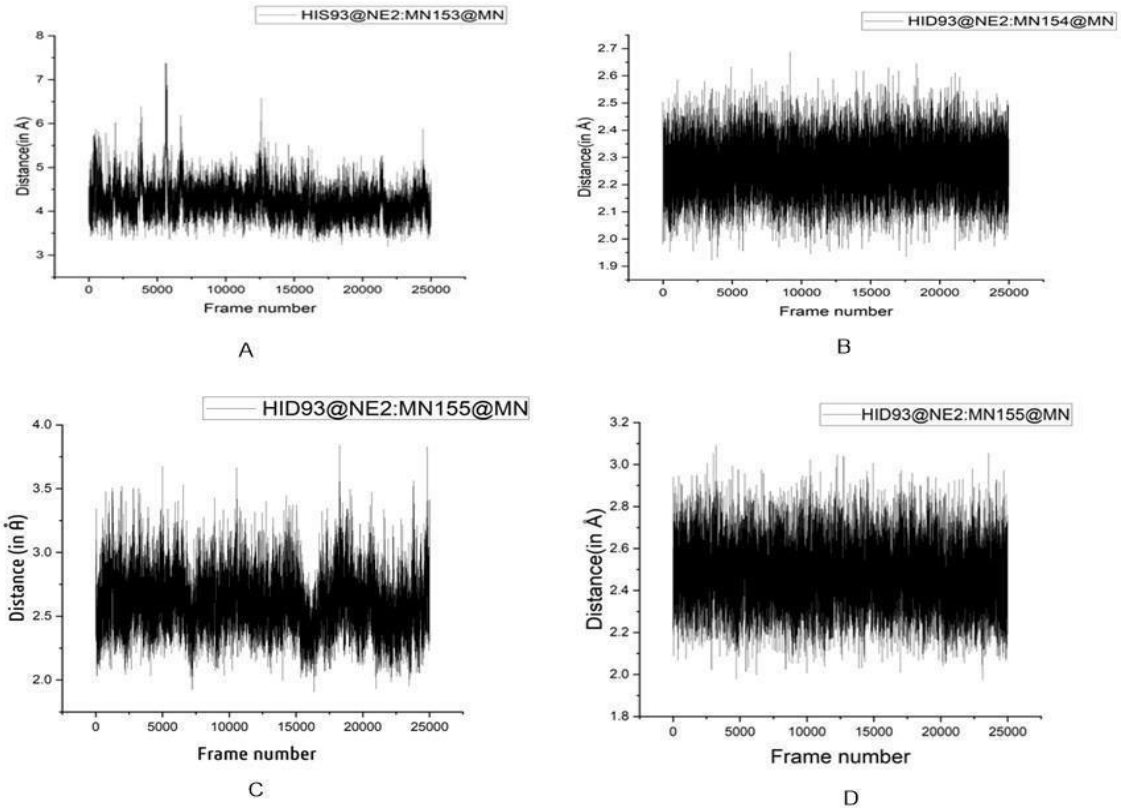


Figure 10: The variation of metal ion and HIS93 distances vs frames for the (A) rMb(MnPc), (B) nMb(FePor), (C) rMb(FePc) and (D) rMb(MnPor).

We have found that the higher reactivity of rMb(MnPc) can be attributed due to the axial ligand effect of HIS93, which was found to be far away from Mn=O during the entire simulation. Distance between HIS93 and Mn largely varies between 3.5 to 6.5 Å. In comparison the axial HIS93 is observed to be intact in case of other three counter-part i.e., nMb(FePor), rMb(FePc) and rMb(MnPor), where distance remains between 3-4 Å as depicted in Figure 10. This can be further confirmed by analyzing the most populated conformations, showing the distance between metal ion and HIS93 as presented in Figure 11.

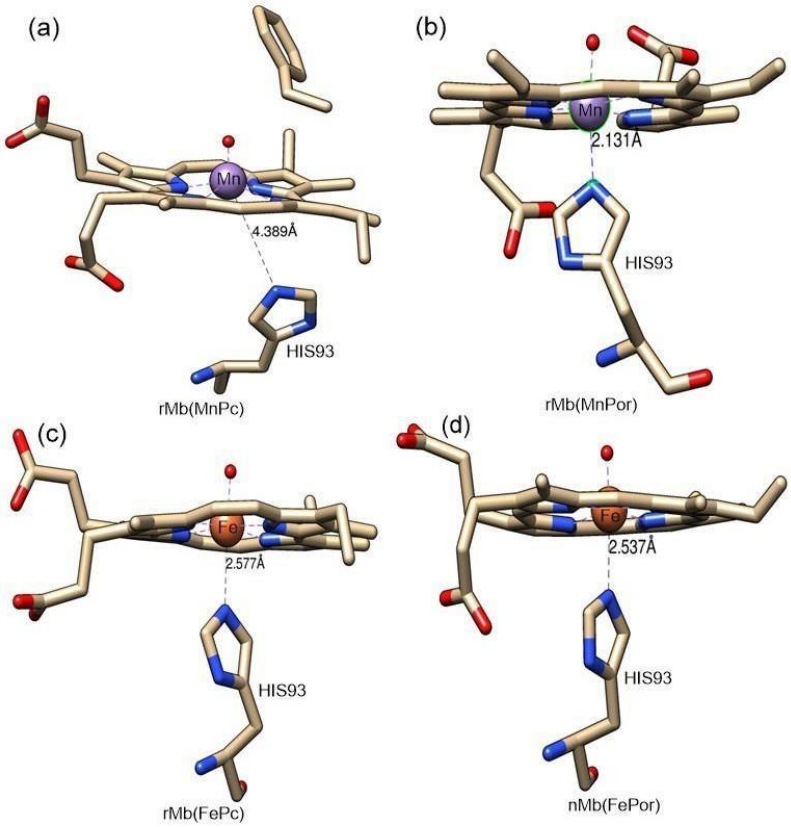


Figure 11: Snapshots from the most populated complexes, showing the distance between the Mn/Fe central metal ion with HIS93 in all the four cases.

5.2: QM/MM calculation results:

We then proceed to check the electronic structure and the reactivity of rMb(MnPc) towards C-H activation using hybrid quantum mechanics/molecular mechanics calculations. The QM/MM calculations were initiated by optimizing the most probable snapshot obtained from the molecular dynamics (MD) simulation. The energy scan was then conducted to investigate the abstraction of the benzylic hydrogen.

In contrast to the experimental reports we have found the Mn(IV)($Pc^{\cdot+}$) state instead of Mn(V) state. This Mn(IV) state resembles with the Fe(IV) state of CpdI of cytochrome p450 which undergoes HAT reaction as shown below in Figure 12.

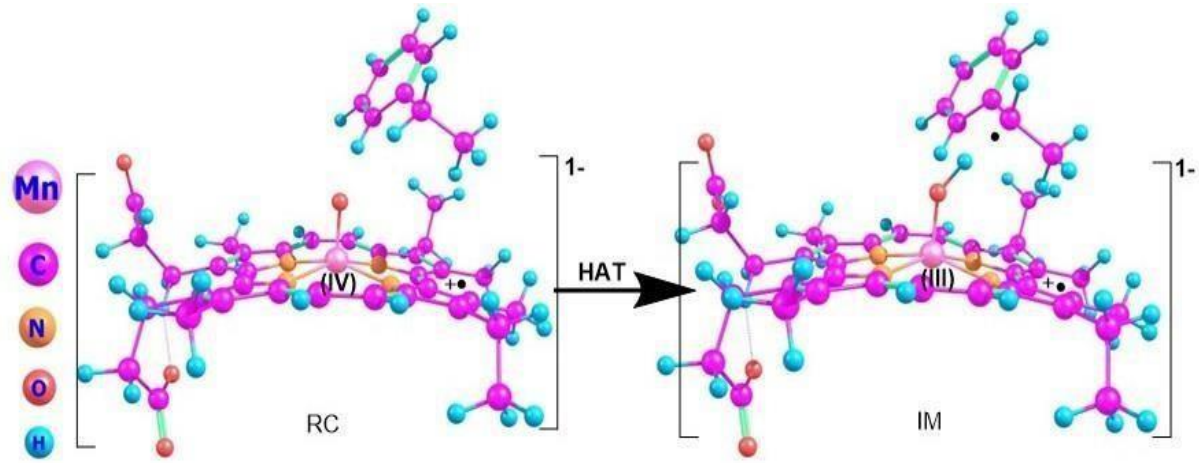


Figure 12: Schematic representation of hydrogen atom abstraction by Oxo-Mn(Pc)(Pro$^{\bullet}$) complex.

$$Mn(IV)=O(Pc^{\cdot+}) + R{-}H \rightarrow Mn(III)\text{-}OH(Pc^{\cdot+}) + R^{\cdot} \qquad (2)$$

A possible pathway for this reaction is shown in Figure 12. In the initial step, the $Mn^{IV(}Pc\cdot+)$ species undergoes a C(sp^3)–H abstraction reaction with ethyl benzene, resulting in the formation of the intermediate referred to as Cpd II. Simultaneously, a radical substrate is generated.

The complex under consideration, i.e. $[Mn(IV)=O(Pc\cdot+)]^{1-}$ can exist in two distinct electronic states, the low-spin singlet state as well as the high-spin triplet state. The energy profile diagram depicting both the spin states and relevant geometries is illustrated in Figure 13.

In the reactant's optimized geometry of the triplet state, the distance between the Mn(IV)=O oxo and hydrogen to be abstracted was 2.46 Å . The angle (Mn-O-H) exhibited a value of 122.7°, consistent with prior investigations proposing the ideal angle for the abstraction of a hydrogen atom during aliphatic or aromatic hydroxylation to be within the range of 110-130°.This range facilitates favorable orbital overlap. The hydrogen abstraction process of the ethyl-benzene C−H bond is facilitated by the transition state ^3TS_H, which exhibits a barrier energy of 12.7 kcal mol^{-1} and 11.4 kcal mol^{-1} at the (def2-SVP + ZPE) and (def2-TZVP +ZPE) levels, respectively.

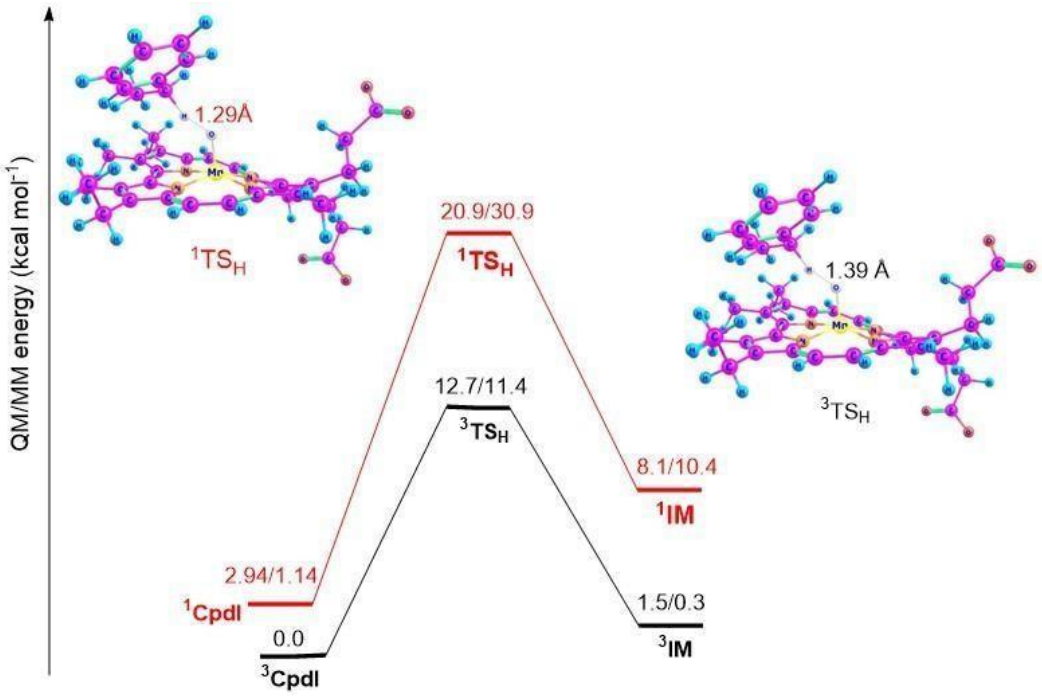


Figure 13: Potential energy profiles of HAT by reconstituted myoglobin with manganese porphycene calculated by hybrid QM/MM level of theory. The QM/MM energy values are given in QM(def2-svp/def2-tzvp) format.

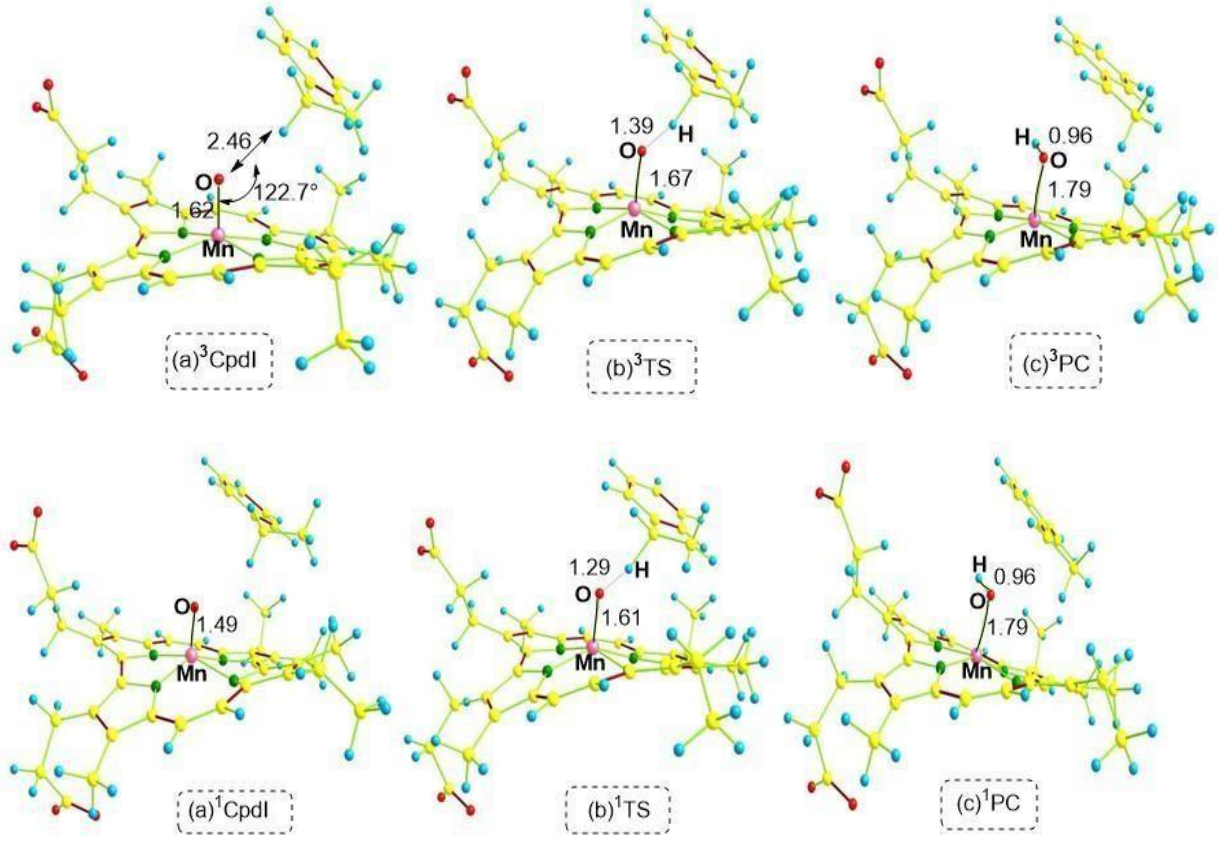


Figure 14: Optimized geometries of the species involved in HAT reaction for both singlet and triplet states calculated using the QM/MM method. Bond lengths are in Ångstroms and dihedral angle (D) is in degrees.

In the optimized transition state (TS), it was observed that the Mn=O bond experienced an increase in length, transitioning from 1.62 Å in the reactant to 1.67 Å in the transition state as shown in Figure 14. The interatomic separation between the carbon atom and the hydrogen atom that will be abstracted has been extended from 1.10 Å to 1.27 Å, and to the oxygen atom bonded to manganese (Mn=O) has been reduced from 2.46 Å to 1.39 Å. Furthermore, the proton abstraction potential energy surface (PES) scan involving the CpdI complex and ethylbenzene is visually depicted in Figure 15. It is evident from the plot that this process involves a high energy cost of 26.7 kcal mol^{-1}. The Mn(III)OH(Pc·+) intermediate thus formed was found to be with endothermicity of 1.5 kcal mol^{-1}, which was lowered to 0.3 kcal mol^{-1} using a higher basis set.

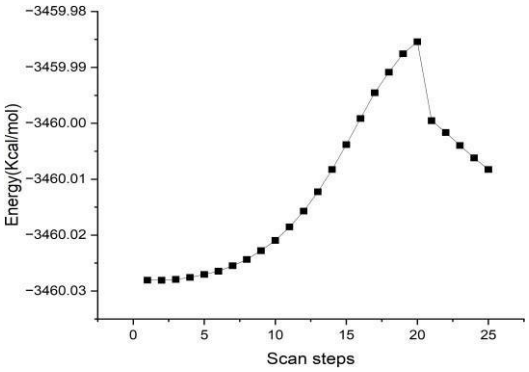
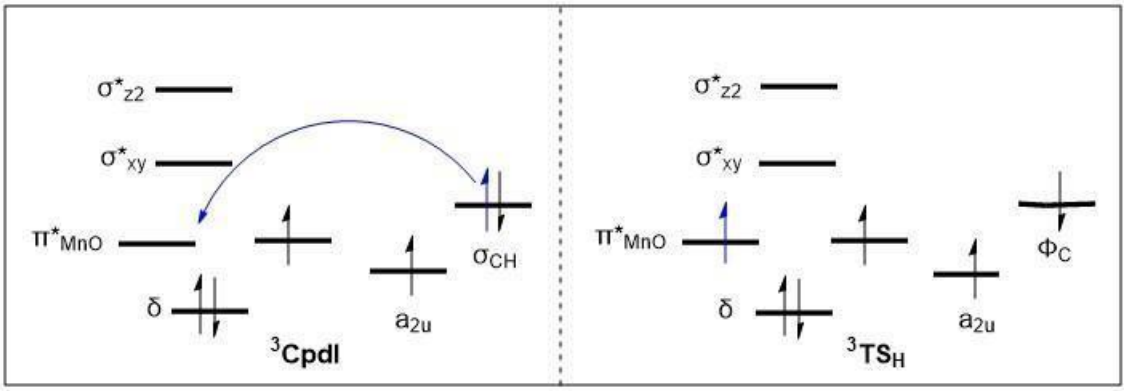


Figure 15: The scanned energy profile for the formation of Mn(III)OH(Pc·+) from CpdI .

The singlet state, $^{1}TS_{H}$ shows very high energy barriers of 20.9 kcal mol^{-1} for the hydrogen atom abstraction. The intermediate formation is more endothermic with 8.1 kcal mol^{-1} energy barrier and less thermodynamically favorable.

The spin natural orbitals (SNO) calculations for ^{3}CpdI and $^{3}TS_{H}$ are shown in Figure 16. An electron initially present in the σ CH bond orbital of ethyl benzene shifts to the unoccupied π*MnO orbital. σ CH is converted to a singly occupied orbital. The observed 1e gain exhibited by the MnO can be attributed to the inherent oxidative characteristics of the hydrogen abstraction phenomenon. This can be understood as a proton-coupled electron transfer mechanism, wherein the abstracted hydrogen atom transfers its electron to the Mn while simultaneously forming a bond with an available lone pair on the oxygen atom.



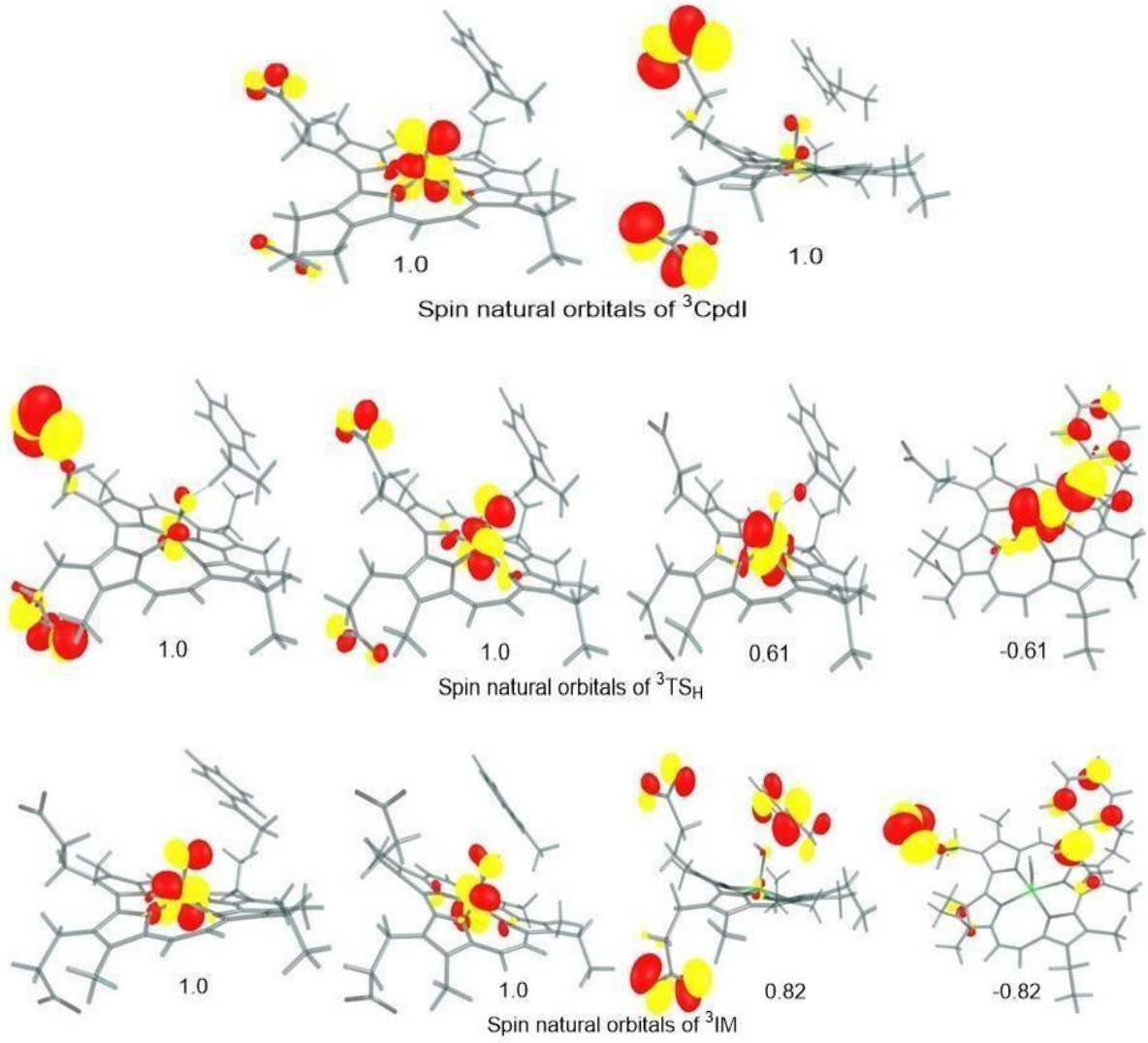


Figure 16: Spin natural orbitals for the active CpdI and its TS_H. The properties of the complex in its triplet ground state were calculated from its electronic structure.

CpdI shows two unpaired identical spin electrons, while TS shows three identical spin electrons and one down spin electron, which is localized on the substrate, as shown in Figure 16.

Chapter 6: Conclusions

Here, we have presented a multiscale computational investigation for the C-H activation reactivity of various myoglobin reconstituted with manganese porphycene, manganese porphyrin and iron porphycene. It was observed that only the reconstituted myoglobin with manganese porphycene i.e., rMb(MnPc) has enhanced reactivity towards ethyl benzene hydroxylation. In this investigation, we have explored the reason behind this typical reactivity pattern using MD and QM/MM techniques. The present study highlighted that rMb(MnPc) has a proper binding site, as the substrate was shown to be stabilized within the ring pocket during the entire simulation. In contrast, we have found that the substrate is unstable and going far away from the active site in the case of rMb(FePc). Further, the effect of axial histidine may also have a role in the reactivity. The Mn=O moiety within the reactive rMb(MnPc) system exhibited a lack of coordination with the HIS93 throughout the MD simulation, proving that the absence of an axial ligand effect can be responsible for its higher reactivity. Our QM/MM analysis shows that modified myoglobin follows a P450-like H atom abstraction process, whereCpd I abstracts the hydrogen from ethyl benzene to form the reactive intermediate Mn(III)- OH(PC). There are two closely lying singlet and triplet spin states, from which triplet state is most reactive as found from the potential energy surface.

Chapter 7: References

(1) Poulos, T. L. Heme Enzyme Structure and Function. *Chem. Rev.* **2014**, *114* (7), 3919–3962.

(2) KARLIN, K. D. Metalloenzymes, Structural Motifs, And Inorganic Models (Vol 261, Pg 701, 1993). *Science.* **1993**, *262* (5139), 1499.

(3) Shook, R. L.; Borovik, A. S. Role of the Secondary Coordination Sphere in Metal-Mediated Dioxygen Activation. *Inorg. Chem.* **2010**, *49* (8), 3646–3660.

(4) Oohora, K.; Kihira, Y.; Mizohata, E.; Inoue, T.; Hayashi, T. C(Sp3)-H Bond Hydroxylation Catalyzed by Myoglobin Reconstituted with Manganese Porphycene. *J. Am. Chem. Soc.* **2013**, *135* (46), 17282–17285.

(5) Van Dyke, B. R.; Saltman, P.; Armstrong, F. A. Control of Myoglobin Electron-Transfer Rates by the Distal (Nonbound) Histidine Residue. *J. Am. Chem. Soc.* **1996**, *118* (14), 3490–3492.

(6) Yonetani, T.; Asakura, T. Studies on Cytochrome c Peroxidase. *J. Biol. Chem.* **1969**, *244* (17), 4580–4588.

(7) Oohora, K.; Hayashi, T. Myoglobins Engineered with Artificial Cofactors Serve as Artificial Metalloenzymes and Models of Natural Enzymes. *Dalt. Trans.* **2021**, *50* (6), 1940–1949.

(8) Oohora, K.; Onoda, A.; Hayashi, T. Hemoproteins Reconstituted with Artificial Metal Complexes as Biohybrid Catalysts. *Acc. Chem. Res.* **2019**, *52* (4), 945–954.

(9) Hayashi, T.; Hisaeda, Y. New Functionalization of Myoglobin by Chemical Modification of Heme-Propionates. *Acc. Chem. Res.* **2002**, *35* (1), 35–43.

(10) Oohora, K.; Meichin, H.; Zhao, L.; Wolf, M. W.; Nakayama, A.; Hasegawa, J. Y.; Lehnert, N.; Hayashi, T. Catalytic Cyclopropanation by Myoglobin Reconstituted with Iron Porphycene: Acceleration of Catalysis Due to Rapid Formation of the Carbene Species. *J. Am. Chem. Soc.* **2017**, *139* (48), 17265–17268.

(11) Hayashi, T.; Dejima, H.; Matsuo, T.; Sato, H.; Murata, D.; Hisaeda, Y. Ja0265052.Pdf. *Si* **2002**, 11226–11227.

(12) Bernard, C.; Gisselbrecht, J. P.; Gross, M.; Vogel, E.; Lausmann, M. Redox Properties of Porphycenes and Metalloporphycenes. A Comparison with Porphyrins. *Inorg. Chem.* **1994**, *33* (11), 2393–2401.

(13) Oohora, K.; Meichin, H.; Kihira, Y.; Sugimoto, H.; Shiro, Y.; Hayashi, T. Manganese(V) Porphycene Complex Responsible for Inert C-H Bond Hydroxylation in a Myoglobin Matrix. *J. Am. Chem. Soc.* **2017**, *139* (51), 18460–18463.

(14) Shaik, S.; Kumar, D.; de Visser, S. P.; Altun, A.; Thiel, W. *Theoretical Perspective on the Structure and Mechanism of Cytochrome P450 Enzymes*; 2005; Vol. 105.

(15) Rittle, J.; Green, M. T. Cytochrome P450 Compound I: Capture, Ch. *Science.* **2010**, *330* (6006), 933–937.

(16) Stone, K. L.; Hua, J.; Choudhry, H. Manganese-Substituted Myoglobin: Characterization and Reactivity of an Oxidizing Intermediate towards a Weak C-H Bond. *Inorganics* **2015**, *3* (2), 219–229.

(17) Cai, Y. B.; Li, X. H.; Jing, J.; Zhang, J. L. Effect of Distal Histidines on Hydrogen Peroxide Activation by Manganese Reconstituted Myoglobin. *Metallomics* **2013**, *5* (7), 828–835.

(18) Mesbahi, E.; Safari, N.; Gheidi, M. Investigation of Axial Ligand Effects on Catalytic Activity of Manganese Porphyrin, Evidence for the Importance of Hydrogen Bonding in Cytochrome-P450 Model Reactions. *J. Porphyr. Phthalocyanines* **2014**, *18* (5), 354–365.

(19) Hollingsworth, S. A.; Dror, R. O. Molecular Dynamics Simulation for All. *Neuron* **2018**, *99* (6), 1129–1143.

(20) Li, P.; Merz, K. M. MCPB.Py: A Python Based Metal Center Parameter Builder. *J. Chem. Inf. Model.* **2016**, *56* (4), 599–604.

(21) Quesne, M. G.; Silveri, F.; de Leeuw, N. H.; Catlow, C. R. A. Advances in Sustainable Catalysis: A Computational Perspective. *Front. Chem.* **2019**, *7* (APR), 1–23.

(22) Sholl, D. S.; Steckel, J. A. Density Functional Theory: A Practical Introduction. *Density Funct. Theory A Pract. Introd.* **2009**, 1–238.

(23) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97* (40), 10269–10280.

(24) Metz, S.; Kästner, J.; Sokol, A. A.; Keal, T. W.; Sherwood, P. ChemShell-a Modular Software Package for QM/MM Simulations. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4* (2), 101–110.

(25) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. Electronic Structure Calculations on Workstation Computers: The Program System Turbomole. *Chem. Phys. Lett.* **1989**, *162* (3), 165–169.

(26) Smith, W.; Forester, T. R. DL-POLY-2.0: A General-Purpose Parallel Molecular Dynamics Simulation Package. *J. Mol. Graph.* **1996**, *14* (3), 136–141.

Thesis

ORIGINALITY REPORT

8%	4%	6%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	pubs.rsc.org Internet Source	1%
2	www.escholar.manchester.ac.uk Internet Source	1%
3	Vandana Kardam, Surajit Kalita, Kshatresh Dutta Dubey. "Computations reveal a crucial role of an aromatic dyad in the catalytic function of plant cytochrome P450 mint superfamily", Journal of Inorganic Biochemistry, 2022 Publication	1%
4	Oohora, Koji, Yushi Kihira, Eiichi Mizohata, Tsuyoshi Inoue, and Takashi Hayashi. "C(sp3)–H Bond Hydroxylation Catalyzed by Myoglobin Reconstituted with Manganese Porphycene", Journal of the American Chemical Society, 2013. Publication	<1%
5	"Biomolecular Simulations", Springer Science and Business Media LLC, 2013 Publication	<1%

Dmondal
24/07/23