AN INVESTIGATION INTO BIODEGRADABILITY OF IONIC LIQUIDS

EMPLOYING ACTIVE SITE MODELS

By

ATIYA BANERJEE

Bachelor of Technology in Biochemical Engineering
Harcourt Butler Technological University
Kanpur, Uttar Pradesh, India
2013

Master of Technology in Chemical Engineering
(Computer Aided Process Plant Design)
Indian Institute of Technology
Roorkee, Uttarakhand,India
2015

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
July, 2020

AN INVESTIGATION INTO BIODEGRADABILITY OF IONIC LIQUIDS

EMPLOYING ACTIVE SITE MODELS

Dissertation Approved:

Jindal K. Shah

Dissertation Adviser

Heather Fahlenkamp

Marimuthu Andiappan

Mario Borunda

ACKNOWLEDGEMENTS

I want to thank firstly the almighty god for his role in formulating the inner workings of the universe and giving the author the gift called life and the opportunity to live it. We live in a blanket of reality and often overlook the spiritual foundations of our being and our existence on this planet.

The author wishes to thank his family and parents that have been a constant source of strength, care, love and a collection of emotions that words cannot clearly do enough justice. Their prayers and well wishes at every step of my journey in life is beyond any composition or collection of phrases that any language might be able to grant to its practitioner.

I express my deepest gratitude to my advisor, Dr. Jindal Shah, Associate Professor, Chemical Engineering for being a constant source of support and being helpful through every part of the work performed as part of the doctoral degree. I express my sincere thanks to him for granting me an opportunity to work with him on his CAREER project that was unique in many ways and shaped my foundation and thought process during the work. My various discussions on research, teaching, academia and other wise with him were extremely helpful for me in order to develop my professionalism. The day to day interactions with Dr. Shah were invaluable in terms of developing research acumen as well as methodology for the me. He has been a big part of my success and development here at Oklahoma State University and career in these four and a half years or so. I specially thank him for his encouragement, his cooperation and his patience in advising me throughout my degree and being a role model for me to follow in each step of the way.

Acknowledgements reflect the views of the author and are not endorsed by committee members or Oklahoma State University

I also thank my committee members, Dr. Heather Fahlenkamp, Professor, Chemical Engineering, Dr. Marimuthu Andiappan, Assistant Professor, Chemical Engineering and Dr. Mario Borunda, Associate Professor, Physics for their constant support, encouragement, thoughtful questions and suggestions as part of my work and kindly serving on my dissertation committee. Their role as scholars in evaluating my research and its broader impact beyond molecular simulations has been extremely helpful for me to compose the general parts of the thesis.

I thank all the group members, both past and present of the Shah research group that granted many helpful discussions, both technical and informal, being supportive during events, conferences, seminars and giving me a chance to share my expertise with them. Their well wishes and company was particularly special to me during my time here at Oklahoma State University.

I thank all my fellow Graduate and Undergraduate Students in the Department of Chemical Engineering for sharing their respect, love and experiences with me during PhD. I specially remember my tenure as a teaching assistant to the undergraduates and have fond memories of guiding them towards their respective courses, and careers in the long run.

I wish to acknowledge the Oklahoma State High Performance Computing Center for kindly allocating the resources for carrying out all of the calculations reported in the work. Special thanks and mention to Jesse Schafer for being really enthusiastic and responsive in resolving issues and sharing technical knowhow and Evan Linde for his help in developing some of the scripts used in the development of the thesis work.

Acknowledgements reflect the views of the author and are not endorsed by committee members or Oklahoma State University

I thank the wonderfully supportive and friendly staff at the Edmon Low Library that were instrumental in providing access to various resources and references related to research in a timely manner that helped me compose this research.

My journey towards my PhD degree has been graced by a multitude of individuals that hold a very special place in my heart. Although it is a very challenging task to pin point experiences, but I would like to mention some very unique names, whose company was particularly helpful to me as a student coming into a new country and absorbing a new culture. Will, who has always remained a rock and like a father figure to me in so many different times and circumstances. My friend Suman and Ankit that have been like brothers to me. Fellow members in the various organizations that I was fortunate to participate in, namely, Oklahoma State Secular Organization, Malaysian 24 Drum Troupe and Indian Student Association during my time here. I thank my landlord, Mr. Jim Brown, for being so patient and cooperative with me during my stay at my address here for four years in Stillwater. I finally would like to thank Abby for being really supportive and graceful to me during the final stages of my dissertation. I appreciate each and everyone for their help towards making this endeavor and degree work successful and complete in its utmost sense.

Acknowledgements reflect the views of the author and are not endorsed by committee members or Oklahoma State University

Name: ATIYA BANERJEE

Date of Degree: JULY, 2020

Title of Study: AN INVESTIGATION INTO BIODEGRADABILITY OF IONIC LIQUIDS EMPLOYING ACTIVE SITE MODELS

Major Field : CHEMICAL ENGINEERING

Abstract: Ionic liquids have the potential for driving industrial processes in a sustainable fashion due to their environmentally benign characteristics like inherent low vapor pressure and flammability. Their extremely low vapor pressures translate to the fact that they have negligible role in air emissions as compared to conventional industrial solvents. Ionic liquids can also be flexibly designed in order to obtain desirable physical properties. This design could be carried out by choosing from thousands of possible combinations of cations and anions. Out of the myriad of a wide range of ionic liquid classes, imidazolium-based ionic liquids have been the most successfully utilized in diverse applications. Suitable variants of this type of ionic liquid have been applied in catalysis, extractive distillation, liquid-liquid extraction among many others. Though being proved to be highly effective in driving complex industrial processes, experimental investigations have raised questions on their environmental degradability. Thus, it becomes imperative to include rational design into their synthesis. From an experimental point of view, considerable efforts have been put in this direction but molecular level details have not been explored in detail computationally. The present work aims to provide physical insight into the phenomena of ionic liquid biodegradability to aid in the future design of these solvents.

Cytochrome P-450 has been identified and widely studied for their role in oxidation of a wide variety of molecules in aerobic and anaerobic environments. Thus, to develop a computational framework related to biodegradability, it was considered important to capture the effects of the cytochrome on imidazolium-based ($[C_n mim]^+$) and other cations for which biodegradability has been studied. For this, the enzymatic center of the P-450 molecule (heme) was modeled as an iron porphyrin molecule and enhancing the complexity of the active site at every step. The $[C_n mim]^+$ was included in the model as a potential substrate for the P-450 enzyme in complexation with the heme receptor. This interaction was firstly modeled using DFT calculations by adapting a purely quantum mechanical framework to provide insight into geometrical and electronic properties of the complex arising from different conformations. In the next step, the entire catalytic cycle of the P450 was treated using a synchronous approach using pure DFT modeling technique. To include favorable conformations in the presence of the protein environment, docking calculations were performed that elucidated the plasticity of the binding pocket. The docked structures were then subjected to QMMM calculations and the resulting geometries were analysed.

TABLE OF CONTENTS

Chapter Page

Chapter Page

Chapter Page

Chapter Page

Chapter Page

LIST OF TABLES

Table Page

Table Page

Table Page

LIST OF FIGURES

Figure **Page**

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

CHAPTER I

INTRODUCTION

1.1 Overview

One of the most necessary tasks for mankind is to strive for sustainable growth and development [13]. The advent of the modern age comes with a unique set of challenges that prove detrimental to this philosophy. As such, the scientific community all over the globe is engaged in creating design practices that produce environmentally benign processes and materials. Chemical processes have an intimate relationship with the global economic state [14, 15, 16] and have been acting as a vital contributing factors to the progress of mankind in the last 150 years [17, 18]. One of the central tasks to sustainable chemistry is creating ecologically friendly products that involves combination of various design principles [19, 20]. Boethling et al. [21] suggests some of these principles in his seminal work titled "Designing Small Molecules for Biodegradability" with regards to Green Chemistry and pollution management. In his work, he states that 'inherently safe design' is one of the critical facets in establishing green processes and design principles have to percolate down to the most basic levels to include molecules.

Prior investigations on the effect of industrial processes [22] have shown that less persistent chemicals on the downstream side of the process make them green. This persistence is directly linked to the natural degradation or biodegradability of the substance. Out of the salient features of safe design, biodegradability is a key factor whose inclusion is critical to the success of the overall design process [23]. Figure 1.1 gives a broad description of the life cycle of a sustainable design process. The first step indicates the analysis of the chemical structures which is followed by their synthesis. From a structural perspective, the presence of chemical moieties and entities susceptible to enzymatic hydrolysis, especially oxygen containing, promote overall biodegradability of the molecule. As such, esters, hydroxylated phenyls and even cyclic hydrocarbons having oxygenated substituents have been shown to be highly biodegradable. Due to this, structural modifications to small molecules have been a highly popular area of research in the last two decades [24].

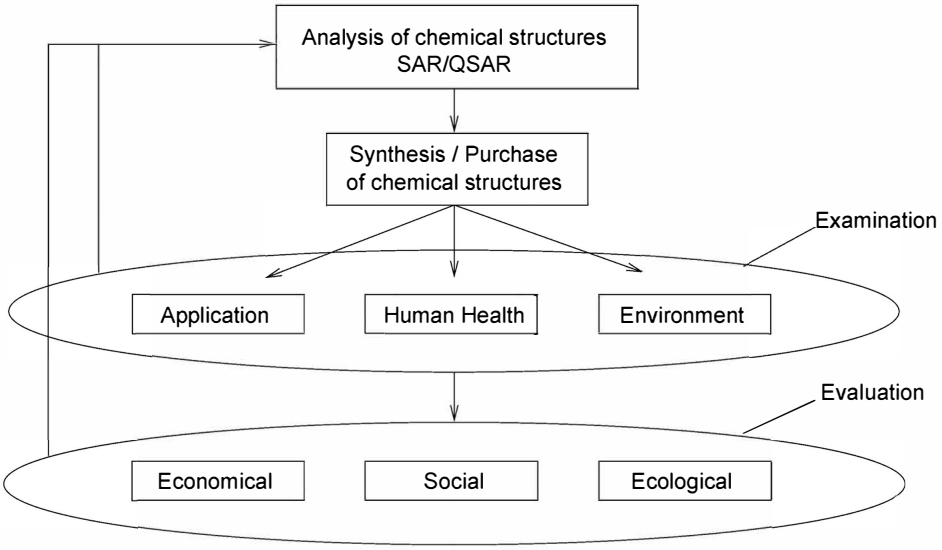


Figure 1.1: Sustainable chemical design schematic adapted from [1]

More often, structure activity relationships have been described for a wide range of these molecules in order to understand their potential for environmental persistence. In terms of structure-activity relationships, the interaction of intermediate reaction products and metabolites with biological macromolecules have a direct effect on the overall reactivity of the parent molecule. The biodegradability of a given substance is a highly involved property that is a function of its biological activity, oxidation potential and mobility among many others. Thus, incorporating biodegradability into the rational design process becomes a multi-lateral problem that requires efforts on many different fronts or disciplines. Figure 1.2 shows the importance of the chemical structure of an ionic liquid (IL) and the surrounding aspects of it pertinent to their rational design. Out of these, the first step is defined as 'molecular interactions' which is followed by 'interactions with biologically relevant macromolecules', both of which are key parts in the present thesis work. The small molecules considered in this work are sub-entities of ILs that have been described in the following section.

1.2 Ionic Liquids

Before delving into the description of ionic liquids, it is appropriate to understand volatile solvents. Solvents in general are environmentally damaging chemicals because of their persistence and volatility. These contribute to air pollution and their use as reaction media for different processes is inevitable. For this reason, they need to be designed in a manner which ensures reduction of harmful emissions of volatile organic compounds (VOCs) and carbon dioxide (CO_2). Among all possible solvents, ILs have emerged as a potentially 'green' choice and attention towards them has enhanced manifold in recent years. Ionic liquids (ILs) are liquid electrolytes that are entirely composed of ions [25] with melting points below 100°C. They have received wide attention from the scientific community as innovative fluids [26] due to their diverse suite of behaviors. Numerous combination of cations and anions can be established for tailoring their physical properties, that

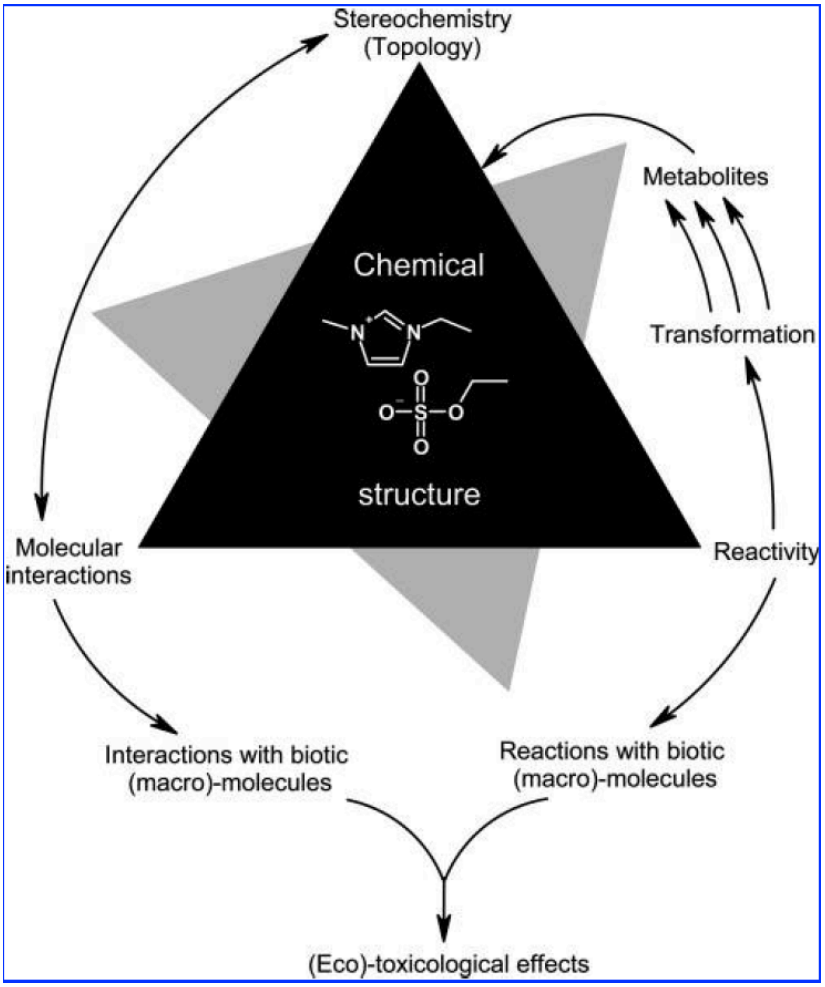


Figure 1.2: Structural Design of Ionic Liquids adapted from [1]

have led to a very emerging interdisciplinary field involving chemistry, chemical engineering, material and environmental sciences. Due to their growing chemical variety [27] over the years, they have been further divided into various classes [28] such as task specific ILs (TSILs), room temperature ILs (RTILS), and many others. Structurally speaking, IL cations are generally structures of low symmetry [29] often stabilized by resonance, involving a charged center of nitrogen or phosphorus. Further, these cations may be of heterocyclic, functionalized or chiral nature, giving rise to a plethora of synthesis options and configurations. Some of the common types of cations are namely, phosphonium, imidazolium, pyridinium, ammonium, sulphonium, amongst many others.

On the other hand, IL anions that are integral to RTILs are mostly weak organic or inorganic chemical entities of the basic nature. During the last decade, various types of functionalized ionic liquids [30] have been designed for serving specific purposes [31] and a precursor for advanced materials such as nanostructures. Ionic liquids have been employed in a wide variety of industrial processes including catalysis, pharmaceutics [32], gas absorption [33] and liquid-liquid extraction. Their extremely low vapor pressure [34, 35] and tunability have projected them as likely 'greener' candidates to conventional solvents. Although, that being true, recent literature reports [36, 37, 38] that these novel solvents, are neither intrinsically non-toxic, and their large scale industrial usage might lead to negative environmental impact. Various aspects of their biodegradability [39, 40] and toxicity have been addressed in recent works, involving detailed characterization, their impact on various media, and their structure-activity [41, 42] relationships. The overarching conclusion from these scientific studies point to the fact that structural modifications to the ionic liquid moiety and the presence of suitable functional groups [43, 40] promote their biodegradation. Importantly, the molecular design of the IL cation often dictates the degradability of the overall fluid [44] and directly influences mechanism behind this phenomena.

1.2.1 Background

The reason for the rise in popularity of ionic liquids can be attributed to their physical properties [45, 46]. These are solids that exist as liquid at room temperature and are entirely made up of ionic entities. Among the various useful properties of these chemicals, some of the most important have been discussed for reference. Their extremely low volatility corresponds to usability in low pressure environments. Also, high thermal stability means that carrying out reactions at high temperature ranges is quite feasible. From a reactive point of view, their polar and noncoordinating nature translates to a very high potential for site-selectivity and stereochemically specific chemical processes. Molecular interaction of ILs amongst themselves and other members in the overall system exhibit their versatility and

complexity. They have been shown to be capable of interacting via a plethora of schemes including intra and inter molecular H bonding, dipolar, selectively electrostatic, hydrophobic, dispersive amongst others. Due to the various kinds of possible interaction schemes, they serve very well as a conducive medium to solubilize harmful greenhouse gases such as CO, CO_2 and CH_4. Due to their high ionic character, they can enhance organic synthesis reaction rates for processes such as polymerization. Their solubility depends on the nature of ions combining to form the ionic liquid. Their synthesis can be fine-tuned to produce desirable physicochemical properties with emphasis on the choice of ionic entities employed. The number of available combinations for producing Ionic liquids suggests that they occupy a highly variable, complex and detailed 'chemical space' which has been exploited for various applications. A graphical representation of some of this space as provided in [3] has been given in the Figure 1.4. The types of ions range from cyclic to non-cyclic, homocyclic to heterocyclic, halogenic to non-halogenic having various kinds of backbones and chemical organization. As mentioned in the previous section, these structural features are the central piece to the study of their biodegradability.

1.2.2 Types of Ionic Liquids

In the last two decades, there has been a steady increase in the attention towards ionic liquids. Due to this, they have been synthesized, functionalized and employed in a variety of applications in the industry. Due to the scope of the present work, the focus towards some of these applications in the chemical industry will be described in this section. Figure 1.3 from the work of Plechkova and Seddon [2] gives an excellent overview of the potential applications of ILs in different concentrations within production and engineering. Several comprehensive review articles have been published on the synthesis and applications of room temperature ionic liquids (RTILs). These have focused on catalysis, biocatalytic transformations, crystallization, voltammetry among many others. Among the wide variety of possible ILs, the ones consisting of N-heterocyclic cations are the most widely used. Salts containing imidazolium and pyridinium head groups are some of the most

prominent out of them. They are often structurally modified to obtain suitable solvent properties. These modifications are performed on the ring head as well as counter ions. Based on the application, the behavior of the ions can be tuned by chemically altering their identity. Thus, several functionalized and derived ILs have been produced using heterocyclic ring heads for employing them in CO_2 capture, stabilizing and activating enzymes, hydrocarbon synthesis and promoting polarity in a reactive medium among others. Other less popular types are pyrrolidinium, ammonium, phosphonium, sulfonium and thiazolium out of the estimated 10^6 possible combinations available. Each of these types have been successfully synthesized, employed and studied in literature. Much of the former investigations are dedicated to the role of geometry of the ionic entities within the IL and its effect on the function of the liquid. Hence, taking into consideration the scope of the present work, some of the geometrical properties of N-heterocyclic cation-containing ILs will be discussed in the next section.

The following section would focus on the different types of ionic liquids and their applications.

1.2.3 Structures of Ionic Liquids

Almost in all cases, an ionic liquid has a cationic and anionic counterpart. These parts makeup the whole nature and set of unique properties for an ionic liquid. Taking that into consideration, the structure of ILs has been discussed in brief in this section.

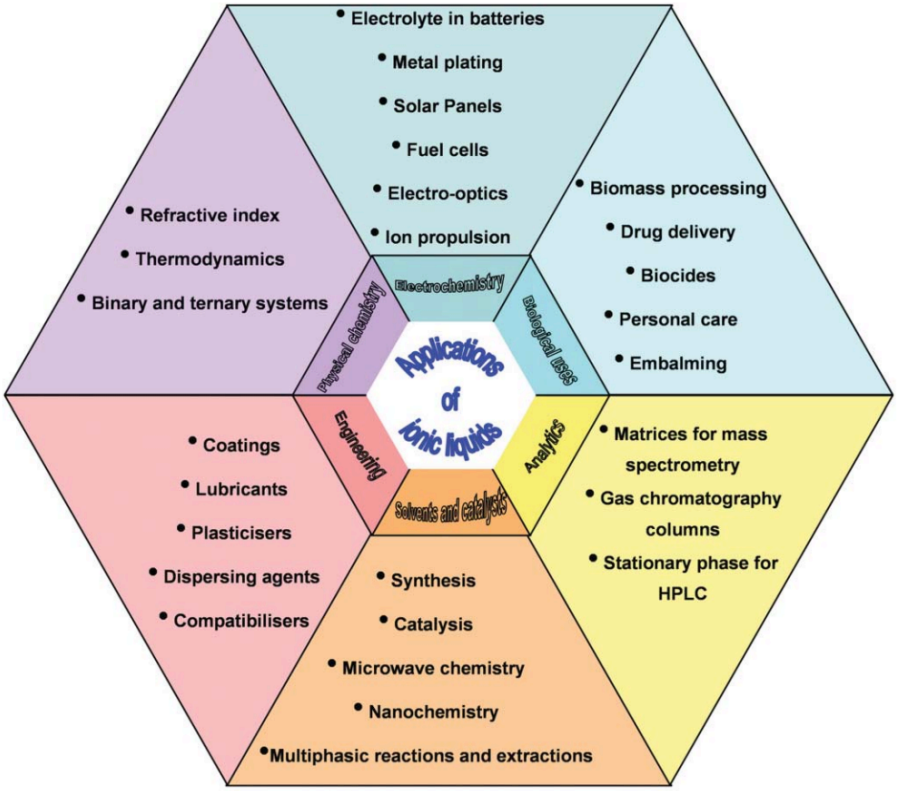


Figure 1.3: Various industrially relevant applications of room temperature ionic liquids (RTILs) from [2]

1.2.3.1 Cations

The most common class of IL studied in literature are ones having an imidazolium-based cation. They can be directly synthesized by alkylating a suitable N-alkylimidazole, followed by appropriate metathesis to incorporate the desired anion. Modifications to the cations most commonly include substitutions to the attached alkyl chain, linking various moieties to the aromatic head and the ring. Property wise, the glass transition temperatures of imidazolium cation containing ILs have been shown to not have any identifiable correlation with the length of the N-alkyl chain. Viscosity of these ILs increase with the length of the chain and so does the molar volume and charge density. Addition of a hydroxyl (-OH) group to the side chain enhances all of the above properties, due to the promotion in the ability of the cation to participate in hydrogen bonding with other groups of the ionic liquid. Introduction of electrophilic groups to the ring head result in higher densities and

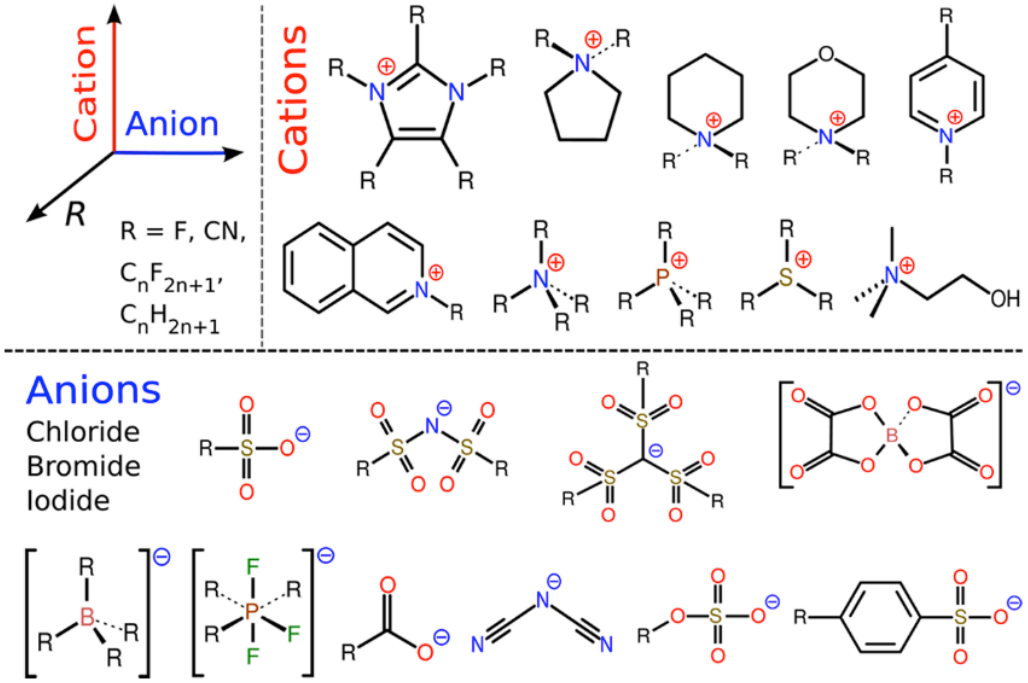
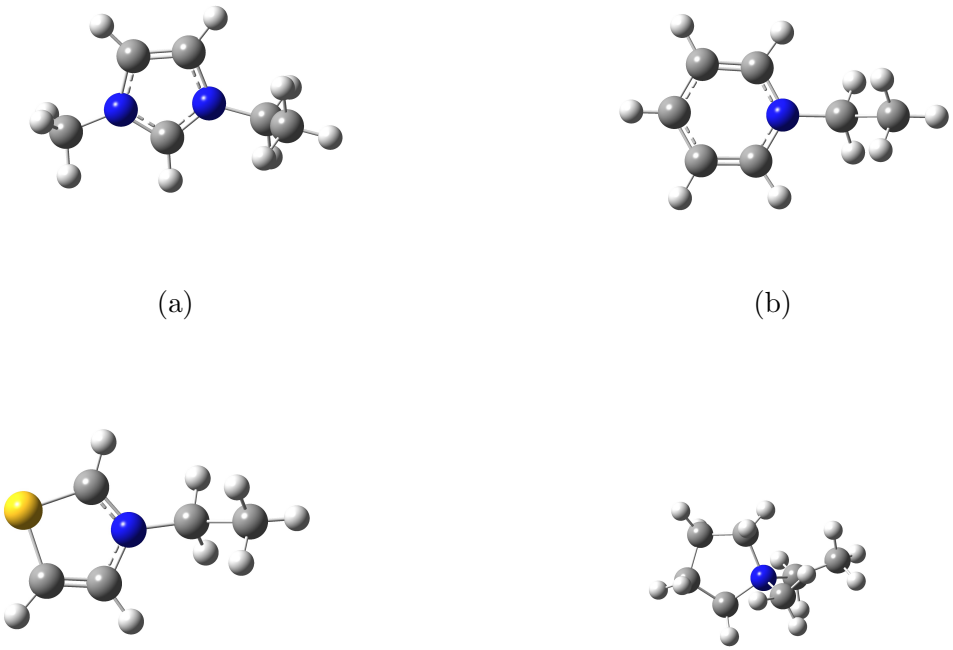


Figure 1.4: Some possible functional groups in ionic liquids adapted from [3]

viscosities.

The pyridinium family of ILs have been synthesized at a lower cost than imidazolium ones and have enhanced thermal stability. They are synthesized through alkylation of pyridine followed by anion metathesis. Having a higher density than imidazolium ILs, they have a better packing efficiency. Due to less delocalization of the electronic charge placed on them, they exhibit a higher viscosity. Increase in the alkyl chain of the cation results in reduced surface tension possibly due to increased electrostatic forces between the separate ionic entities. In terms of modifications, placement of the methyl group is significant in determining their viscosity. Enhancing the length of the side chain serves to decrease the density and promote their viscous nature. Addition of e^- withdrawing substituents on the pyridinium ring have also shown to affect the physicochemical properties dramatically. Similar to imidazolium ILs, hydroxylation of aliphatic chain connected to the ring enhances hydrogen bonding interactions among ionic members.



(a) (b)

(c) (d)

Figure 1.5: Some ionic liquid cations (a) 1-ethyl-3-methylimidazolium ($[C_2mim]^+$) (b) 1-ethylpyridinium ($[C_2py]^+$) (c) 1-ethylthiazolium ($[C_2th]^+$) (d) 1-ethyl-1-methylimidazolium ($[C_2pyrr]^+$)

1.2.3.2 Research Focus

The performed research is focused towards understanding the geometrical and electronic aspects of ionic liquids in a biological environment to unravel the physical insight into their overall oxidative ability. This oxidation is key to their degradation in a suitable mode to enable their large scale industrial usage and process control. To predict the reaction characteristics of these ionic liquid molecules, a purely computational approach is undertaken in the research. Major emphasis of the research has been dedicated to the development of a model that could mimic the activity of the binding pocket of oxidizing protein and its effect on the ionic liquid molecules.

The key focus areas of the stated research are the following:

(1) Elucidating the geometrical and electronic properties of active site model by

including the ionic liquid cation as the substrate in different electronic environments.
(2) Enhancing the description of the active site model at every step to achieve more of the physical aspects of the actual protein environment and isolating the effect of additions at all of these steps.
(3) Description of the thermodynamics process by performing a variety of analysis that highlight the key descriptors of the process and quantities that illuminate its characteristics.

All of these focus areas are expected to provide insight into molecular level details so that suitable structural motifs can be induced into the substrate as well as the protein to enhance the biodegradation potential of the former entity.

1.2.3.3 Objectives

Different objectives of the research are related to the understanding of the systems' thermodynamic aspects at every step upon addition of progressive elements to the active site for developing the model. The design of these objectives is performed by taking into consideration the size and complexity of the model and then evolving the objective to meet the research tasks.
The most important objectives of the research work are the following:
(1) To understand the conformational effects of the ionic liquid cation in the presence of the active site molecule of the protein by evaluating binding affinity using electronic structure calculations.
(2) Infer the electronic effect of the presence of cation on the protein both gas and condensed phase environments by calculating reactivity indexes using conceptual density functional theory approach.
(3) Evaluate the effects of the protein environment by including the amino acid residues pertinent to the active site embedded within the protein and evaluating free energies of binding and other pertinent quantities.

All of these objectives are expected to form a design principle at the in silico level for screening suitable structures of both the protein and substrate for making the ionic liquid more biodegradable.

1.2.3.4 Potential Environmental Benefits

The ability to produce biodegradable ionic liquids will enhance the friendliness of a lot of industrial processes that employ ionic liquids due to their desirable physical properties. As these solvents are have inherent design flexibility, biodegradability can be included as one of the design principles among others for them to be used to drive complex processes. Potential benefits to environment would include reduced water pollution, soil pollution and cost-effectiveness at the downstream side of the plant/industry. Other foreseeable effects of the research work are as follows:

(1) The increased use of safer ionic liquids would ensure a shift from conventional solvents to friendlier alternatives that would lead to lower environmental footprint.
(2) Use of low volatile solvent will contribute to the reduction in CO_2 emissions that is one of the most challenging problems faced by the globe today.
(3) Integration of the design from a molecular to industrial level would ensure that at each step the design has been evaluated for risks and strategies have been put in place to ensure swift retrofitting and re-design across levels.
(4) An inverted pyramid approach takes care of the fact that the screening is conducted at the most basic level and all of its effects can be translated easily to other levels.
(5) The research work aims to integrate computational and experimental work to create a robust rational design approach that can be used for designing other environmentally friendly materials.

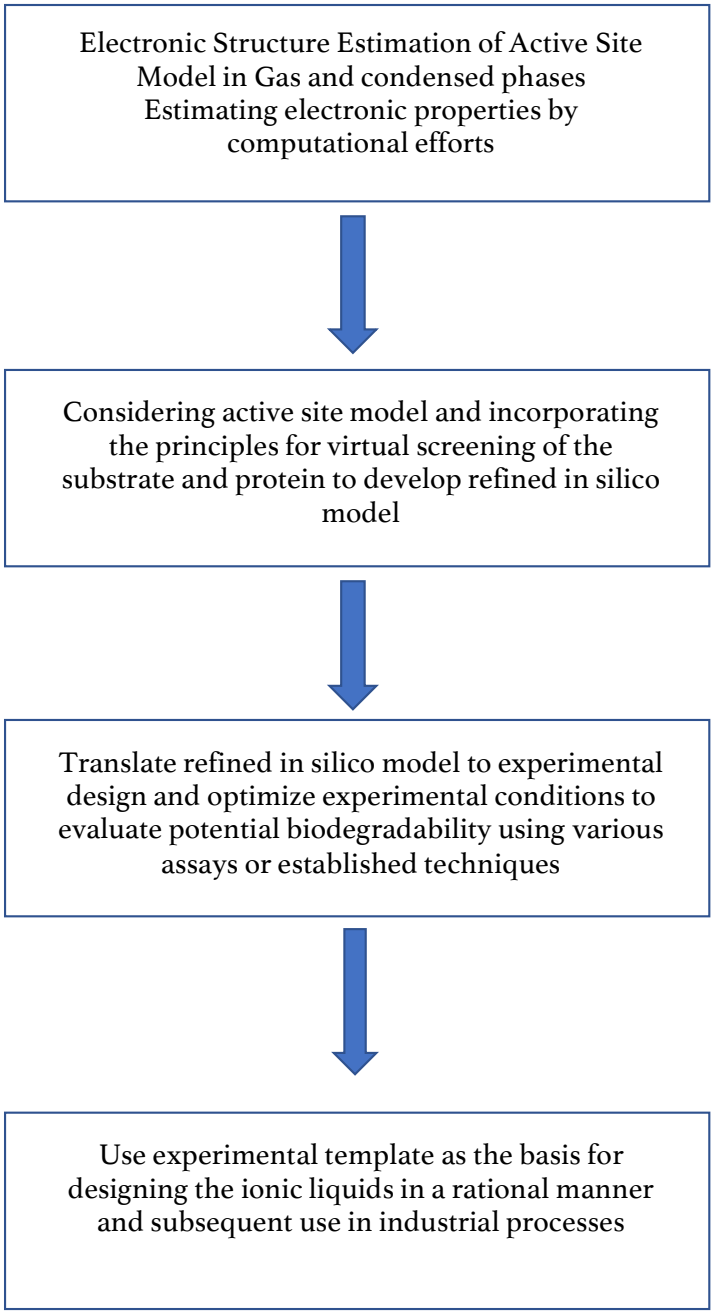


Figure 1.6: Rational design approach for Ionic Liquids from the present research work

1.2.3.5 Potential Applications

Inherently biodegradable ionic liquids are expected to have a myriad of potential applications across all sectors of society. Some of these applications are the following:

Academia (1) Inclusion of design approach courses in green chemistry with focus towards electronic structure estimation and then extrapolating the results and translating them to higher model size scales. This could potentially include part theory and part experimentation involving both sides of research.
(2) Delivering hands-on experience to students for better understanding of the design process for safer chemicals from the molecular level to bridge concepts from various scientific disciplines such chemistry, biochemistry, thermodynamics, biophysics and others.
(3) Highlighting the computational research is also expected to increase visibility of Green Chemistry principles among computational scientists that are engaged strictly in finding molecular level solutions to challenging simulation problems.

Industry (1) Widespread use of biodegradable ionic liquids will help industries meet environmental regulations with regards to air pollution and greenhouse gas emissions due to their low volatility. This is expected to enhance the viability of a lot of the more cost-affected sectors in the industry.
(2) Having a rational design approach as the backbone of a solvent being used will promote collaborations between academia and industry to gain further insight and produce even more greener chemicals in the same light as ionic liquids.
(3) Since Ionic liquids are highly flexible and task specific, their rational design will open even more avenues for formulating effective, green and robust solutions to tough and complex industrial challenges to increase efficiency of present day processes.

Society (1) Development of a rational design approach can promote the values of green chemistry among the general public so that they are aware of the virtues like 'reduce, reuse and recycle'.
(2) Use of safer chemicals in general will advance chemical sciences and create scope for nascent solvents from biomimetic materials that have gained popularity in the last two decades or so.
(3) On a philosophical note, reducing impact on the environment will give society more breathing space and create a sustainable tomorrow for coming generations. The effects of green chemistry are not limited to academicians, scientists and researchers but to the general public as well. The shift in the trend towards benign chemicals is expected to reduce burden on more populous parts of the world while preserving manufacturing ability.

1.3 Innovation and Scientific Merit

Ionic liquids have been projected to be green chemicals due to their low vapor pressure and several other environmentally benign characteristics. Although, many ionic liquids are water soluble and experimental advances have shown that they have significant toxicity and ability to bioaccumulate. Thus, their design needs to incorporate rational methods in order to successfully propagate their large scale industrial usage. Former experimental works have shown that introduction of suitable structural moiety into their cationic part can make the overall ionic liquid to be biodegradable. The mechanism put forth by pioneers in the field has been related to site-specific hydroxylation. Presently, there is a dearth in molecular level details for this mechanism and the present research works is an attempt to uncover some of those details. The development of a computational framework for understanding the biodegradability of a novel solvent forms the innovation behind this work. Such a framework has not been reported in literature related to ionic liquids and this work aims to bridge the molecular principles into macromolecular design level.

The association of the active site model from the protein cytochrome P-450 to ionic liquid cations forms the scientific merit and novelty of the work. Although cytochrome P-450 is one of the most widely studied enzymes in nature, ionic liquid cations are not considered a natural substrate to their binding site. Thus, no crystal structure has yet been reported for the association of ionic liquid cation with cytochrome P-450. The present research work begins from a reductionist approach consisting of a very minimalist representation of the active site and elucidating interactions with the cationic substrate. The key conclusions derived from this reductionist model were then used to form further objectives and analysis to be conducted on larger models. Also, the reductionist model benchmarks conformational and electronic effects to be expected from the active site model in the gas phase. These results are expected to form the principles for rational design of these liquids in the future. A tentative design approach has been depicted in Figure 1.6 that ties molecular principles and moves towards industrial application at each level.

1.4 Outline

The dissertation is laid out in a total of six chapters. The first chapter is an introduction to the basics of ionic liquids, their structural composition and the key principles of rational chemical design. Some of the key ionic components of ionic liquids have been discussed in brief. In the latter part of the section, the focus of the present research is detailed to serve as a guide for the reader. To gauge the impact of the research, potential environmental benefits have been expressed. The influence and novelty of the research presented in the work has also been detailed which distinguishes it from the available body of work in literature.

The second chapter is dedicated to a thorough review of the existing literature from two different aspects, namely, theoretical and experimental. Within this chapter, the first part is dedicated to inform the reader about the biodegradability of ionic liquids. The propensity of degradation of ionic liquids is discussed with

respect to the structural motifs present on the ionic components of the liquid. Effect of varying both the cationic and anionic parts of the molecule on the overall biodegradability is discussed. This subsection is culminated by a detailed tabulation of the available degradation data for all the cationic groups considered in this work. In the second part of the chapter, the enzyme in focus, cytochrome P-450 is introduced. In the interest of both computational chemists as well biophysicists, structural and mechanistic aspects of the workings of its active site are enumerated in a semi-detailed manner with references attached for a detailed exploration, if desired. As a key part of this research is dedicated to the catalytic cycle of P-450, it is discussed in detail and the behavior of the active site species in the presence and absence of substrate molecules is also examined.

Literature review is followed by a section explaining the theory behind the calculations and results presented in the work. The fundamental relationships governing electronic structure calculations are briefly laid out and new concepts are introduced at every step leading to theory behind DFT (density functional theory). Within DFT, the functionals employed in the present work have been discussed in some detail. This is followed by the theory related to the different analysis performed in the work to aid the reader in understanding the importance and need for each of them. The final part of the chapter is dedicated to the hybrid QMMM modeling technique and a brief explanation of its theory and implementation in the present work.

The methodology behind the calculations and remarks about the implementation are discussed in this section. At every step, the scheme behind them are discussed and program specific details are also given in some instances. The first part of this section is related to pure DFT calculations and the subsequent part is dedicated to QMMM modeling and inferences drawn in the case of inclusion of the neighboring residues of the protein.

Methodology is followed by two separate chapters; the first of which is dedicated to the minimalist representation of the active site and results obtained through a purely quantum mechanical treatment of the system. This would include binding energies, charge distribution, and electrophilicity indexes. The fifth chapter

succinctly presents the results obtained from QMMM modeling of the cytochrome P-450 including imidazolium cation as the substrate. Before molecular analysis of the optimized geometries, docking results have been given to showcase the generation of initial geometries. This chapter is completed by a brief discussion of a comparison between the active site region from the QM and QMMM models and substituted metal porphyrins.
The sixth chapter enunciates the key conclusions from the modeling efforts from the three former chapters and gives crucial details regarding the electronic structure and some mechanistic aspects of the system. It also expands on the characteristics of the active site at different levels starting from iron porphyrin to the P450 BM3 chain. The dissertation is wound up by citing recommendations for future research in this direction from a computational point of view. Also ideas for expanding this field of research are discussed in brief.

The following chapter presents an overview of the authors' experience for all the different sets of calculations performed in the work. The various approaches, their intricacies and recommendations based on experience have been mentioned in this section. Special attention has been devoted to geometry optimization as it is the backbone of the thesis and the electronic structure calculations laid out in the dissertation.

Ionic liquids have been projected to be green chemicals due to their low vapor pressure and several other environmentally benign characteristics. Although, many ionic liquids are water soluble and experimental advances have shown that they have significant toxicity and ability to bioaccumulate. Thus, their design needs to incorporate rational methods in order to successfully propagate their large scale industrial usage. Former experimental works have shown that introduction of suitable structural moiety into their cationic part can make the overall ionic liquid to be biodegradable. The mechanism put forth by pioneers in the field has been related to site-specific hydroxylation. Presently, there is a dearth in molecular level details for this mechanism and the present research works is an attempt to uncover some of those details. The development of a computational framework for understanding the biodegradability of a novel solvent forms the innovation behind this work. Such a framework has not been reported in literature related to ionic liquids and this work aims to bridge the molecular principles into macromolecular design level.

CHAPTER II

LITERATURE REVIEW

As the present work is dedicated to the development of a computational framework, the literature review will focus on two key aspects; first, the biodegradability of ionic liquids and second, the description of cytochrome P-450 and some of the studies based on the different substrates associated with it.

2.1 Biodegradability of Ionic Liquids

The four categories for estimating the biodegradability of a certain substance are defined as; (1) Primary Biodegradation - when the substance loses a specific structural part of its parent state, (2) Inherently biodegradable - if the substance degrades about 20% then its potential to biodegrade is assumed, (3) Readily biodegradable - this is a more direct measure of the degradation in which the substance is shown to degrade to a certain percentage within a given time frame, (4) Ultimately biodegradable - the substance is completely broken down into chemically simple entities in the environment, (5) Mineralization - the substance is said to decompose into organic and gas molecules that might be utilized by plants. The three steps set by the Organization of Economic Co-operation and Development (OECD) to guide towards obtaining definitive biodegradation data are ; (1) Measure the degradation in an aerobic environment to assess if it is readily biodegradable. (2) If the first step fails, then the substance is examined by using other suitable tests (under simulated conditions representing degrading environ-

ment) and (3) in the final test, the overall potential aerobic biodegradability is determined.

Some of the tests that are supported by OECD are given in detail in the excellent review given by Gathergood and co-workers [47] that include the fresh water test, inherent biodegradation, sewage (anaerobic), seawater, solid tests (soil and sediment experiments) and anaerobic biodegradation tests. Many researchers have raised concerns over how ionic liquids interact will soil and water matrix and some of the prime findings are:

(1) The mobility of the IL in the environment is inversely related to its lipophilicity (the ability of the IL to dissolve in fats, lipids, oils or non-polar solvents.
(2) The cationic part of ILs bind less strongly to soils that do not have organic matter in abundance.
(3) Concentration of minerals in soil affects the ability of the ionic liquid to pass through the soil into the water underneath.
(4) The sorption coefficient is directly proportional or linked to the alkyl chain length of the cation in the IL.
(5) The IL variants that have hydroxylated moieties associated with them adsorb weaker than pure IL compounds.
(6) The phytotoxic effect if the IL is directly related and promoted by its hydrophobicity.

One of the most important outcomes from the plethora of tests conducted on various classes of ionic liquids is that the increase in the alkyl side chain promotes biodegradability. The reason behind this being the presence of extra oxidizable carbons present in the molecule [23]. However, presence of longer chain lengths has been shown to increase antimicrobial activity so there seems to be a trade off between toxicity and biodegradability for ionic liquids. Interpretation of biodegradation data can be a challenging task due to the fact that each IL consists of both the cationic and the anionic part in itself. As an example, if an inorganic anionic

moiety is present in the IL, it does not serve as a carbon source. On the contrary, if an IL consists of an anionic part that is carbon rich, it should not be interpreted as having high overall biodegradability. Because of these reasons, it becomes very important to assess the effects of individual ionic components to identify the overall biodegradability of the ionic liquid. Some of the key inferences for both anions and cations will be discussed below that will also form the motivation behind carrying out the present work.

2.1.1 Anions

Halides are some of the most abundant class of anions. For these ILs, biodegradation is solely dependent on the cation. In this direction, [bmim] (1-butyl-3-methylimidazolium) derivatives have been shown to give poor biodegradability. Many organic anions have been synthesized and tested from natural chemicals such as amino acids and carbohydrates. Alkylsulphate anions have been shown to exhibit very good rates of biodegradation. Of these, dodecylsulphate is popular for assessment of activated sludge activity. When paired with a cholinium cation, some of the variants have been shown to have high rates of biodegradation. Similarly, natural organic acids and equivalent sugars have been paired to yield appreciable biodegradation rates ($\geq$ 50% having dibutyl-dimethyl ammonium cation). Structurally, a free carboxylate group promotes biodegradability while branching on amino acid side chains impede the same. Overall, the structure of cation plays a very salient role even in the degradability of the IL having an organic anion. In terms of the design of ILs, researchers suggest that presence of a halide anion can serve as the role of mineralized ion in a potentially biodegradable liquid. Among the different ions, anions derived from organic acids appear to be the most appropriate selection for the design of environmentally friendly ILs.

2.1.2 Cations

The selection of cation is essential due to the fact that most of the ILs synthesized till date have more carbon content in the cation rather than the anion. This

content translates to a higher overall CO_2 evolution when analyzed through a certain biodegradation screening method such as the CO_2 headspace test. Similar to anions, the research efforts in order to select benign cations has been extensive involving a variety of scaffolds, tests and microbial environments.

One of the most widely used ILs consist of the imidazolium group, which forms majority of the basis for the present work. To improve the biodegradability of imidazolium-based ILs, two distinct modifications have been carried out to the native structure; firstly, the modification of nitrogenous substituents in the ring of the cation and secondly, substitution at the C2 and C4/C5 positions [48] on the ring. Initial tests on these compounds have shed light on their biodegradability levels. Modifications to these cations by means of introducing amide and ester groups at the C2 positions did not enhance their biodegradation levels (compared to the situation in which an ester linkage was added to the compound). In the structural perspective put forth by Stolte [49, 50], it has been outlined that while the ring and its carbon derivatives can be classified as ultimately biodegradable, nitrogenous substitutions are bio-accumulative. Possible reason behind this being that the presence of N-moiety reduces the propensity of chemical attack by enzymes. On the positive side, appreciable biodegradability was noted for compounds having side chain lengths of C6 and C8, that have been termed inherently biodegradable. This might be attributed to their favorable uptake into microbial communities [51, 52], which is promoted by their hydrophobic nature [53, 54]. This presents an opportunity for their metabolization by appropriate enzymatic systems such as cytochrome P-450. Moreover, the addition of functional groups makes the cation more reactive and in turn, more susceptible to biotransformation. Although that being said, it is now widely understood that even the incorporation of functional groups into short alkyl chain containing cations (butyl or less) does not make them significantly biodegradable. Even inclusion of an ester group into the side chain at the C2 position only enhances their degradability slightly. The imidazolium ring remains intact in most of the tests and the inclusion of electron-rich groups only make the overall ring structure even more stable.

The second most widely studied cation is the pyridinium with biodegradation studies conducted on a lot of different variants of it. Some of them include -amino , -carboxy and -alkoxy derivatives. Similar to imidazolium-based ILs, increase in the side chain of the alkyl moiety is shown to promote the biodegradability of these compounds. Although, the degradation of pyridinium containing cations has been observed to be significantly higher than its imidazolium counterparts. Also, a correlation has been drawn between the length of side chain and overall mineralization for N-alkyl-3-methyl pyridinium compounds. For chain lengths of ≥ 6, complete primary biodegradation has been observed for the parent compounds not containing the 3'-methyl moiety. Another subtle difference arises from the fact that in the case of pyridinium, even the core ring head is accessible to the biodegradation process. On the other hand, as already stated, in case of imidazolium, the ring remains recalcitrant following the degradation of the side chain. In the case where an ester group was incorporated into the 1-N position of the cation, it made the overall IL readily biodegradable.

Thiazolium-based ILs have only been sparingly investigated [55] but it was included in the present studies to uncover the effect of the presence of a heavy S atom in the ring moiety. It is the third and final aromatic system to be considered in the thesis work. The only biodegradation studies performed on this system have shown that they are poorly biodegradable with only ≤ 7% inferred from a CO_2 headspace test.

Due to the wide range of structural possibilities associated with the IL cation, it was decided to also investigate a cation containing no aromaticity in its ring head. Pyrrolidinium based cations were chosen due to the presence of biodegradability data from former studies in literature. Even for different derivatives of pyrrolidinium cations (not containing -OH group), shorter side chain lengths (≤ C_4) have yielded very low biodegradation values. The extension of the alkyl chain to about C_8 has shown to make the IL readily biodegradable [56]. For pyrrolidinium,

the design rules align with the ones set for imidazolium and pyridinium classes of ILs thus leading to a formation of a unified concept for rational design.

The final class of ILs considered in this work are the linear cations that are devoid of any cyclic or aromatic groups. Firstly, ammonium containing compounds such as cholinium have been widely synthesized and studied for their biodegradability. Among these, two compounds, namely, choline 2-naphthoxyacetate and choline anthracene-9-carboxylate were shown to be poorly biodegradable. One of the rules laid out in the seminal work by Boethling [21] stating that presence of polycyclic residues hinder natural breakdown. Although, when the anions are non-organic, cholinium containing compounds were shown to exhibit high levels of biodegradability $\geq$ 80% which is attributed to the presence of the hydroxyl group at the terminal position of the cation. Another popular variant of the straight ammoniums, the tetraalkylammonium cations are widely used in the surfactant industry. These compounds, on the other hand, show negligible biodegradability. The maximum estimate has been about 20-25 % for a tetrabutyl ammonium cation consisting of an amino acid derived anion. The biodegradability of this class of compounds have been shown to be driven by the choice of the anion, contrary to the rules laid out for aromatic and cyclic species.
Phosphonium ILs have shown no significant biodegradability, with or without functionalized alkyl side chains. The only few observations of high biodegradation rates have been attributed to the presence of octylsulfate as the anion which cannot be interpreted as the biodegradability of the cation. Even the inclusion of ester groups did not increase the propensity of this class of cations to degrade naturally. The maximum observable degradation has been only about 30 % , that is much lower than some of the other classes of cations mentioned in the present work. The remaining class of cations has been selected as the sulphonium type owing to the fact that this motif has been utilized in a load of applications as electrolytes. Although, there seems to be no available biodegradation data for this type of cation.

2.2 Metabolite Studies of IL degradation

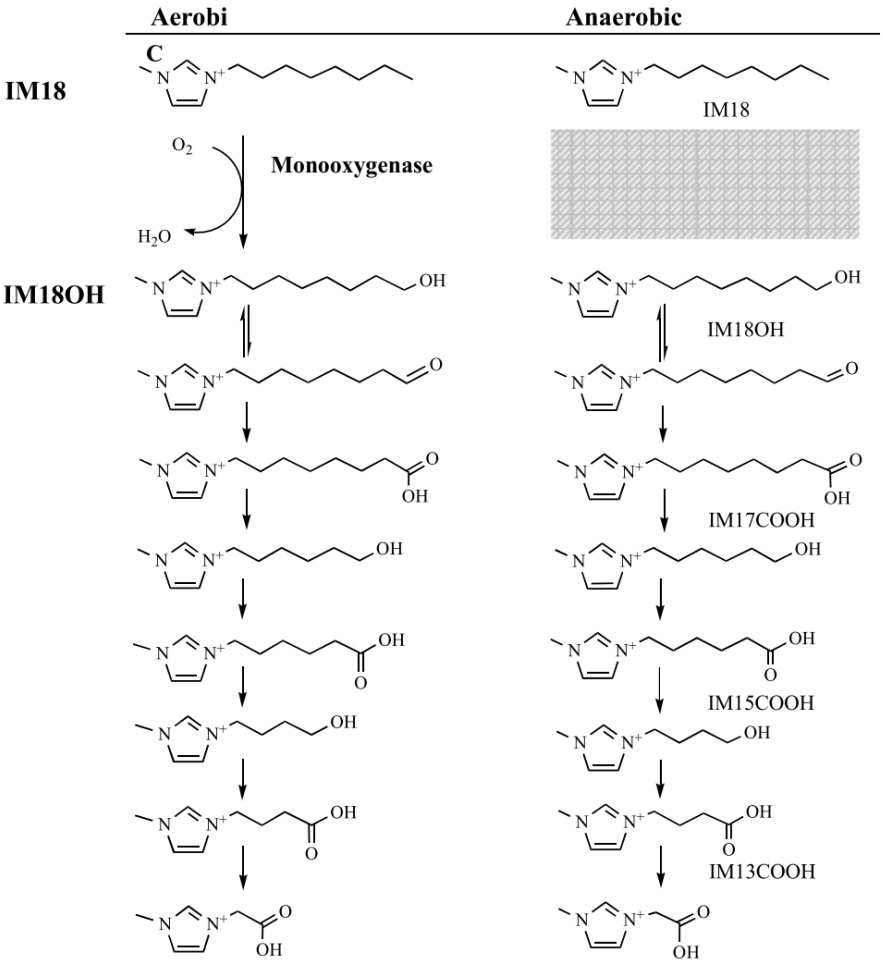


Figure 2.1: IL degradation mechanism postulated in both aerobic and non-aerobic environments. Adapted from Neumann et al. [4]

Based on the work by Jastorff and co-workers [40], a model for predicting the metabolites for a butyl imidazolium cation ($[C_4mim]^+$) was formulated. A wide variety of oxidized products were generated to check for their toxicity. An HPLC-

MS analysis revealed that the postulated structured were in accordance with their experiments. The detections suggested a mechanism by which the biodegradation of imidazolium-based cations might occur in both aerobic and anaerobic environments. The conversion of the hydrophobic side chain of the cation would seem to start from the oxygen insertion at the terminal site. This is probably going to be assisted by a monooxygenase such as cytochrome P-450. The alcohol that is formed in the first step then would subsequently oxidize to aldehyde and carboxylic acids by appropriate dehydrogenases. The cationic acids then would undergo β-oxidation and the remaining fragments would enter the TCA cycle via a coenzyme acetyl Co-A. The probability of non-terminal alcohols has also been mentioned and that they would not be further degraded by the β-oxidation process leading to either ketones or other hydroxylations. The products that have been identified from the initial oxidation have degraded or functionalized chains, all indicating that the reactivity must be conducive towards the hydrophobic part of the cation. Although, the converted products are all shown to be less toxic than its parent with the final fragment consisting of the ring intact.

A similar outcome was made in case of a butyl pyridinium cation using an LC-MS technique. After a month of incubation in the presence of a microbial consortium, hydroxylation at the chain was observed leading the authors to infer that the same metabolic pathway as the imidazolim cased would be probably in this case. Another key conclusion from the work suggested that the conversion pathways are directly related to the size of the external side chain.

In an anaerobic environment, the effective biodegradation was observed to be very low even in the case of long alkyl chain containing cations and only a select few yielded secondary metabolites that were oxidized. Even in this case, literature strongly suggests that in the absence of aerobic environment, the oxidation cannot proceed through a molecular oxygen by a monooxygenase and thus, biodegradation is not readily initiated. Some of the key suggestions have been outlined that form the basis for the hypothesis made in this work and the foundation for the

research work carried out in this present dissertation.

(1) The biodegradability of cations containing different head groups needs to be studied systematically that would involve pathways, metabolites, mechanisms and reaction kinetics in a variety of the environments.
(2) The focus of the studies should not be strictly towards readily biodegradable compounds as ILs are not readily biodegradable in most cases. The test conditions utilized should not be constricted or biased in any manner to reflect the true nature of these compounds
(3) The microbial consortia or family would have a strong impact on the overall biodegradability of the compounds and hence their identification, isolation and applicability needs to be understood in order to degrade ILs and design them in a more rational manner.
(4) Finally, biodegradation data should be carefully analyzed to avoid misinterpretation. Key conclusion such as erroneous pass levels might lead to ineffective design principles and geometrical classification.

2.3 Cytochrome P-450

As indicated in the previous section, the biodegradability of ionic liquids is mechanistically related to the first step which is introduction of oxygen into a suitable position along the alkyl chain of the cation. It has been established that this oxygen insertion is brought about by an alkyl monooxygenase in the aerobic environment [4, 9, 57, 58, 59, 60, 61]. Out of the various enzyme families capable of oxygenation [62], cytochrome P-450 has been shown to have great diversity in its chemical reactivity [63, 64, 65, 66]. It is capable of serving as an agent for a wide range of biotransformations [67, 68, 69, 70]. One of the most of important characteristics of the P-450 family is the presence of a variety of isoforms of the same enzyme, but sharing the same active site [71]. The present work is dedicated to the modeling of the active site of P-450 in conjunction with the IL cation as

its substrate to develop an understanding of the molecular mechanism of binding *in silico* that would ultimately enable the rational redesign of both the enzyme and the substrate in order to facilitate IL biodegradability. The catalytic cycle of cytochrome P-450 (Figure 2.2) is a collection of seven separate steps leading to the final hydroxylation of the substrate compound [72]. The series of processes leading to oxygen insertion are advocated by a collection of chemical and physical changes made to the active site of the protein, i.e. the heme. The changes in the molecular make-up of each of the species in the cycle have been represented in the Figure 2.2. Each of the species will be discussed in the following subsections to help the reader understand the processes involved in the cycle and a deeper understanding of the working mechanism of cytochrome P-450.

2.3.1 Resting State (S_1)

The resting state species consists of the heme ligated to a water molecule at the distal end [73, 74, 75]. Experimentally, the species has been shown to occupy a doublet state with a closely echoing spin equillibrium existing between the doublet and sextet states. Although, the relative energetics between the two spins is directly dependent on the functional used to evaluate the model, it is now firmly established that the doublet state changes to a sextet only on the migration of the water molecule from the distal end and leaving a pentacoodinated Fe in the porphyrin. In the presence of other water molecules in the binding pocket [76, 77], the water molecule attached to the porphyrin shows a slight tilt of about 30-50 degrees as a direct consequence of intermolecular hydrogen bonding. The medium surrounding the active site molecule has varying degree of effects on it as shown by the bond lengths, that are significantly shorter in the protein embedding. The protein environment influences orbital mixing of the d_{z^2} on iron with the participating p_z orbital on underlying sulphur reducing the antibonding character.

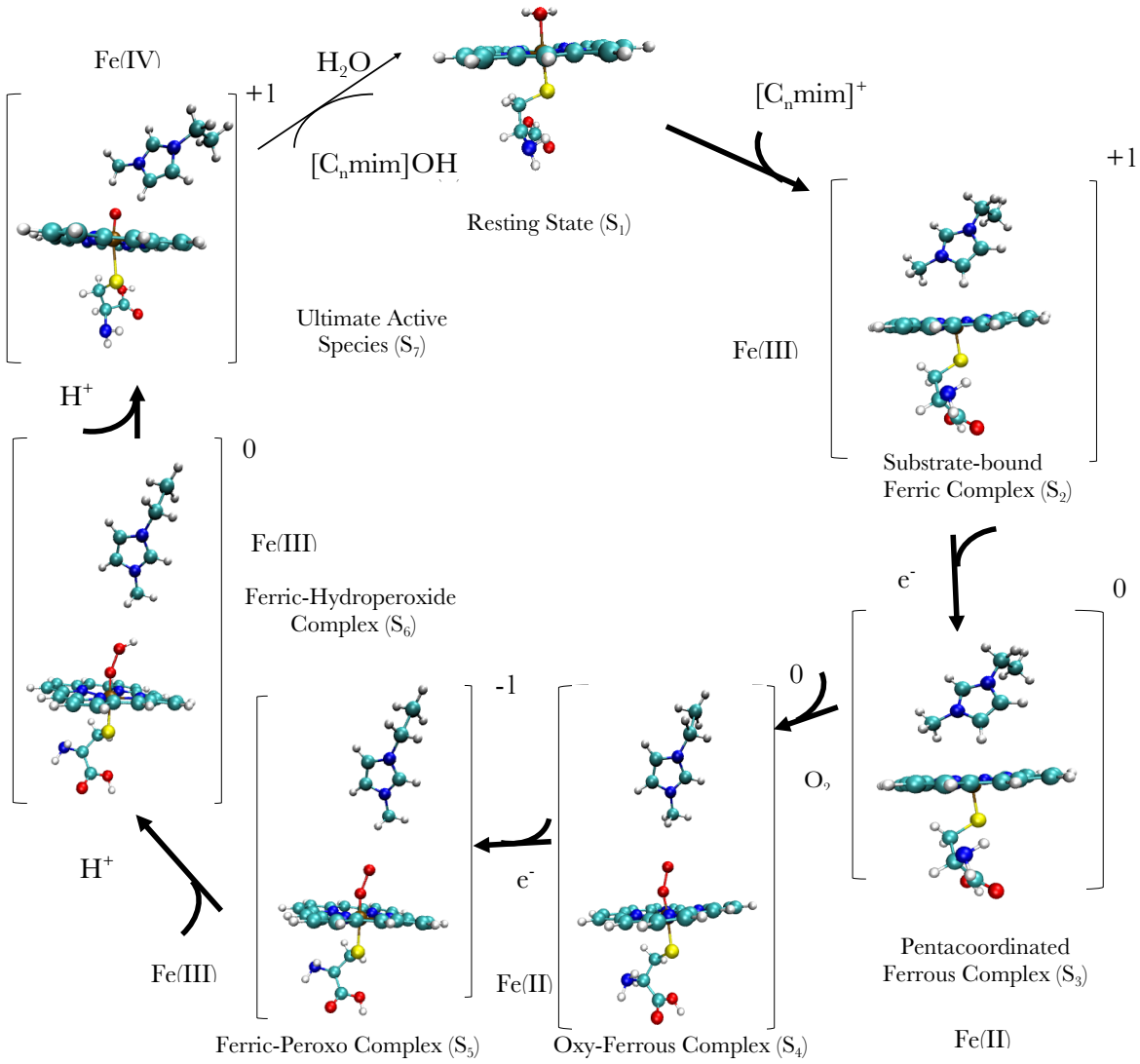


Figure 2.2: Catalytic cycle of cytochrome P-450 showing imidazolium-based IL cation ($[C_2mim]^+$) as substrate

2.3.2 Substrate-bound Ferric Complex (S_2)

The binding of substrate to the resting complex removes the water molecule from the distal position and changes the spin state of the heme complex [78, 79, 80, 81]. Experimental evidence suggests that the orbital make up of this species in the ground state is $d^1_{x^2-y^2} d^1_{xz} d^1_{yz} \sigma^{*1}_{z^2} \sigma^{*1}_{xy}$ with closely lying quartet and doublet states. DFT calculations have suggested that the removal of water causes the Fe atom to be pushed towards the sulphur containing moiety underneath as a direct conse-

quence of the alteration in orbital mixing from the resting species [82, 83, 84, 85]. The sextet ground state is valid for both the SCys and SH ligand representations, i.e. the proximal end containing sulphur bound to a cysteine and sulphur bound to a single hydrogen atom respectively [86, 87, 88]. The bound ferric complex then undergoes an electron addition changing the oxidation state of iron to Fe(II).

2.3.3 Pentacoordinated Reduced Ferrous Complex (S_3)

The reduced ferrous complex following the reduction of the substrate binding yields a quintet ground state that is reported to have two possible electronic configurations, $d^2_{x^2-y^2}d^1_{xz}d^1_{yz}\sigma^{*1}_{z^2}\sigma^{*1}_{xy}$ and $d^1_{x^2-y^2}d^2_{xz}d^1_{yz}\sigma^{*1}_{z^2}\sigma^{*1}_{xy}$. As with the two previous species, there exist multiple closely lying electronic states in addition to the ground state [89, 90]. Both singlet and triplet states are possible that have both the primitives $d_{x^2-y^2}$ or d^1_{xz} having a set of electrons. The reduction potentials have been studied extensively for different bound and unbound models of P-450 and vary according to the media and active site presented. Thermodynamically, it has been shown to be exothermic in gas phase, protein phase and embedding with the with the protein phase being most favorable among the three. Structure calculations have also shown that the reduction of $S_2 - S_3$ is more exothermic than the reduction of S_1 to its anionic counterpart. This becomes significant as it supports the experimental gating mechanism that is influenced by the proximal residue. In the absence of the proximal ligand, both S_1 and S_2 can be easily reduced and might disrupt the catalytic cycle. The presence of thiolate or its variant in the proximal side makes the resting species a poorer electron acquiring agent, so that only the substrate bound ferric complex may be reduced. DFT studies for the electron transfer event have suggested that there seems to be a redox connection between the binding and reductase domains within P450 and the minimum distance between the metal redox centers between the two domains is about 14 Å. The electron pathway has been shown to be connected to the proximal cysteine rather than the propionate chains connected to the central heme.

2.3.4 Oxy-Ferrous Complex(S_4)

The first reduction step is followed by dioxygen addition to the distal side of the active site molecule. The formation of this complex is rather kinetically driven and it is known to be a singlet electronically. This species has an inherent radical nature which presents a unique challenge for its structure calculation. To determine its correct orientation and electronic field, the symmetry of the paired electron needs to be broken by using the wavefunction estimated from the next higher spin and imposing it on a singlet state. The structure of this complex has been of much debate but crystallographic evidence points towards an end-on complex with a bent O_2 attachment to be the appropriate ground state. The bonding for S_4 complexes suggest that it is an open shell singlet species with two unpaired electrons having dominant metal and dioxygen character. From the Weiss model [91], it has been established that the ferrous porphyrin donates a single electron to the π^*_{xz}(OO) orbital to make it doubly occupied and it in turn coordinates to Fe through a non-ionic dative σ bond. This is maintained by an orbital overlap between the filled 2p atomic orbital on the Fe linked oxygen and the Fe d^2_z vacant orbital. Simultaneously, the singly filled d_{yz} orbital of ferric porphyrin links with π^*_{yz}(OO) orbital to form a distinct singlet pair.

2.3.5 Ferric-Peroxo Complex(S_5)

The formation of the Ferric-Peroxo complex is the second reduction event within the catalytic cycle. It has shown to only exist in the doublet state by both experiments and structure calculations. Experimentally, it has also been shown that the Fe-O bond length in S_5 is stronger than that of the previous oxy-ferrous complex which is difficult to capture through DFT calculations. The reduction potential for this step also poses a significant challenge for validation with the values ranging from -1.04 to -2.28 V based on the chosen environment, model and protonation state. The electronic structure of S_5 is simpler than S_4 with only a single unpaired electron in the π^*_{yz}(OO) orbital. It also has potential for hydrogen bonding with nearby substrates such as arginine. Multiple works have indicated that this pro-

cess is not facile and establish it to be the rate determining step for the whole cycle. The presence of different substrates in the binding pocket although, have the ability to reduce the energetic barrier and exothermicity of the process.

2.3.6 Ferric-Hydroperoxide Complex(S_6)

The hydroperoxide complex is generated by the first proton addition step in the cycle and it structure has been well established through crystallographic efforts. The ground state of this species has been shown to be a doublet with a single unpaired electron in the π^*_{yz} orbital. In addition to this, one of the key aspects of the experimentally resolved structure is the hydrogen bonding between the proton of the adjoining OH group and one of the nitrogens of the porphyrin. Unlike the Ferric-Peroxo species, different QM models for this species validate to similar structural parameters in the active site. They are not shown to be sensitive to the model size and method chosen for the calculation. Some of these parameters are the Fe-O and Fe-S bond lengths that are in accord with the experimental values for a variety of models evaluated at both QM and QMMM levels. The proton addition to the peroxo complex indicates the participation of key residues such as Asp251 and Thr252 in addition to a crystallographic water molecule in the binding pocket. Indeed, the formation of the complex involves the migration of a single proton from the bulk solvent phase to the active site with a significant reorientation of the proton network. The inbuilt proton transfer machinery involves an acid alcohol pair coupled with a network modulated by a water molecule. The free energy associated with the formation of the hydroperoxide complex seems to be highly variable and sensitive to the model description and protonation states involved. A number of models have been developed for understanding the protonation using molecular dynamics, QM and QMMM and they yield different energy barriers and conformational aspects related to this process.

2.3.7 Ultimate Active Species(S_7)

The second protonation step leads to the formation of the ultimate active species that participates in the hydroxylation of the substrate [92]. This step shows a very

high exothermicity due to the fact that the former S_6 complex is a very strong base with a high proton affinity. Several mechanisms for this step have been suggested for this protonation and the most popular one involves two steps in series. The first step, as the -OH moiety departs, is the transport of an electron from the porphyrin to outgoing OH group converting it to an anionic OH. It is followed by the protonation of OH^- by Thr252 and the simultaneous protonation of Thr 252 by a water molecule. This proton-coupled electron transfer is made possible by the barrierless proton transfer in the Asp251-Thr252 channel and the push effect of the proximal sulphur containing group. The ultimate active species in an elusive chemical entity owing to multiple possible spin states and electronic configurations that are closely related in terms of energy [93]. Several experiments have been successful in isolating this species and indicate the presence of the active species in both doublet and quartet states. Early investigations on determining the ground state were shown to be sensitive to the model type and choice of the proximal attachment. It was shown to possess three orbitals having single occupation, namely, π^*_{xz}, and π^*_{yz} and the third one. The third singly occupied orbital depended on the proximal model and interactions with the protein environment. Upon considering the SH^- or $SCys^-$, the model yields a quartet configuration having the orbital description $\pi^*_{xz}\pi^*_{yz}a^1_{2u}$.

2.3.8 Substrate Hydroxylation by Cytochrome P-450

In this section, the hydroxylation of small molecules to cytochrome P-450 is discussed. It is compiled to review the basics of the interaction between the active site molecule(s) of P-450 proteins with substrates having widespread chemical nature.

A typical P450 oxidation needs one molecule of oxygen and two subsidiary electrons from its redox partner, NADPH to add the oxygen to a given substrate. If the oxygen consumed is greater than the amount of product formed, then the enzyme is uncoupled. This occurs in the case where the dioxycomplex (S_4) reverses to the ferric reduced state and dissociates with the adjoining superoxide, or when a set of electrons reduce S_7 to an aqueous molecule before it has the chance to

hydroxylate the associated substrate [94, 95, 96]. Uncoupling reduces the catalytic efficiency of the enzyme giving low hydroxylation rates.Thus, there are competing mechanisms at play in the case of cytochrome P-450 catalytic cycle that depend on the chemical conditions during the process.

Next, we discuss some of the accepted hydroxylation mechanisms for different substrates. In the most widely accepted scheme [97, 98], a hydrogen abstraction takes place from the substrate by the ferryl oxygen in S_7 which is followed by a recombination of the carbon radical with the oxygen in a rebound manner. The iron(IV) oxo porphyrin radical cation has been shown to be responsible for most of the oxidations by cytochrome P-450 that has been supported by both experiments and calculation in both QM and QMMM schemes. The S_7 species first abstracts an hydrogen atom from the most conducive site on the substrate producing a hydroxyl species and a representative carbon radical. The hydroxyl species at the porphyrin can be viewed as a complex consisting of an Fe(III) porphyrin with an attached hydroxyl radical. In the next step, the -OH radical from the porphyrin undergoes a recombination step in which it combines with the carbon radical at the substrate end.

This radical association often involves change in the stereochemistry of the substrate with the most cited example being the abstraction of either the 5-exo or endo hydrogen from camphor by P-450 cam [99, 100] to yield a selective 5-exo-hydroxycamphor. In the same vein, allylic rearrangements leading to migration of resonance and double bonds have also been noted in hydroxylations mediated by cytochrome P-450. The rate of the -OH radical rebound association step has been measured by the use of radical clocks. In a typical clock, a strained ring is directly connected to the carbon that acts as the site for the proposed hydroxylation. Ring strain associated aids in the irreversible rearrangement to the radical. The product formed is also in equillibrium with an unrearranged product, both of which give rates that can be independently measured giving the lifetime of these products. One of the concerns of this process was the impediment of the rearrangement by the radical probe attached to the structure [101]. This was confirmed to be false owing to observations such as; high rates with other similar clocks, large substrates

also undergoing significant motion under the influence of P-450 and intramolecular isotopic effects shown by some of the probes involved. Further advancements led experimentalists to believe that the extent of rearrangement is dependent on both the initial transition state formed by hydrogen abstraction from substrate and the radical recombination state. In view of this, a two-state reactivity theory was proposed which involves two competing pathways leading to hydroxylation. The dominant pathway although, has been shown to be dependent on the substrate and the environment involved in the process.

The hydroxylation mechanism proposed by Shaik [102] revolves around the two state reactivity [103] of the ultimate active species (S_7). The orbital make up of the active species has been shown by Shaik and co-workers to be a triradicaloid [104] with the configuration of $\pi^*_{xz}\pi^*_{yz}a^1_{2u}$ having a pair of energetically closely related ground states. The two state reactivity states that each of these ground states would give rise to a unique set of specific products with their individual stereo and regioselectivities [105]. To summarize, the ultimate active species is a two state oxidant whose electronic activity can be tuned by the overall polarity of the binding pocket and hydrogen-bonding scheme within the pocket. Different substrates feature different modes of reactivity with alkane epoxidation [106, 107] following two state reactivity while benzene hydroxylation and sulfoxidation show a doublet and quartet state being dominant. Substrate deformation or geometrical changes has been shown to relate to the electronic state of the underlying reactant species [108, 109, 110]. The porphyrin macromolecule is also shown to not be a mere spectator ligand but acts as an electron acceptor during the oxidation process. Also, during the later protonation steps, it acts as a proton shifting agent that shuttles them to the substrate. Other unusual reactions have also been observed for compounds such as alleles and alkynes. A diversity of reaction outcomes are possible that make P-450 to be extremely challenging as well as exciting in terms of molecular design [111].

Ionic liquids have been projected to be green chemicals due to their low vapor pressure and several other environmentally benign characteristics. Although, many ionic liquids are water soluble and experimental advances have shown that they have significant toxicity and ability to bioaccumulate. Thus, their design needs to incorporate rational methods in order to successfully propagate their large scale industrial usage. Former experimental works have shown that introduction of suitable structural moiety into their cationic part can make the overall ionic liquid to be biodegradable. The mechanism put forth by pioneers in the field has been related to site-specific hydroxylation. Presently, there is a dearth in molecular level details for this mechanism and the present research works is an attempt to uncover some of those details. The development of a computational framework for understanding the biodegradability of a novel solvent forms the innovation behind this work. Such a framework has not been reported in literature related to ionic liquids and this work aims to bridge the molecular principles into macromolecular design level.

CHAPTER III

THEORY

3.0.1 Molecular Modeling

3.0.1.1 Electronic Structure Calculations

The end goal of most quantum mechanical modeling approaches is to approximate the solution of the temporally independent Schrodinger equation devoid of any relativistic effects:

$$\hat{H}\psi_i(\vec{x_1}, \vec{x_2}, ...\vec{x_N}, \vec{R_1}, \vec{R_2}, ...\vec{R_M}) = E_i\psi_i(\vec{x_1}, \vec{x_2}, ...\vec{x_N}, \vec{R_1}, \vec{R_2}, ...\vec{R_M}) \quad \text{(III..1)}$$

where $\hat{H}$ represents the Hamiltonian operator for a system of subatomic particles consisting of M nuclei and N electrons in an isolated mode that is away from the influence of any magnetic or electric fields. The above stated Hamiltonian is a differential operator that represents the total energy of the system:

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \frac{1}{2}\sum_{i=1}^{N}\frac{1}{M_A}\nabla_A^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{R_{iA}} + \sum_{i=1}^{N}\sum_{j>1}^{N}\frac{1}{R_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}} \quad \text{(III..2)}$$

where the letters A and B run over the M nuclei, i and j sum over the N electrons in the system. The first an second terms of the equation represent the kinetic energies of the electrons and the nuclei employing the Laplacian operator ∇_i^2 which is the sum of directional differential operators:

$$\nabla_i^2 = \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} \tag{III..3}$$

M_A is the mass of the nucleus A proportional to the mass of a single electron. The third, fourth and fifth terms define the potential energy part of the Hamiltonian and they stand for the electrostatic interaction between the nuclei and the electronic particles that is attractive, while the final two terms give the repulsive potential energy due to similarly natured particles. R_{AB} is the distance between the particles in consideration, i.e. A and B. The term ψ_i stands for the wave function of the given state of the system, that depends on the position of the particles, i.e. the 3N spatial coordinates $\vec{r_i}$. Also, the electrons are represented by N spin coordinates ,s_i which are collectively attributed to x_i. For the nuclei, the 3M spatial coordinates are given by $\vec{R_i}$. All the energetic information about the system is embedded within the wavefunction while E_i is the energy described by that wavefunction.

This fundamental equation representing the motions of the subatomic particles can be effectively reduced if we consider the relative masses between an electron and a nuclei. Due to its heaviness of the order of about 1800 even for the simplest element H, the nuclei move much slower than electrons [112]. Thus, the negatively charged electrons may be considered as moving in a scheme of fixed nuclear coordinates. This statement is known as the Born-Oppenheimer approximation that aids in the bifurcation of nuclear and electronic motion in a molecular system. Due to fixed nuclei, their kinetic energy is zero and the potential energy due to their subsequent repulsion is constant. The overall Hamiltonian then reduces to an electronic nature given by:

$$\hat{H_{elec}} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{R_{iA}} + \sum_{i=1}^{N}\sum_{j>1}^{N}\frac{1}{R_{ij}} = \hat{T} + \hat{V_{Ne}} + \hat{V_{ee}} \tag{III..4}$$

where $\hat{T} + \hat{V_{Ne}} + \hat{V_{ee}}$ represents the sum of electronic kinetic energy, potential energy due to interaction between nuclei and electron and electronic repulsion. The solution to the above equation yields the electronic wave function that explicitly depends on electronic coordinates while nuclear coordinates are parametric in nature. Hence, the total energy of the system is the sum of the electronic energy and the nuclear repulsion term that is constant.

$$\hat{H_{elec}}\psi_{elec} = E_{elec}\psi_{elec} \tag{III..5}$$

$$E_{total} = E_{elec} + E_{nuc} \tag{III..6}$$

where E_{nuc} is the energy of nuclear repulsion i.e.

$$E_{nuc} = \sum_{A=1}^{M} \sum_{B>A}^{M} \frac{Z_A Z_B}{R_{AB}} \tag{III..7}$$

The potential exerted on the electrons due to the heavier particles is termed the external potential $\hat{V_{Ne}}$. It is important to note that the wavefunction is not physically observable [113]. In order to mathematically determine an observable associated to this function for a system, its squared term is considered which stands for the probability of finding electrons in appropriate elements of volume $|\psi(\vec{x_1}, \vec{x_2}, ...\vec{x_N})|^2 d\vec{x_1} d\vec{x_2} .....d\vec{x_N}$. Upon switching the order of electrons, the probability must remain the same, i.e.

$$|\psi(\vec{x_1}, \vec{x_2}, ...\vec{x_i}, \vec{x_j}..\vec{x_N})|^2 = |\psi(\vec{x_1}, \vec{x_2}, ...\vec{x_j}, \vec{x_i}..\vec{x_N})|^2 \tag{III..8}$$

In case of electrons, the change in consideration of their orders renders the wavefunction antisymmetric, i.e. the wavefunction is not equal with respect to change in order and spin. Electrons are fermions whose functional description changes upon interchange of the spatial and spin coordinates of any two particles, i.e.

$$\psi(\vec{x_1}, \vec{x_2}, ...\vec{x_i}, \vec{x_j}..\vec{x_N}) = -\psi(\vec{x_1}, \vec{x_2}, ...\vec{x_j}, \vec{x_i}..\vec{x_N}) \tag{III..9}$$

This antisymmetry principle represents the fact that no two electrons can simultaneously occupy the same spin or spatial state and is more generally known as the Pauli's exclusion principle. The direct consequence of this relationship is that the integral of this wavefunction over the full range of system volume comes out to be exactly unity. i.e. the wavefunction is said to be normalized.

$$|\psi(\vec{x_1}, \vec{x_2}, ...\vec{x_N})|^2 d\vec{x_1} d\vec{x_2}.....d\vec{x_N} = 1 \tag{III..10}$$

In order to solve the general Schrodinger equation for any molecule, initially a Hamiltonian needs to be setup for it. Upon looking at the Hamiltonian, it is evident that the molecule provides information about the number of electrons and nuclei to be considered and the external potential (calculated from the positions and charges of the nuclei). The remaining operators for the kinetic energy as well as the electron repulsion are independent of the molecule that is under consideration. Then, the eigenfunctions and corresponding eigenvalues of this Hamiltonian need to be determined. The exact calculation of the wavefunction is not possible for systems larger than Helium and for this reason, it is approached through various principles of approximation. To this end, the variational principle developed in quantum mechanics states that the energy computed via any observable using a normalized wavefunction is greater than the ground state wavefunction. The trial wavefunction operated on the initial Hamiltonian yields an energy that is always higher in value than the one computed using the ground state wavefunction. The equation holds true only for a single solution for which the trial and ground state wavefunctions are equal. The eligibility of a wavefunction is defined by two factors; it being continuous in all space and integrable quadratically, otherwise the normalization principle would be violated.

$$\langle \psi_{trial} | \hat{H} | \psi_{trial} \rangle = E_{trial} \geq E_0 = \langle \psi_0 | \hat{H} | \psi_0 \rangle \tag{III..11}$$

The exact wavefunction cannot be identified out of all the N-electron wavefunction by eliminating all but the most eligible function, the variational principle can be used to determine subsets of all the possible solutions. One of the schemes of determining subsets of eligible wavefunctions is the Hartree-Fock approximation which states that the construction be made from all antisymmetric products consisting of 'N' spin orbitals. The $\hat{H}$ can be uniquely determined by obtaining Z_A and R_A which gives the external potential composed of the interaction between electrons and nuclei. Through this operator, the ground state wave function can be determined in principle

$$E_0 = \min_{\psi \to N} E[\psi] = \min_{\psi \to N} \langle \psi | \hat{T} + \hat{V_{Ne}} + \hat{V_{ee}} | \psi \rangle \tag{III..12}$$

$$[N, Z_A, R_A] \to \hat{H} \to \psi_0 \to E_0, E_0 = E[N, V_{ext}] \tag{III..13}$$

Hartree-Fock Approximation The HF-approximation gives the N-electron wavefunction in the form of an antisymmetrized product of N separate 1-electron sub wavefunctions $\chi_i(\vec{x_i})$ whose product is the Slater determinant :

$$\psi_0 \approx \phi_{SD} = \begin{vmatrix} \chi_1(\vec{x_1}) & \chi_2(\vec{x_1}) & ... & \chi_N(\vec{x_1}) \\ \chi_1(\vec{x_2}) & \chi_2(\vec{x_2}) & ... & \chi_N(\vec{x_2}) \\ ... & ... & ... & ... \\ ... & ... & ... & .. \\ \chi_1(\vec{x_N}) & \chi_2(\vec{x_N}) & ... & \chi_N(\vec{x_N}) \end{vmatrix} \tag{III..14}$$

The one-electron $\chi_i(\vec{x_i})$ functions are the spin orbitals , that are composed of an orbital representing their spatial coordinates and a spin function. Hence, each term of the Slater determinant is given as a product of the spatial and spin orientation. The spin functions are orthonormal in nature which gives :

$$\chi_i(\vec{x_i}) = \phi_i(\vec{r_i})\sigma_i(\vec{s_i}), \langle \alpha | \alpha \rangle = \langle \beta | \beta \rangle = 1, \langle \alpha | \beta \rangle = \langle \beta | \alpha \rangle = 0 \tag{III..15}$$

For determining the minimum energy wavefunction, the various spin orbitals are varied and the expected Hamiltonian is constructed using the expanded determinant and evaluating individual energetic contributions. The energy derived from the HF approximation can be stated as:

$$E_{HF} = \langle \phi_{SD} | \hat{H} | \phi_{SD} \rangle = \sum_{i}^{N} (i|\hat{h}|i) + \frac{1}{2} \sum_{i}^{N} \sum_{j}^{N} [(ii|jj) - (ij|ji)] \tag{III..16}$$

$$\sum_{i}^{N} (i|\hat{h}|i) = \int \chi_i^*(\vec{x_1}) \{ -\frac{1}{2} \nabla^2 - \sum_{A}^{M} \frac{Z_A}{r_{1A}} \} \chi_i(\vec{x_1}) d\vec{x_1} \tag{III..17}$$

is evaluated to be the contribution to the kinetic energy and the energy due to attraction between electron and nuclei and

$$(ii|jj) = \iint |\chi_i(\vec{x_1})|^2 \frac{1}{r_{12}} |\chi_j(\vec{x_2})|^2 d\vec{x_1} d\vec{x_2} \tag{III..18}$$

$$(ij|ji) = \iint \chi_i(\vec{x_1}) \chi_j^*(\vec{x_1}) \frac{1}{r_{12}} \chi_j(\vec{x_2}) \chi_i^*(\vec{x_2} d\vec{x_1} d\vec{x_2} \tag{III..19}$$

the two integrals are Coulomb and exchange integrals respectively. The Hartree fock energy (E_{HF}) is determined as a functional of the spin orbitals and hence, the choice of orbitals determines the energy of the system. The HF equations are stated as eigenvalue problems similar to the Schrodinger equation in which the orbitals take the place of the overall wavefunction and the orbital energies for the overall energy, respectively. The Fock operator employed in this equation is in fact, a one electron operator composed of the kinetic energy, sum of potential energies (between a given electron and all the nuclei) and the HF potential, which gives the repulsive potential experienced by a certain electron due to the presence of all the other electrons.

$$\hat{f_i} = -\frac{1}{2} \nabla_i^2 - \sum_{A}^{M} \frac{Z_A}{r_{1A}} + V_{HF}(i) \tag{III..20}$$

For evaluating the Coulombic integral, the repulsion is calculated by consider-

ing a reference electron and another one at positions 1 and 2 over space and is also weighted by the associated probability of finding the second electron at the given position in space. The application of this integral depends solely on the value of the orbital function χ_i at $\vec{x_1}$ and is thus termed local. On the contrary, the terms involved in the exchange integral lead to the χ_i being dependent on the position and spin of the second electron as well, making it non-local. These contributions only exist for electron of like spin bearing in mind the antisymmetry principle. The key idea here is that the evaluation of the Fock operator itself depends on the spin orbitals or the solution to the Fock equation. Hence, this problem is of an iterative nature and the technique used to solve this is called the self-consistent field method. The orbital energies are derived from their own effective potential in a self-consistent manner. The method begins with a set of guessed orbitals, ψ_{trial} and the HF equations are solved using these orbitals in the first stage. The resulting set of orbitals obtained from this stage serve as the initial guess for the next iteration and then the orbitals are compared until its less than a predetermined threshold.

It is important to note that the Slater determinant never corresponds to the exact wavefunction solution to the problem. The energy calculated by the guess set of orbitals is always larger than the exact energy of the ground state E_0. The correlation energy is defined as the difference between the estimated and this ground state energy.

$$E_C^{HF} = E_0 - E_{HF} \tag{III..21}$$

In the scheme of electronic structure calculations, it serves as the error that is introduced by the basic HF technique. One of the reasons behind this is the instantaneous repulsion between electrons, which is not accounted for by the HF potential. This electron correlation can be either static or dynamic depending depending on whether the error arises due to the state or dynamic effects of electrons. Because of the size of molecules and the number of functions needed to

estimate their behavior, it is not possible to use wavefunction as the basic variable for calculations. To address this problem, the electron density is used as the flagship quantity upon which all the other computations are directly dependent. It is convenient and only a function of the physical space. Some of the key take aways for electron density are:

(1) $\int \rho(\vec{r_1})d\vec{r_1} = \text{N}$ which essentially means that density, when integrated over all physical space yields the number of electrons in the system.
(2) $\rho(\vec{r})$ has maxima placed at the positions of the nuclei in space.
(3) The electron density at the position of the individual nuclei contains all the necessary information about the nuclear charge Z. These three facts together are necessary and sufficient conditions for setting up a Hamiltonian for a certain system and when solved can yield all necessary molecular properties.

In terms of estimating the electron density, Thomas and Fermi put forth a model based on a quantum statistical description taking into consideration only the kinetic energy of the Fermions (or electrons) while the other two associated terms, i.e. nuclear-electron and electron-electron energies are treated in a classical manner [114, 115]. Based on the assumption of an uniform electron gas, the expression for the kinetic energy involving just the electron density is :

$$T_{TF}[\rho(\vec{r})] = \frac{3}{10}(3\pi^2)^{\frac{2}{3}} \int \rho^{\frac{5}{3}}(\vec{r})d\vec{r} \tag{III..22}$$

which when combined with the terms for attractive nuclear-electron potential and repulsive electron-electron potential gives the total energy of an atom solely dependent on the electron density,

$$E_{TF}[\rho(\vec{r})] = \frac{3}{10}(3\pi^2)^{\frac{2}{3}} \int \rho^{\frac{5}{3}}(\vec{r})d\vec{r} - Z\int \frac{\rho(\vec{r})}{r}d\vec{r} + \frac{1}{2}\iint \frac{\rho(\vec{r_1})(\vec{r_2})}{r_{12}}d\vec{r_1}d\vec{r_2} \tag{III..23}$$

3.0.1.2 Hohenberg-Kohn Theorems and Density Functional Theory

The entire density functional theory is formulated from the two basic theorems laid out by Hohenberg and Kohn in their paper in 1964. The first theorem states that '$V_{ext}(\vec{r})$, which is the external potential formed by the nuclei, is a unique functional of the electron density $\rho(\vec{r})$(to within a constant) ; and in turn, the full many particle ground state Hamiltonian is also a unique functional of it'. The energy of the ground state of a molecule can be expressed as :

$$E_0[\rho] = T[\rho_0] + E_{ee}[\rho_0] + E_{Ne}[\rho_0] \tag{III..24}$$

The third term of the right hand side may also be expressed as $E_{Ne}[\rho_0] = \int \rho_0 V_{Ne} d\vec{r}$ which is dependent on the composition of the actual system and the kinetic energy along with the electron-electron repulsion are not dependent on N, R_A or Z_A. $T[\rho_0]$ and $E_{ee}[\rho_0]$ are collectively called the Hohenberg-Kohn functional $F_{HK}[\rho_0]$. For a certain electron density, the wavefunction that gives the lowest energy can be used to find the value of this functional. The HK functional is a function of the density and can be expressed as an expectation value of the kinetic and electron-electron energy operator when applied to the ground state wavefunction.

The second theorem is dedicated to the evaluation of the electron density and the convergence of the problem towards the ground state energy. It simply states that for any trial density, $\tilde{\rho}(\vec{r})$ that satisfies the two essential conditions, $\tilde{\rho}(\vec{r}) \geq 0$ and $\int \tilde{\rho}(\vec{r}) d\vec{r} = N$, and which is part of a trial external potential $\widetilde{V_{ext}}$ always gives an upper bound to the actual ground state energy. The ground state energy is only obtained if the true ground state density is applied to the relationship :

$$E_0 \leq E[\tilde{\rho}] = T[\tilde{\rho}] + E_{Ne}[\tilde{\rho}] + E_{ee}[\tilde{\rho}] \tag{III..25}$$

This is known as the variational principle upon which the self consistent ap-

proach is based and the techniques for estimating the electron density will be discussed in the remaining part of the section. These two theorems mentioned above allow the representation of the many body function using just the electron density as the key quantity instead of the wavefunction. The ground state energy can be expressed as :

$$E_0 = \min_{\rho \to N}(F[\rho] + \int \rho(\vec{r}) V_{Ne} d\vec{r}) \tag{III..26}$$

In the equation, the universal Hohenberg-Kohn functional $F[\rho]$ contains the terms representing the kinetic energy, the Coulombic interaction of classical nature and the vital quantum portion due to exchange and correlation effects,

$$F[\rho(\vec{r})] = T[\rho(\vec{r})] + J[\rho(\vec{r})] + E_{ncl}[\rho(\vec{r})] \tag{III..27}$$

Among the terms involved in calculating the universal HK function, only the $J[\rho(\vec{r})]$ is known explicitly while the other two are estimated by using appropriate orbital guesses. To this end, the Slater determinant ϕ_{SD} is looked upon as the exact wavefunction describing a considered system of N non-interacting electrons (not participating in Colulombic repulsion) that move under the influence of the effective Hartree-Fock potential. Using the above fact, the kinetic energy can be written as

$$T_{HF} = -\frac{1}{2} \sum_{i}^{N} \langle \chi_i | \nabla^2 | \chi_i \rangle \tag{III..28}$$

In the next step, the Hamiltonian is divided to include an effective, non-interacting local potential

$$\hat{H}_s = -\frac{1}{2} \sum_{i}^{N} \nabla_i^2 + \sum_{i}^{N} V_s(\vec{r}_i) \tag{III..29}$$

which makes the ground state wavefunction Θ ,

$$\Theta_S = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_1(\vec{x_1}) & \varphi_2(\vec{x_1}) & ... & \varphi_N(\vec{x_1}) \\ \varphi_1(\vec{x_2}) & \varphi_2(\vec{x_2}) & ... & \varphi_N(\vec{x_2}) \\ ... & ... & ... & ... \\ ... & ... & ... & .. \\ \varphi_1(\vec{x_N}) & \varphi_2(\vec{x_N}) & ... & \varphi_N(\vec{x_N}) \end{vmatrix} \tag{III..30}$$

where the spin orbitals are determined from the eigenvalue relationship

$$f^{\hat{KS}}\varphi_i = \varepsilon_i\varphi_i \tag{III..31}$$

where $f^{\hat{KS}}$ is the effective one-electron Kohn-Sham operator that are distinct from HF orbitals. For estimating the kinetic energy accurately, the HK functional was separated into a non-interacting kinetic energy term, the $J[\rho(\vec{r})]$ and a third term known as the exchange-correlation energy E_{XC} , which is defined as the sum of the non-classical contribution to the kinetic energy and electronic repulsion.

$$E_{XC}[\rho] = (T[\rho] - T_S[\rho]) + (E_{ee}[\rho] - J[\rho]) = T_c[\rho] + E_{ncl}[\rho] \tag{III..32}$$

Hence, the energy of the interacting systems can be expressed as below.

$$E[\rho(\vec{r})] = T_s[\rho] + J[\rho] + E_{XC}[\rho] + E_{Ne}[\rho] \tag{III..33}$$

$$= T_s[\rho] + \frac{1}{2}\iint \frac{\rho(\vec{r_1})\rho(\vec{r_2})}{r_{12}} + E_{XC}[\rho] + \int V_{Ne}\rho(\vec{r})d\vec{r} \tag{III..34}$$

$$(-\frac{1}{2}\nabla^2 + [\int \frac{\rho\vec{r_2}}{r_{12}}d\vec{r_2} + V_{XC}(\vec{r_1}) - \sum A^M \frac{Z_A}{r_{1A}}])\varphi_i = \tag{III..35}$$

$$(-\frac{1}{2}\nabla^2 + V_{eff}(\vec{r_1}))\varphi_i = \varepsilon_i\varphi_i \tag{III..36}$$

As state above, the exchange correlation potential does not have an explicit form. Hence, it is just defined as a derivative of the associated energy with respect to the electronic density ,

$$V_{XC} = \frac{\delta E_{XC}}{\delta \rho} \tag{III..37}$$

The main steps in the so called Kohn Sham approach for calculating the ground state energies are the following:
(1) A non-interacting system of N particles is defined that acts as a reference whose exact ground state can de represented by a Slater determinant Θ_s and whose electronic density is said to be equal to the actual interacting physical system.
(2) The non-interacting kinetic energy can then be determined by solving the Kohn Sham equation whose solutions are elements of the Slater determinant in Step (1). The effective potential involved in the single electron Hamiltonian operator needs to be chosen so that the condition of the fictitious electron density, $\rho_s = \rho_0$.
(3) The energy related to the actual interacting system is parsed into the classical explicit T_s of the non-interacting system, the electron-electron repulsion energy J, nuclei E_{Ne}, the exchange correlation E_{XC} and the remainder of the kinetic energy not accounted for in the explicit T_s term.
(4) The variational principle is invoked in the expression in step (3). With this, the effective potential is calculated as a sum of the nuclear potential, V_{Ne}, the Coulombic potential (V_C) and the exchange-correlation potential E_{XC}.
(5) Solving for V_S and by evaluating the one-electron equations, the Kohn Sham orbitals may be obtained. These then define the non-interacting system said to have the same density as the actual interacting system. As emphasized before, the exchange correlation functional is unknown as has to be estimated to calculate the ground state energy.

All exchange-correlation functionals are dependent on the idea of an uniform electron gas. It is a hypothetical system on which the negatively charged electrons

move such that the net ensemble is neutral. In this, the number of electrons and the volume of the gas approach infinity, but their ratio remains finite., i.e. $N \rightarrow \infty$, $V \rightarrow \infty$, $N/V = \rho$. It can be visualized as a metal consisting of a perfectly formed crystal having valence electrons and their counterpart as positive cores that are smeared to create a background charge (local density approximation). This background is said to have a uniform distribution and the exchange-correlation energy can be expressed as

$$E_{XC}^{LDA}[\rho] = \int (\rho(\vec{r})\varepsilon_{XC}(\rho(\vec{r}))d\vec{r} \tag{III..38}$$

$$\varepsilon_{XC}(\rho\vec{r}) = \varepsilon_X(\rho(\vec{r})) + \varepsilon_C(\rho(\vec{r})) \tag{III..39}$$

$$\tag{III..40}$$

Here the $\varepsilon_{XC}(\rho(\vec{r}))$ stands for the exchange-correlation energy belonging to a single particle within the aegis of the uniform gas $\rho(\vec{r})$. Out of the two terms (exchange and correlation), the exchange part has an explicit form while the correlation part does not have any such term. If the local density approximation is extended to the situation where the spin of the molecule is polarized, then the E_{XC} is a function of both density and the spins. In this case, the local spin-density approximation can be expressed as

$$E_{XC}^{LSD}[\rho_\alpha, \rho_\beta] = \int \rho(\vec{r})\varepsilon_{XC}(\rho_\alpha(\vec{r}), \rho_\beta(\vec{r}))d\vec{r} \tag{III..41}$$

$$\xi = \frac{(\rho_\alpha(\vec{r}) - \rho_\beta(\vec{r}))}{\rho\vec{r}} \tag{III..42}$$

where ξ is the spin polarization parameter that takes the values 0 and 1 for spin compensated and fully polarized systems respectively. Most of the systems in computational chemistry do not enjoy the situation where the electron density is

constant everywhere, hence is change w.r.t. position needs to be accounted. The previous rationale is then enhanced to also include the gradient of the electron density at a certain position within the system. This is popularly known as a generalized gradient approximation and is expressed as

$$E_{XC}^{GGA}[\rho_\alpha, \rho_{beta}] = \int f(\rho_\alpha, \rho_{beta}, \nabla\rho_\alpha, \nabla\rho_{beta})d\vec{r} \tag{III..43}$$

$$E_{XC}^{GGA} = E_X^{GGA} + E_C^{GGA} \tag{III..44}$$

The exchange part among the two terms in the second equation above can be written as

$$E_X^{GGA} = E_X^{LDA} - \sum \sigma \int F(s_\sigma)\rho_\sigma^{4/3}(\vec{r})d\vec{r} \tag{III..45}$$

$$s_\sigma\vec{r} = \frac{|\nabla\rho_\sigma(\vec{r})|}{\rho_\sigma^{4/3}\vec{r}} \tag{III..46}$$

The s_σ is known as the reduced density gradient which will be discussed in detail in the following section. This parameter will be used extensively when weak intermolecular interactions will be explained in the Results for the different active site models constructed in the present work. The exchange correlation energy can also be expressed as an integral of the non-classical contribution E_{ncl}^{λ}. Here, at $\lambda = 0$, the system is said to be non-interacting and only consists of the exchange contribution. It can be computed accurately by applying the Kohn Sham orbitals. For $\lambda = 1$, in a fully interacting mode, both the exchange and correlation parts are included. The non-classical contribution is not exactly known at intermediate values of the interacting parameter. Hence, it is approximated in many different ways, one of which is the half and half combination of both the exchange and correlation parts given by Becke. In the next level, semiempirical coefficients were introduced to determine the weights in the following equation where E_X^B is

Becke's 1988 functional and E_c^{PW91} is Perdew and Wang's 1991 correlation functional [116]. (a,b and c are parameters used to fit data from various molecule sets).

$$E_{XC}^{HH} = \frac{1}{2}E_{XC}^{\lambda=0} + \frac{1}{2}E_{XC}^{\lambda=1} \tag{III..47}$$

$$E_{XC}^{B3} = E_{XC}^{LSD} + a(E_{XC}^{\lambda=0} - E_X^{LSD}) + bE_X^B + cE_c^{PW91} \tag{III..48}$$

In the same vein, the popular B3LYP functional [117] used in this work (both with and without the dispersion correction) was proposed that has been show to perform well for transition metal containing systems.

$$E_{XC}^{B3LYP} = (1-a)E_X^{LSD} + aE_{XC}^{\lambda=0} + bE_X^{B88} + cE_c^{LYP} + (1-c)E_c^{LSD} \tag{III..49}$$

3.0.2 Treatment of Dispersion

One of the challenges of working with DFT has been its inability to tackle dispersion interactions. In our system, majority of the interactions are due to formation of instantaneous dipoles due to the charge distribution which make them particularly typical for treatment with traditional DFT. In this section, the nature of dispersion interactions, the shortcomings of DFT treatment and appropriate corrections will been discussed. The dispersive forces are attractive forces that act between molecules that are said to arise from electron correlation of separate subsystems within a complex. As they have a pure correlational nature, they cannot be approximated by HF theory for supermolecules. The leading term in the dispersion energy put forth by Drude in his model is represented as

$$U_{dis} = \frac{-3\alpha^2 E_1}{4(4\pi\epsilon_0)^2 r^6} = \frac{C_6}{r_6} \tag{III..50}$$

where the polarizability and ionization energy of the molecule are α and E_1. The constant term originates from the interaction between the induced dipole on one molecule and the instantaneous dipole on another. At large separations, the leading term consisting of C_6 dominates the attractive part of the potential. At shorter separations, when electron clouds tend to overlap, other significant effects arise due to the distortion of the individual charge clouds of the molecules. The inadequacy arises because modern DFT describes these correlation effects at the ground state with reasonable accuracy but are unable to do so for long range effects debitum the local density approximation. The LDA estimates the exchange-correlation functional at shorter distances and the general concept being that the sum of two distant charge distributions is the overall sum of their own contributions. A detailed analysis presented by Pulay and co-workers [118] for local and semi-local theories points out that they fail to properly describe dispersive interaction near the vDW minimum, though being highly effective for the repulsive part of the potential. Therefore, for weakly interacting systems, the E_{XC} term needs to be enhanced to produce hybrid methods incorporating the above explained long-range effects. In this direction, Grimme and co-workers [119, 120, 121] published their seminal work on corrections to the traditional DFT to include dispersion theory. The general form of this energy is an extension of the model given by Drude and can be represented as

$$E_{disp}^{DFT-D} = -\sum_{AB}\sum_{n=6,8,10...} s_n \frac{C_n^{AB}}{R_{AB}^n} f_{damp}(R_AB) \qquad \text{(III..51)}$$

which is based on a pairwise treatment on all atoms in the system and is added to the Kohn-Sham DFT energy. The term C_n^{AB} represents the n-th order dispersion coefficient for a given atom pair AB, with R_{AB} as the distance between the two nuclei. For avoiding errors at intermediate distances, certain damping functions are used that dictate the range of dispersion corrections. A typical expression for such a function is given as

$$f_{damp}(R_{AB}) = \frac{1}{1 + 6(R_{AB}/(s_{r,n}R_0^{AB}))^{-\gamma}} \quad \text{(III..52)}$$

$$C_n^{AB} = \sqrt{C_n^A C_n^B} \quad \text{(III..53)}$$

$$R_0^{AB} = R_0^A + R_0^B \quad \text{(III..54)}$$

where R_0^{AB} is the vDW cutoff radius for atom pair AB, $s_{r,n}$ is a scaling factor depending on the functional theory employed and γ is a constant determining the steepness of the function at small internuclear distances.

3.0.3 The M06 functional

The M06 functional has been shown to be very effective [122, 123, 124] in treating non-covalent interactions for a variety of molecular systems. In this section, the theory related to this functional will be discussed. The parts related to the local spin for the M06 functional depends on the spin density (ρ_σ), spin related to kinetic energy density (τ_σ)and reduced spin density (x_σ). Out of these, the kinetic energy density given by τ_σ can be expressed as

$$x_\sigma = \frac{|\nabla \rho_\sigma|}{\rho_\sigma^{4/3}} \quad \text{(III..55)}$$

$$\tau_\sigma = \frac{1}{2}\sum_i^{occup} |\nabla \psi_{i\sigma}|^2, \sigma = \alpha, \beta \quad \text{(III..56)}$$

The functional form of the overall M06 functional is a linear combination of the PBE exchange model and the local spin density exchange

$$E_X^{M06} = \sum_{\sigma} \int dr[F_{X\sigma}^{PBE}(\rho_\sigma, \nabla\rho_\sigma) f(w_\sigma) + \varepsilon_{X\sigma}^{LSDA} h_X(x_\sigma, z_\sigma)] \quad \text{(III..57)}$$

$$f(w_\sigma) = \sum_{i=0}^{m} a_i w_\sigma^i \quad \text{(III..58)}$$

$$w_\sigma = (t_\sigma - 1)/(t_\sigma + 1) \quad \text{(III..59)}$$

$$t_\sigma = \tau_\sigma^{LSDA}/\tau_\sigma \quad \text{(III..60)}$$

where $f(w_\sigma)$ is defined as the spin kinetic energy density enhancement factor [125]. The correlational functional for M06 (meta GGA) is expressed piecewise by combining parallel-spin and opposite-spin and the total energy is given by

$$E_c = E_c^{\alpha\beta} + E_c^{\alpha\alpha} + E_c^{\beta\beta} \quad \text{(III..61)}$$

The non-linear parameters involved in the calculation of these three energy terms have been evaluated in their previous works by Truhlar and co-workers [126]. The training function used to derive these parameters involved a variety of databases consisting of thermochemistry, interaction and excitation energies.

3.0.4 Optimization and SCF Calculations

In the typical geometry optimization scheme, for each of the electrons the orbitals are optimized in a potential field of all the other particles. This step is repeated until the difference in energy obtained from the previous set of orbitals and the current energy is less than a pre-determined threshold. In the simple HF model, the calculated wavefunction for a system containing n-electrons is represented by a single Slater determinant, $|\psi_0\rangle$ of n spin orbitals ϕ_i

(1) $|\psi_0\rangle = |\phi_1\phi_2.....\phi_n\rangle$

The orbitals from the previous step are optimized in accordance to the variational principle so that the energy in Step (2) is minimized.

(2) $E_0 = \langle\psi_0|\hat{H}|\psi_0\rangle \,/\, \langle\psi_0|\psi_0\rangle$

in which $\hat{H}$ denotes the all-electron Hamiltonian. Now, this equation is decomposed into functions involving single electrons

(3) $\hat{F}(i)\phi(\tau_i) = \epsilon\phi(\tau_i)$

where $\hat{F}(i)$ represents an effective one-electron Hamiltonian which is a sum of the kinetic operator, nuclear potential and the HF potential

(4) $\hat{F}(i) = \hat{h}(i) + \hat{V}^{HF}(i)$

where $\hat{h}(i) = -\frac{1}{2}\nabla_i^2 - \sum_A^M \frac{Z_A}{r_{iA}}$ and $V^{\hat{HF}}(i) = \sum_j \hat{J}_j(i) - \hat{K}_j(i)$

and

$$\hat{J}_j\phi_i(1) = \int \phi_j(2)\frac{1}{r_{12}}\phi_j(2)\phi_i(1)d\tau_2 \qquad \text{(III..62)}$$

$$\hat{K}_j\phi_i(1) = \int \phi_j(2)\frac{1}{r_{12}}\phi_i(2)\phi_i(1)d\tau_2 \qquad \text{(III..63)}$$

where ∇_i^2 is the KE operator for electron i, Z_A is the nuclear charge on nucleus A, r_{iA} is the separation between nucleus A and electron i and $V^{\hat{HF}}(i)$ is the average potential felt by one electron due to a field created by all the others.

To reduce the complexity of the eigenvalue equation expressed in step (3), each of the spin orbitals are written as a linear combination of separate one-electron basis functions χ_ν

(5) $\phi_i = \sum_\nu^N C_{\nu i}\chi_\nu$

which is commonly referred to as the restricted HF method as all of the orbitals are subjected to a condition to have the same spatial description. Upon

substituting the relationship in Step (5) into Step (3) gives

(6) $\sum_{\nu}^{N} \hat{F}\chi_{\nu}C_{\nu i} = \epsilon_i \sum_{\nu}^{N} \chi_{\nu}C_{\nu i}$

which can be decomposed further by multiplying χ_{μ} and upon integration yields

(7) $\sum_{\nu}^{N}(F_{\mu\nu} - \epsilon_i S_{\mu\nu})C_{\nu i} = 0,\ i = 1, 2, ...N$

where $F_{\mu\nu}$ and $S_{\mu\nu}$ are the integrals over the space consisting of μ and ν
The matrix involved in the previous step can also be written in terms of the combination of the density matrix, one and two electron integrals $(h_{\mu\nu} and \mu\lambda||\nu\sigma)$

(8) $F_{\mu\nu} = h_{\mu\nu} + \sum_{\lambda\sigma}(\mu\lambda||\nu\sigma)P_{\lambda\sigma}$ where
(9) $P_{\lambda\sigma} = \sum_i C^*_{\lambda i}C_{\lambda i}, h_{\mu\nu} = \int \chi_{\mu}(1) * \hat{h}(i)\chi_{\nu}(1)d\tau_1$

$(\mu\lambda||\nu\sigma) = \int \chi_{\mu}(1) * \chi_{\lambda}(2) * \frac{1}{r_{12}}(\chi_{\nu}(1)\chi_{\sigma}(2) - \chi_{\sigma}(1)\chi_{\nu}(2))d\tau_1 d\tau_2$

that reduces the orbital expansion to matrix equations given by $FC = SC\epsilon$ where F is the basic Fock operator, molecular orbital coefficient matrix is given by C , the overlap matrix is represented by S and ϵ is the diagonal matrix consisting of the individual electron orbital energies. The matrix equation above becomes nonlinear w.r.t the molecular orbital coefficients as both the operators $\hat{J}$ and $\hat{K}$ in the Fock matrix depend on C. The initial Fock matrix is constructed by employing an initial guess for these orbital coefficients and the equation $FC = SC\epsilon$ is solved which are used to recompute the Fock matrix. This cycle is performed until the difference between orbital energies and molecular orbitals from consecutive steps reaches a tolerance value. This forms the basis for the self consistent field approach where every electron is influenced by an average potential (known as the SCF) .

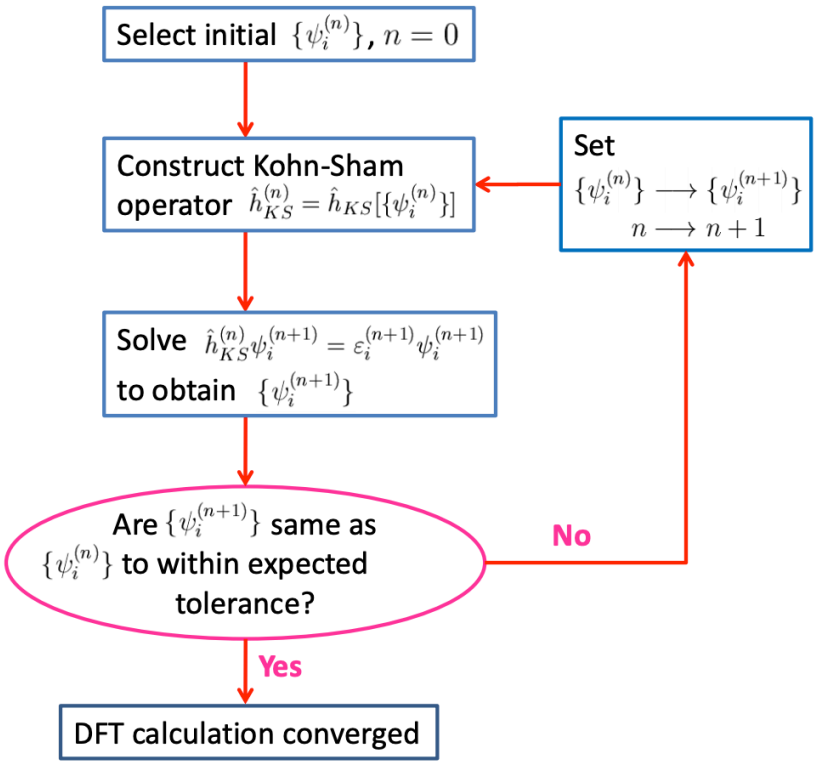


Figure 3.1: Steps involved in the SCF procedure with DFT (adapted from [5])

3.0.5 Basis Sets

Performing any structure calculation involves description of the space that is covered by the system considered. To construct the orbital matrices for a certain system, individual particle functions are needed. Mathematically, a basis set can be expressed as a unique collection of vectors that define a certain space in which a problem is conceptualized. Basis sets in quantum chemistry refer to collection of functions used to construct molecular orbitals. In this direction, linear combination of atomic orbitals is a key concept which simply translates to the fact that molecular orbitals can be constructed from atomic orbitals. These atomic orbitals are represented by Gaussian functions or by Slater-type functions. A basic Slater type orbital is given by the expression //

$$\phi_{abc}^{STO}(x,y,z) = Nx^a y^b z^c e^{-\zeta r} \tag{III..64}$$

where N is a normalization constant, the exponents a,b and c dictate the order of the function where the angular momentum is $L = a + b + c$, ζ denotes the width of the orbital and r the radial distance. Although, traditionally they were able to mimic the atomic orbitals for atoms like H, however when the orbitals are centered on different atoms, they fail to give accurate results. To overcome this, Gaussian type functions were used that have the functional form

$$\phi_{abc}^{GTO}(x,y,z) = Nx^a y^b z^c e^{-\zeta r^2} \tag{III..65}$$

These are advantageous to STO's as the Gaussian product theorem can be applied to them which states that product of two Gaussians can be expressed as a single Gaussian, that is located along the straight line joining them. This linear combination is essentially called a contraction which is a collection of Gaussian primitives having its coefficients and exponents fixed. The two forms of contraction are general and segmented. The first one allows for each of the present primitives to appear in each of the possible basis functions while for segmented, one specific primitive only occurs in a single contraction. The most basic minimal basis set contains only one basis function (either Slater or Gaussian) for each orbital centered on an atom. Double and triple zeta have two and three basis functions for each orbital, respectively. These were enhanced to give more weightage to the valence shell of the atoms rather than the core. In split valence basis sets, more functions are assigned to the valence orbitals as compared to the core atomic orbitals. STO-nG is the most popular minimal basis set where n gives the numbers of functions (or primitives) in the resulting contraction. Some of the notations used for describing basis set involves elaborating the number of functions for each orbital contraction. As an example, (63111,4311,1) represents the basis set where 5 s-type contractions are present that consist of 6,3,1,1 and 1 primitives. The p-orbital has 4 such contractions with 4,3,1, and 1 primitives with the d-orbital

having only one primitive. In this direction, Pople and co-workers [127, 128] formulated a convention in which the structure was given for the whole molecule, rather than being centered on an atom.

3.1 Molecules

All the electronic structure calculations performed in the work involved both charged and uncharged molecules. The structure of these molecules were initially constructed in the Avogadro and a pre-optimization was performed on them using the tool present in the graphics program. The molecular models were constructed taking into consideration the fact that the ionic liquid cation is in complexation with the macromolecule present in the active site of the protein. Hence, the part of the model dedicated to the protein was created in accordance with the crystal structure of available cytochrome P-450 BM3 proteins. Most of the calculations presented in the work have been performed on some variant of the central active site molecule of P-450 i.e. the heme. It is a coordination complex that consists, at its center, an iron atom (with variable oxidation state) attached to a porphyrin (which forms four coordination sites). In addition to this iron porphyrin, one or more axial ligands are attached that serve as chemical descriptors, reaction sites or key groups responsible for triggering mechanisms for the P450 biocatalyst. In nature, the heme is an important constituent of biologically relevant proteins. Some of these proteins are invested in various functions such as intracellular transport, electron source, detection of diatomic gases among others. Figure 3.3 shows the structure of Heme (without added hydrogens) consisting of the iron atom connected to four pyrrole rings. These rings are conjugated to methyl groups at two positions, two vinyl groups and two propionate chains. To start from a very basic structure, this heme was truncated for most of the gas phase DFT calculations and reduced to just iron porphyrin (FeP). One more simpler variant was considered in the initial calculations which had two hydrogens in place of the central iron atom, the so called free base porphyrin (FBP). Figure 3.4 gives the two models used for the initial DFT calculations performed for the active site in gas phase.

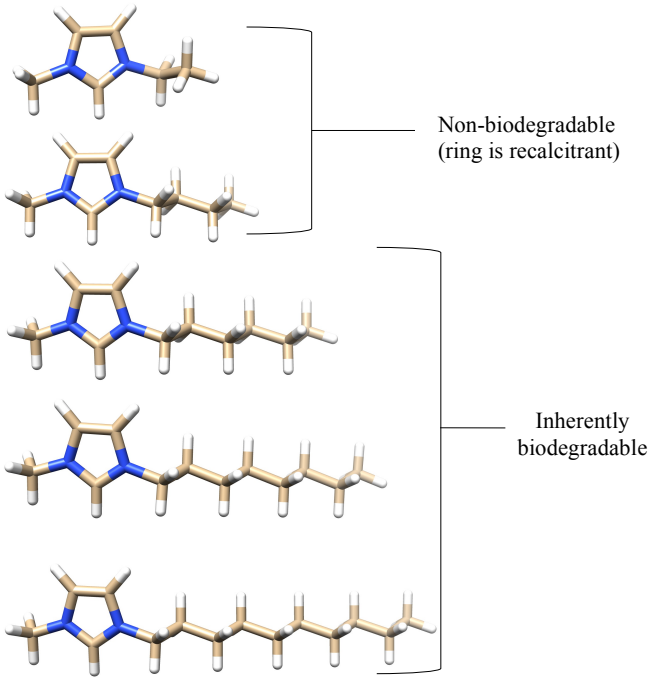


Figure 3.2: Imidazolium cations considered in the work as substrates (from top to bottom, $[C_2mim]^+$, $[C_4mim]^+$, $[C_6mim]^+$, $[C_8mim]^+$ and $[C_{10}mim]^+$

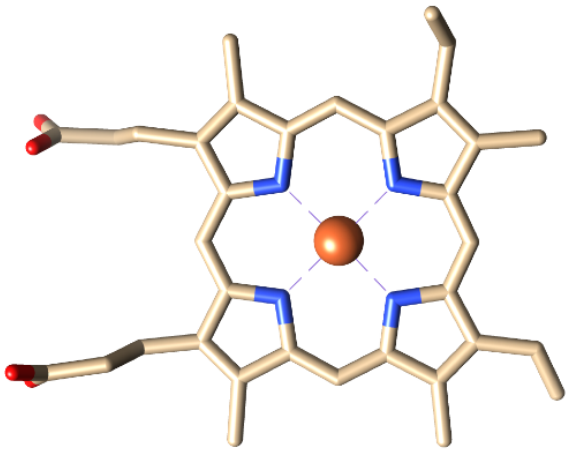
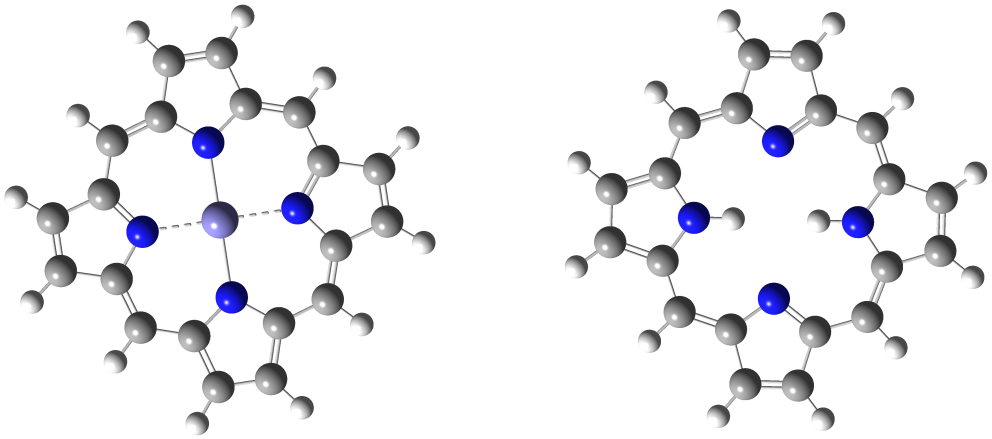


Figure 3.3: Structure of heme consisting of all side groups and linkages



(a) Iron Porphyrin (b) Free base porphyrin

Figure 3.4: Figure showing molecular models used for gas phase calculations in the work

3.1.1 Evolution of Active Site Model

The first step in formulating the computational framework for biodegradability is to unearth the interactions between a biological entity and the molecule chosen to be investigated. In the present work, the biological molecule was the heme, as discussed in the previous section. The truncated heme was chosen as the key reactive species in the model. The reason behind this being that it is a part of cytochrome P-450's active site. By definition, an active site is the enzymatic region where binding of substrates yield suitable products. The formation of these products is dictated by the catalytic cycle of the concerned protein. The active site of cytochrome P-450 is unique in the fact that it is chemically self-sufficient and does not require external co-factors for undergoing local chemical processes. The local environment of the active site directly influences the pattern of the reaction pathway and progress. From an evolutionary point of view, each active site is optimized to bind specific substrate molecules and is in turn, highly specific to these molecules or moieties. The P-450 BM3 active site is well-defined and located safely within the protein and comprises less than 40 percent of the volume of the whole protein in most cases. Although, the binding space is well defined, the substrate considered in the work, the ionic liquid cation, is a non-native substrate to cytochrome P-450 and there is an absence of crystal structure with the particular small molecule embedded within it.

To overcome this, a reductionist approach to the whole modeling process was adopted. The first step of this approach involved calculating the ground states of the monomers involved in the binding i.e. iron porphyrin and ionic liquid cation using a pure QM description. The DFT calculations performed gave a structural and electronic description of the molecules in gas phase devoid of any external disturbances or contacts. The next step in the scheme was to combine the optimized monomers in the gas phase in a complex to gain insight into the binding and electronic coupling of the two species. Majority of the work presented in the thesis is dedicated to the results and discussion of the binding process. Although,

other steps in the catalytic cycle, such as reduction, oxidation and hydroxylation have also been addressed in some detail in order to understand them separately. In all subsequent steps from the first one, the description of the active site is enhanced progressively to approach the crystal structure of P-450 BM3. The second step consists of the simplest description of the active site, i.e. the iron porphyrin. To mimic the proximal site of P-450, this FeP molecule is ligated to a de-protonated cysteine residue and named FePCys thereafter. In the sub-study involving the catalytic cycle of P-450, the FePCys is mutated to mimic the different species in each catalytic step. As all pure electronic structure calculations have a size limitation associated with them, the next step in the description involved coupling the active site with the neighboring protein in a QMMM (quantum mechanical molecular mechanical) scheme. The details of the QMMM calculations will be provided in a separate section within the document. Also, subtle details for the calculations for metal-containing systems such as P-450 will also be provided.

3.1.1.1 Role of Key Residues

Arg47 and Tyr51 The role of these two residues has been shown initially in a palmitoyl derivative bound crystal structure where two polar residues were found near the mouth of the access channel. They regulate the insertion of substrates, water and co-solvents into the binding pocket acting as "gate-keepers" to the reactive center of the protein. Different substrates are lined in their unique manner and possible binding sites are dictated by the geometrical orientation provided by these two residues.

Phe87 This is the residue located closest to the actual haem site and is situated in between the co-factor and the substrate in some of the BM3 crystal structures. Its motion has implications in reducing distance of closest approach (distance between centers of two bodies when they are externally tangential to each other) and other forms of active site dynamics. The role of Phe87 is very powerful in dictating site-specificity for a myriad of substrates. Strong evidence suggests that it is responsible for stimulating the substrate to be oxidized/ reacted close to the

terminus rather than focus on its center.

Glu267 and Thr268 Threonine is one of the few polar components of the active site that is generally hydrophobic in nature. Its role has been connected to delivery of proton for the formation of species leading to the ultimate active compound (refer Figure 2.2 for details on each species) and their stabilization. The side-chain belonging to glutamate is connected to a water molecule through hydrogen bonding in some crystal structures and is also instrumental in proton delivery to the haem center.

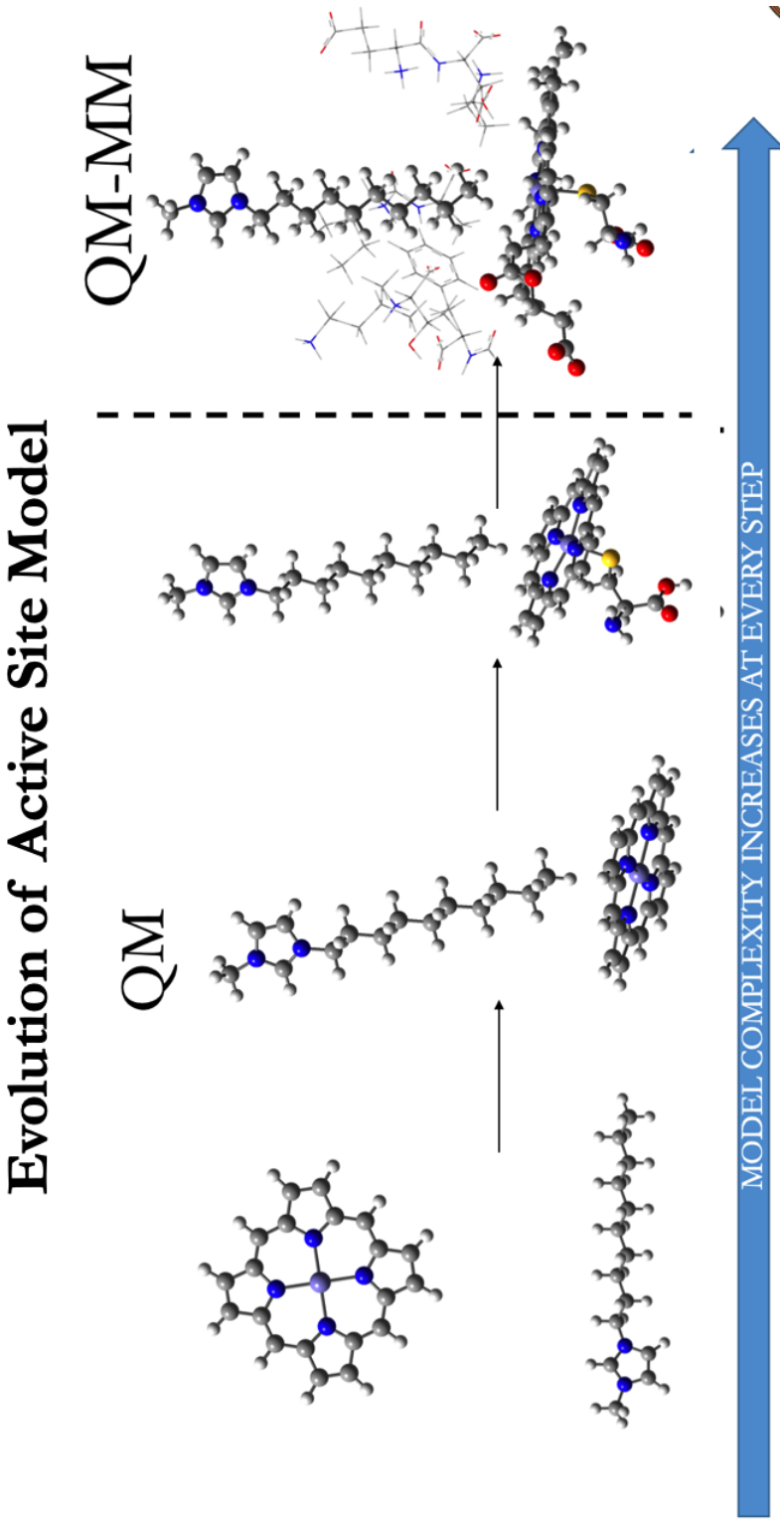


Figure 3.5: Diagram showing the evolution of the active site model addressed in the work

The systems under consideration consisted of 1-n-alkyl-3-methylimidazolium $[C_nmim]^+$ cations having variable alkyl chain lengths (n = 2, 4, 6, 8, and 10). Geometry optimization of the cations and cations complexed with free base porphyrin (FBP) and iron porphyrin (FeP) were carried out in the gas phase using a quantum mechanical approach. For these calculations, a density functional theory (DFT) method incorporating a derived hybrid functional B3LYP [117] with the 6-31G(d,p) basis set [128, 127] was chosen taking into consideration relative computational cost and that the functional has been applied to porphyrin containing systems [129]. Moreover, this level of theory has been found adequate in calculating the optimized geometry and partial charges for a wide range of ionic liquid cations and anions. [130, 131, 132, 133] The calculations were performed using Gaussian 09 software package. [134]

3.1.2 Binding of Cation-Porphyrin Complexes

The binding energies between the imidazolium cation and the porphyrin molecules serve as an indicator of the stability of the overall complexation process. The binding energies of the ionic liquid-FBP and ionic liquid-FeP complexes were calculated using eq. III..66

$$\Delta E_{\text{binding}} = E_{\text{complex,cp}} - E_{\text{IL}} - E_{\text{FBP/FeP}}. \tag{III..66}$$

In this equation, $\Delta E_{\text{binding}}$ refers to the binding energy of the complex; $E_{\text{complex,cp}}$ denotes the counterpoise corrected energy associated with the gas-phase optimized geometries of the complex and E_{IL}, $E_{\text{FBP/FeP}}$ are the gas phase energies of ionic liquid cation, and FBP or FeP, respectively.

The binding process can be thought of as a two step process (Fig. 4.2): (1) The isolated moieties (ionic liquids, FBP, and FeP) undergo a conformational change from their respective optimized gas-phase geometry to that in the complex; (2) the interaction between the two binding partners (de)stabilize the complex. With this partitioning, the binding energy in eq. III..66 can be written in the following

form

$$\Delta E_{\text{binding}} = \Delta E_{\text{c}} + E_{\text{i}} + \Delta E_{\text{cp}}. \tag{III..67}$$

ΔE_{c} is the energy change associated with the change in the conformation energies of the two chemical entities participating in the binding process; E_{i} is the interaction energy between the ionic liquid and FeP or FBP, and ΔE_{cp} refers to the counterpoise correction. Further, each of the energy components can be obtained as follows

$$\Delta E_{\text{c}} = E_{\text{m,c}} - E_{\text{m,g}} \tag{III..68}$$

where $E_{\text{m,c}}$ and $E_{\text{m,g}}$ are the energies of the monomers (m; cation, FBP or FeP) in complex (c) and gas phase (g), respectively, which were determined by isolating the conformation of the monomer in the complex from its partner and performing single point energy calculations. This component serves as an indicator of the energy requirement to induce a conformational change in the monomers to achieve optimum binding.

The interaction energy between the ionic liquid and FBP or FeP was calculated with reference to $E_{\text{m,c}}$ and is given by

$$\Delta E_{\text{i}} = E_{\text{complex}} - E_{\text{m,cation}} - E_{\text{m,FBP/FeP}}. \tag{III..69}$$

Finally, the counterpoise correction contribution to the overall binding energy was computed using eq. III..70

$$\Delta E_{\text{cp}} = E_{\text{complex,cp}} - E_{\text{complex}} \tag{III..70}$$

where $E_{\text{complex,cp}}$ is the counterpoise corrected energy of the optimized complex. (Figure 4.2)

In order to compare the stability of the three conformers for a given cation, the relative binding energies were calculated with respect to the minimum binding

energy conformation (eq. III..71)

$$\Delta\Delta E = \Delta E_{\text{binding}} - \Delta E_{\text{binding,min}} \tag{III..71}$$

where $\Delta E_{\text{binding,min}}$ is the energy of the most stable complex, i.e., the one having the highest absolute binding energy.

3.1.3 Counterpoise Correction

The calculations of binding energies involves a correction to the overall interaction energies that is defined as counterpoise correction. When the intermolecular interaction is particularly weak and dominated by long range forces, the whole system experiences an artificial strengthening [135]. This arises due to the fact that the basis sets used to describe the individual participating species are small in comparison to a complete description (a complete basis set incorporating all the essential features of the individual atomic orbitals) . As the two species approach each other, species 1 has a tendency to use external basis functions from species 2 to enhance its local electron distribution. Thus, at short distances the basis set describing a certain species gets enhanced and at longer distances this effect is absent. Mathematically, the orbital overlap integrals are too low at longer separations and do not provide adequate stabilization [136, 137, 138]. This error is simply defined as the Basis Set Superposition Error (BSSE). One of the most effective way of removing this error is given by the seminal work by Boys and Bernardi. In an ideal situation where each species is perfectly described devoid of any errors, the interaction energy between species 1 and 2 would be calculated by the simple relationship:

$$\Delta E_{\text{i}}(12) = E_{12}^{12}(12) - E_{1}^{1}(1) - E_{2}^{2}(2) \tag{III..72}$$

where the superscripts denote the basis function used, the subscripts represent the species (geometry at that level) and the system considered is in parentheses. Thus, $E_{12}^{12}(12)$ is the energy of the 12 complex evaluated in the basis formed by the

union of individual basis of the two participating species. The second and third terms are the energies of the respective species in their individual basis sets. Now, the energy by which the individual species are artificially stabilized by the orbital overlap can be computed by taking the difference of the energy of each evaluated at the dimer basis and subtracting off its energy at the local basis.

$$E_{BSSE}(1) = E_1^{12}(1) - E_1^1(1) \tag{III..73}$$

$$E_{BSSE}(2) = E_2^{12}(2) - E_2^2(2) \tag{III..74}$$

In this situation it is assumed that the geometries of the individual species do not change as they approach each other to form the complex. In this case, the individual BSSE energies evaluated would give the stabilization of both 1 and 2 as they form a complex (the energies would be negative). If the deformation energies are included, then the difference needs to be evaluated with the geometry of say, 1 as present in the complex. So, the BSSE is computed as :

$$E_{BSSE}(1) = E_{1'}^{12}(1) - E_1^1(1) \tag{III..75}$$

$$E_{BSSE}(2) = E_{2'}^{12}(2) - E_2^2(2) \tag{III..76}$$

where 1' and 2' are the deformed geometries of species 1 and 2 respectively. It must be noted that if the deformation is not considered, the overall binding energy equals the interaction energy. Since, we need to considered the deformation of the species, the overall BSSE then evaluates to

$$\Delta E_{\text{binding}}^{CP}(12) = [E_{12}^{12}(12) - E_{12}^{12}(1) - E_{12}^{12}(2)] - [E_{12}^1(1) - E_1^1(1)] - [E_{12}^2(2) - E_2^2(2)] \tag{III..77}$$

which is basically the sum of the counterpoise correction arising due to interaction of rigid monomers in complex 12, the deformation of the first species and the deformation of the second species.

$$\Delta E^{CP}_{\text{binding}}(12) = E^{CP}_{int}(12) + E^{1}_{def}(1) + E^{2}_{def}(2) \qquad \text{(III..78)}$$

In this work, all of the reported binding energies are counterpoise corrected using the above described approach in which the interaction energy is corrected by this quantity and individual species energy is evaluated at infinite separation.

3.1.4 Population Analysis

The frontier orbitals play an important role in determining the ability of a molecule to lose (ionization energy) or gain (electron affinity) an electron and are important in biodegradation as these orbitals participate in reactions [139, 140]. Therefore, a frontier orbital treatment of the reactivity of all the systems was carried out. Specifically, the HOMO (Highest Occupied Molecular Orbital) and the LUMO (Lowest Unoccupied Molecular Orbital) energies and their variation with respect to the alkyl chain length were determined. In addition, the shifts in the HOMO and LUMO energy levels from their isolated gas phase state were investigated in the presence of FBP and FeP. Charge transfer between the cations and porphyrin molecules were characterized based on the ChelpG charges [141, 142] computed for the optimized geometries. The changes in the partial charge distributions were analyzed to account for the intermolecular charge transfer.

The catalytic cycle of oxidizing agent cytochrome P-450 has been extensively studied by Shaik and co-workers [143, 144, 145, 146, 147]. In the systems studied in the present work, the imidazolium cation in proximity to the porphyrin variants play the role of substrate in the overall catalytic cycle. To understand the role of the imidazolium cation in changing the electronic environment of the porphyrin receptor, it is essential to look at factors that are associated with the reactivity of

the latter. In order to assess the propensity for Fe^{3+} to Fe^{2+} reaction once the substrate, the imidazolium cation in this case, binds, electrophilicity was calculated as an indicator of porphyrin reactivity based on the frontier orbital energies. According to Parr et al., [148] the electrophilicity ω can be obtained from the knowledge of the electron chemical potential μ and the chemical hardness η (Eq. III..81). Electrophilicity is one of the prime indicators of reactivity that may be derived from frontier orbital energy levels. A definition regarding the reactivity of a species in association with another was proposed by Parr that connects reduction in energy with maximum electron flow between the two respective species. In the work, density functional theory was utilized to define global electrophilicity index as a combination of the frontier orbital energies (HOMO and LUMO) associated with different species. Parr's electrophilic ω index serves as a useful tool for understanding the scale of reactivity from a theoretical standpoint based on a simple relationship between the electronegativity and chemical hardness.

While μ may be defined as the chemical potential (i.e. the negative of electronegativity) calculated as the derivative of the change in total energy of a system containing a certain number of electrons, η is its corresponding derivative. The second derivative of total energy serves as an indicator of the change in behavior of the system to an external potential in the form of a density response function. The Koopman's theorem aids in defining the role of HOMO and LUMO energies in approximating reactivities of chemical species. To be realized on a simple theoretical scale, from the standpoint of frontier orbital theory, the same, they are calculated as follows :

$$\mu = \frac{E_{\mathrm{HOMO}} + E_{\mathrm{LUMO}}}{2} \tag{III..79}$$

$$\eta = E_{\mathrm{HOMO}} - E_{\mathrm{LUMO}} \tag{III..80}$$

$$\omega = \mu^2/2\eta \tag{III..81}$$

where E_{HOMO} and E_{LUMO} refer to the HOMO and LUMO energies for FBP or FeP.

3.1.5 Thermodynamics of the FeP and FBP Complexes

The binding energies between the imidazolium cation and the porphyrin molecules serve as an indicator of the degree of interactions between them. The binding energy was calculated as

$$E_{\text{binding}} = E_{\text{complex}} - E_{\text{cation}} - E_{\text{porphyrin}} \quad \text{(III..82)}$$

where E_{binding} is the energy change upon binding; E_{complex} is the energy of ionic liquid cation in complexation with either FBP or FeP; E_{cation} and $E_{\text{substrate}}$ are the energies of the cation and the porphyrins (FBP and FeP) in the gas phase obtained from respective optimizations. The binding energy was corrected for the basis set superposition error (BSSE).

3.1.6 NBO Analysis

A natural Lewis structure describes the occupation of electrons in a molecule. It is a basic layout for showing the bonding relationship between atoms of a molecule and the lone pairs of electrons in the molecule. But often, the Lewis structures of molecules need to account for external effects that are not captured by it. Some of these effects include resonance delocalization, backbonding and intermolecular H-bonding. From the natural bond orbital point of view, the complete basis set (Ω_{i}) can be separated into Lewis ($\Omega_{\text{i}}^{(L)}$) and non-Lewis ($\Omega_{\text{i}}^{(NL)}$) components,

$$\Omega_{\text{i}} = \Omega_{\text{i}}^{(L)} + \Omega_{\text{i}}^{(NL)} \quad \text{(III..83)}$$

Similarly, the total wave function Ψ can said to be composed of Lewis type ($\Psi^{(L)}$) and non-Lewis type ($\Psi^{(NL)}$) corrections of secondary nature :

$$\Psi = \Psi^{(L)} + \Psi^{(NL)} \quad \text{(III..84)}$$

The prediction of non-Lewis type interactions are dependent on the treatment of this secondary correction. This correction is dealt with as a perturbation to the original Lewis solution of the possible electronic interactions in the system. In terms of the general Schrodinger equation, the associated energy can be described by the equation having the Hamiltonian operator (H_{op})

$$H_{op}^{(L)}\Psi^{(L)} = E^{(L)}\Psi^{(L)} \tag{III..85}$$

This is the idealized model in which resonance type effects are absent where $E^{(L)}$ represents the natural Lewis structure eigenvalue. By expressing the above equations in a one-electron eigenvalue equation and by incorporating the NBOs from equation III..83, we get

$$h_{op}^{(0)}\Omega_i^{(L)} = \epsilon_i^{(L)}\Omega_i^{(L)}, i = 1, 2, 3.....N \tag{III..86}$$

where solving for the first N eigenfunctions $\Omega_i^{(L)}$ are the filled Lewis-type NBOs that have the corresponding orbital energies $\epsilon_i^{(L)}$ and the remaining eigenfunctions represent the non-Lewis types NBOs $(\Omega_i^{(NL)})$

$$h_{op}^{(0)}\Omega_j^{(NL)} = \epsilon_j^{(NL)}\Omega_j^{(NL)}, j = N + 1, ... \tag{III..87}$$

In the scheme of orbital interactions, the filled orbitals are 'donor' orbitals and the vacant NBOs are 'acceptor' orbitals, respectively. Mathematically, for these donor and acceptor NBOs, the effective one electron Hamiltonian operator F_{op} is non-zero expressed as :

$$F_{ij} = \int \Omega_i^{(L)*} F_{op} \Omega_j^{(NL)} d\tau \tag{III..88}$$

in which F_{ij} is the total sum of all one electron operators F_{op} over all such possible interactions. Now, for each possible donor-acceptor pair, the second order perturbation $\Delta E_{ij}^{(2)}$ gives the correction to the overall natural Lewis wavefunction. Figure 1 depicts the interaction between donor and acceptor NBOs ($\Omega_i^{(L)*}$ and $\Omega_j^{(NL)*}$). Their unperturbed energy levels are $\epsilon_i^{(L)}$ and $\epsilon_j^{(NL)}$ respectively. All of these energy levels are shown on the vertical scale in the figure. In the presence of the perturbation due to intermolecular interactions, these levels mix and split to give rise to a lower (ϵ_-) and upper (ϵ_+) energy levels. This mixing results in an overall energy stabilization i.e. both the electrons at the donor side are stabilized by partial delocalization into the acceptor NBO. This net perturbative energy lowering of the donor-acceptor pair can be expressed as:

$$\Delta E_{ij}^{(2)} = -q_i |F_{ij}|^2 / (\epsilon_j^{(NL)} - \epsilon_i^{(L)}) \qquad \text{(III..89)}$$

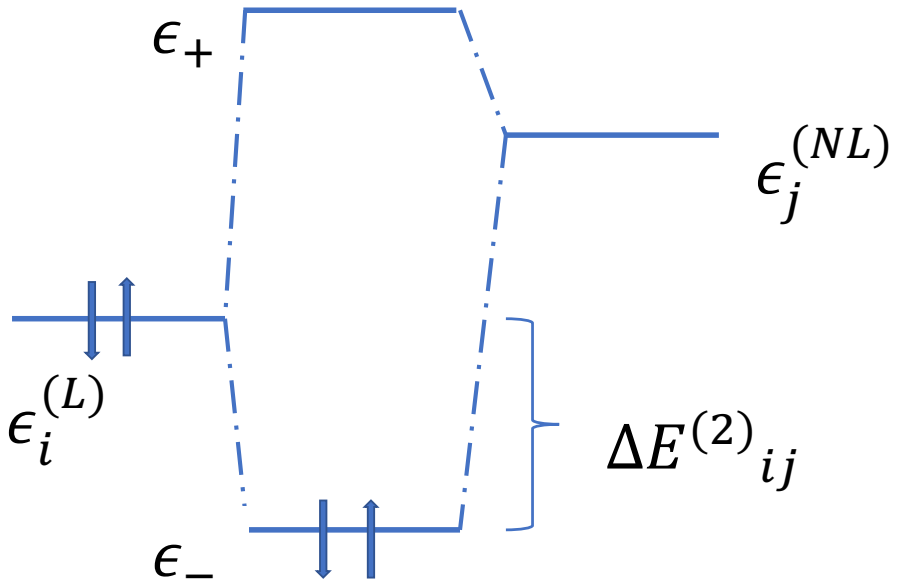


Figure 3.6: Stabilizing interaction between filled donor orbital $\Omega_i^{(L)}$ and unfilled acceptor orbital $\Omega_j^{(NL)}$ leading to lowering of the energy by $\Delta E_{ij}^{(2)}$

where q_i is the occupancy of the donor orbital, F_{ij} is the hamiltonian operator given by equation III..88 and $(\epsilon_j^{(NL)} - \epsilon_i^{(L)})$ represents the difference in the corresponding acceptor and donor orbital energy levels.

For our systems, the second order perturbative energy $(\Delta E_{ij}^{(2)})$ is of importance because it gives a quantitative treatment of the stabilization of the donor orbitals at the IL cation side by the acceptor orbitals belonging to porphyrin. The possible interactions and their order in magnitude will provide a site specific insight into the non-Lewis perturbation when different conformations are exposed to the porphyrin molecule.

3.1.7 Structural Parameters

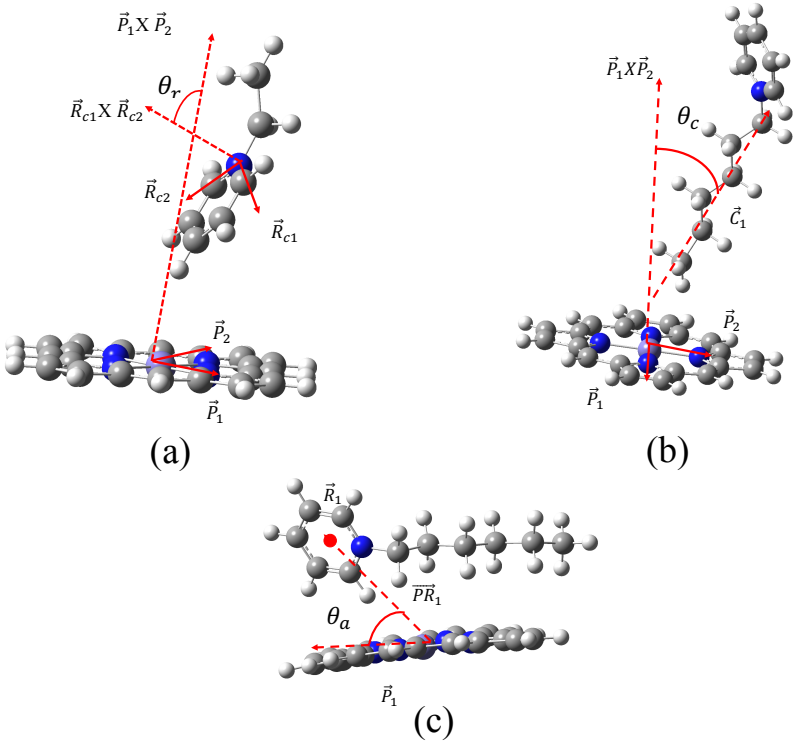


Figure 3.7: Structure parameters computed for different cations to quantify effect of optimization

3.1.8 Analysis of Weak Intermolecular Interactions

As emphasized in the previous sections, the electron density derived from quantum-mechanical treatment can be used to derive all key properties of interest for a molecule. In the same direction, the reduced density gradient (RDG) that can be evaluated using the density and its first derivative, acts as a descriptor of the inhomogeneity of the electronic distribution. Quantitatively, the RDG assumes very high values at regions away from the nuclei where the density decays exponentially. On the contrary, it approaches zero at regions exhibiting both non-covalent interactions and covalent bonding. Upon a simple examination of a RDG vs ρ plot, the region in the top left correspond to high RDG and low density while points on the right (low RDG and intermediate ρ) correspond to bonding region. It is important to note that covalent bonds show a characteristic saddle point in the RDG=0 isosurface. Although, the profile of ρ is a characteristic of the molecular system considered, each of the systems show spikes (or valleys) in the low density, low RDG region, as signatures of non-covalent interactions. In terms of its location in a complex, it corresponds to the intermolecular region between two monomers in electronic contact. It must also be noted that weak interactions within molecular systems occur in multiple perspectives, some of which are hydrogen bonding, vDW interactions and others.

To distinguish between these, density derivatives are employed. Bader and co-workers [149] established that the sign of the hessian of the electron density $\nabla^2\rho$ can be used to discriminate between different interaction modes. This Laplacian function is then parsed into 3 principal axes for assessing their maximal variation. These are simply the three eigenvalues of the above talked about Hessian, that can be expressed as $\nabla^2\rho = \lambda_1 + \lambda_2 + \lambda_3$. Mathematically, at nuclear positions, the electron density reaches a maxima, which renders all the three λ to be negative. Between bonded atoms, one of the eigenvalues assumes a positive value. For covalent regimes, the contribution from negative eigenvalues is dominant while for non-covalent interactions the positive λ dominate. This is irrespective of whether

bonding or nonbonding regimes are encountered. For bonding regions, the indermediate eigenvalue λ_2 assumes a negative value. If in a nonbonded situation, $\lambda_2 > 0$ in the region between atoms and the other two eigenvalues may assume positive or negative values. From the discussion above, the sign of the second eigenvalue λ_2 can be employed to distinguish between bonded and non-bonded noncovalent interactions. The electron density at these regions give an idea of the strength of the underlying interactions.

3.2 Molecular Docking

Molecular docking serves at the template for structure based design [150, 151, 152] and has a wide variety of applications in different kinds of industry. The purpose of docking is to use certain macromolecular targets or receptors and find matching small molecule or ligands that have favorable stereochemical and electrostatic attributes to achieve high binding affinity. This high affinity is expected to yield specific products [153, 154, 155] for use in synthesis and production of industrially valuable chemicals. The docking process begins with fixing a target structure [156, 157] and then predicting affinity of small molecules to it *in silico*. They become highly useful in solving crystallographic structures for intermolecular recognition. From this knowledge, several mechanistic features, conformational changes, binding sites can be established to aid in molecular design.

Most docking programs provide the user an ability to predict structures with a reasonable degree of accuracy to the macromolecular target site. They also provide predictions of the binding affinity and ranking for the docked compounds in the ligand-receptor complexes. Some of the key definitions and concepts for molecular docking will be discussed below.

In the first stage that is dedicated to the conformational search [158, 159, 160, 161, 162], the geometrical features of the associating ligand, comprising of all

the degrees of freedom are optimized. This optimization is achieved by different search methods [163, 164, 165] that might be stochastic or systematic in nature. Systematic algorithms explore the energetic landscape formed by gradually moving through the conformational space to find an appropriate minima. These suffer from problems such as being stuck in a local minima [166, 167, 168, 169] that might be overcome by supplying various initial guesses at different points along the landscape. On the other hand, stochastic methods focus on modifying the geometrical parameters of the ligand itself through random moves. Such an algorithm generates a suitable collection or ensemble of structures in order to populate a wide area within the energy landscape. It has a greater computational cost as compared to the systematic technique but avoids being stuck in a local minima. One of the most widely employed approaches for the systematic nature is the breakdown of the ligand into respective fragments and then regrowing it [170, 171, 172]. In this technique, the key part of the ligand is considered as an anchor and the rest of the fragments are joined in a sequential manner. This reduces the degrees of freedom to be sampled for the molecule and reduces the chances for a combinatorial explosion. Genetic algorithms are used to include stochasticity that use principles ordained from theory of evolution and mutations. In such a strategy, the first step is often the encoding of the geometrical parameters of the ligand into a vector [173, **?**]. This vector becomes the initial point for the random search to begin and populate the energy landscape. The candidates generated within this process reporting lowest energy values become template for the subsequent step. This recursive method usually converges to a vector that is the global minimum [174, 175, 176].

The second stage involves the evaluation of the binding energetics from the geometries that were obtained as a result of the conformational search. For this purpose, docking programs employ several scoring functions [177, 178, 179]. These scoring functions are directly related to the identification of the most important phenomena involved in the binding such as intermolecular interactions, entropic effects or desolvation [180, 181]. Three types of scoring functions are most popular, namely, force-field based, empirical and lastly, the knowledge-based ones. The first one estimates the energy of association between the receptor and the ligand by gath-

ering the contributions made from all kinds of bonded and non-bonded interactions. However, they lack the estimation of entropic and desolvation effects due to the absence of an appropriate physical model. Empirical scoring functions [182] comprise of individual terms dedicated to a certain physical effect present in the binding process often including hydrogen bonding, desolvation, entropic effects and others. These are developed by initially training on a well known data-set containing protein-ligand binding affinities [183, 184, 185, 186, 187]. The third approach is the knowledge based function that employs pairwise energy potentials from known intermolecular complexes to obtain functions [188, 189, 190]. They are mostly distance-based and evaluated by noting the frequency of interaction between a pair of atoms. The docking score from this method is simply the sum of all such interactions.

Now we briefly discuss some of the challenges faced in the prediction of structures by docking programs. The formation of covalent bonds [191, 192] is one such challenge that is overcome by different programs using a collection of strategies. One such method is the evaluation of an interaction energy map first and then using it as a guide for the conformational search. Another one is the determination of key attachment points within the system and using it as a pivot [193, 194, 195].

The fact that small molecules such as ligands are flexible make it important for the program to capture the structural changes undergone during binding [196, 197, 198]. The ligand influences the macromolecular target by shifting the latter's energy surface towards a collective minimum. In these situations, molecular dynamics [199, 200, 201, 202, 203, 204] can be utilized to generate an ensemble of starting structures that may be treated at a higher accuracy at a later stage [205, 206, 207].

The presence of cystallographic waters in the binding site poses another major challenge to the prediction of binding affinity [208, 209, 210]. They usually mediate hydrogen bonding between ligand and the receptor through proton networks [211, 212, 213]. The release of such a molecule is entropy driven but incurs a loss in enthalpic contribution to the overall energy [214]. For this reason, in the present work, the docking has been performed for the wild type target as the receptor both including and excluding waters in the binding pocket.

In this part of the section, the scoring function and some key aspects of the calculation of the binding free energies in the present work will be discussed. This discussion would be important in the understanding the binding affinities calculated and the mode of these calculations for this pair of macromolecule and ligand i.e. cytochrome P-450 and ionic liquid cation (1-alkyl-3-methyl imidazolium).

The docking score can be represented using a simple expression consisting of the linear combination of all the possible physico-chemical effects during binding. The following equation represents such an empirical approach.

$$\Delta G = \Delta G_{vdw} + \Delta G_{hbond} + \Delta G_{elec} + \Delta G_{conform} + \Delta G_{tor} + \Delta G_{sol} \qquad \text{(III..90)}$$

where the separate ΔG terms on the right hand side represent van der waals interaction, hydrogen bondign, electrostatics, conformational, torsional and desolvation, in that oder. Out of these, the torsional terms is dependent on the restriction of rotatable bonds and translation and the desolvation terms stands for the hydrophobic effects that arise during binding. This has been further implemented in Autodock as

$$\Delta G = \Delta G_{vdw} \sum_{i,j} (\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}) + \Delta G_{hbond} \sum_{i,j} E(t)(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}}) + \Delta G_{elec} \sum_{ij} \frac{q_i q_j}{\varepsilon(r_{ij}) r_{ij}} + \Delta G_{conform} + \Delta G_{tor} N_{tor} + \Delta G_{sol} \sum_{i,j} (S_i V_j + S_j V_i) e^{(\frac{r_{ij}^2}{2\sigma^2})} \qquad \text{(III..91)}$$

where the ΔG terms are empirically determine by fitting binding constants collected from a set of known protein-ligand complexes, i and j are the indexes on ligand and protein atoms respectively in addition to the 1-4 interactions for the atoms in the ligand. A Lennard-Jones 12-6 potential is applied for the dispersion term, 12-10 potential for the hydrogen bonding, E(t) is the weight based on the an-

gle , t, between the probe atom and the target atom. The torsional term involved in the equation due to the restriction of the ligand degrees of freedom is dependent on the number of rotatable bonds in the ligand. The desolvation energy term is dependent on the solvent accessible surface areas, S , and the fragment volumes exposed to the protein atoms, V, weighted by a Gaussian distribution function. The final form of the model that was utilized in the Autodock 3.0 program also included an E_{hbond} term in the hydrogen bonding term which was defined as the estimated energy of hydrogen bonding of waters with a given polar atom. It is also worth noting that in this original model form, the intramolecular or internal energy of the ligand is not included in the binding score calculation. Other assumptions include a constant protonation state and charge distribution in the presence and absence of the ligand within the receptor. The receptor is considered to be rigid with the covalent angles and lengths constant while a set of covalent bonds in the ligand is considered free to move.

The conformational part of the scoring function i.e. $\Delta G_{conform}$ is represented in Autodock Vina as

$$c = \sum_{i<j} f_{t_i t_j}(r_{ij}) \tag{III..92}$$

$$c = c_{inter} + c_{intra} \tag{III..93}$$

where the function sums over all pairs of movable atoms but excluding the 1-4 interactions. Each of the atoms involved is assigned a type t_i and a set of suitable interaction functions $f_{t_i t_j}$ having the interatomic distance r_{ij}. The intermolecular and intramolecular contributions are added to give the final energy c. The optimization algorithm tries to find the global energy minimum and rank the other low-scoring conformations that follow the most favorable one. The intermolecular part of the most favorable hit is used to calculate the predicted free energy of binding of the lowest scoring mode

$$s_1 = g(c_1 - c_{intra1}) = g(c_{inter1}) \quad \text{(III..94)}$$

$$s_i = g(c_i - c_{intra1}) \quad \text{(III..95)}$$

where the s_1 is the best score, g is the function used to evaluate the score while c_i is the score of a certain mode. The score of other conformations than the lowest scoring one are evaluated by the intramolecular contribution, intra1, of the best conformation.

3.3 QMMM Modeling

Investigations involving bond changes i.e. breaking or forming is a big part part of studying enzymes. Therefore, it becomes important to study the reactive part of the enzyme through high accuracy methods such as QM for accurate prediction of structure and energies [6, 215, 216]. Molecular mechanics alone is not sufficient for such investigations as they do not describe electron density. Although, *ab initio* methods or DFT is capable of providing insight into reactive processes but they suffer from size limitations. Employment of pure QM methods to system sizes comparable to enzymes or proteins is difficult and intractable in most cases. Oversimplification of the model region would, on the other hand, lead to serious inaccuracies in prediction of molecular interactions. In order to study large biomolecules, a hybrid modeling approach must be employed. This approach in theory, must contain a combination of the highly accurate QM method and a relatively inaccurate classical part. To render such a problem to be computationally realizable, the first idea was developed Warshel and Levitt [217, 218, 219, 220] and improved upon by considering a large variety of systems. The rationale behind the original method was that in case of condensed phases, most chemical reactions occur locally within a region of the biomolecule known as the active site. In most enzymes including P-450 BM3, the catalytic process is confined to this specific region and the surrounding residues provide an electrostatic environment that influences the reaction through long-range interactions. The potential for this

QMMM energy consists of three separate classes of interactions; between atoms in the region described by QM, between atoms described by MM and interactions between atoms in both layers. The interactions between the two layers can be described by two separate coupling approaches that are namely, subtractive and additive schemes.

Subtractive Coupling In this scheme, the QMMM energy is obtained in a sequential manner in which the energy of the whole system (QM+MM) is computed at the MM level [221, 222, 223]. The energy of the isolated QM region is added to the MM energy and finally, the MM energy of the QM region is subtracted from the added energy.

$$V_{QMMM} = V_{MM}(MM + QM) + V_{QM}(QM) - V_{MM}(QM) \tag{III..96}$$

where QM and MM denote the atoms present in the respective regions and the subscripts represent the level at which the energy is calculated. Gaussian employs the ONIOM method developed by Morokuma and co-workers which will be discussed in detail in the subsequent chapter. In this scheme, it must be noted that that the two layers do not communicate with each other and a force field is required for the QM region. Also, the QM density is not polarized by the MM environment in this scheme and as such, charge transfer process cannot be described using subtractive coupling.

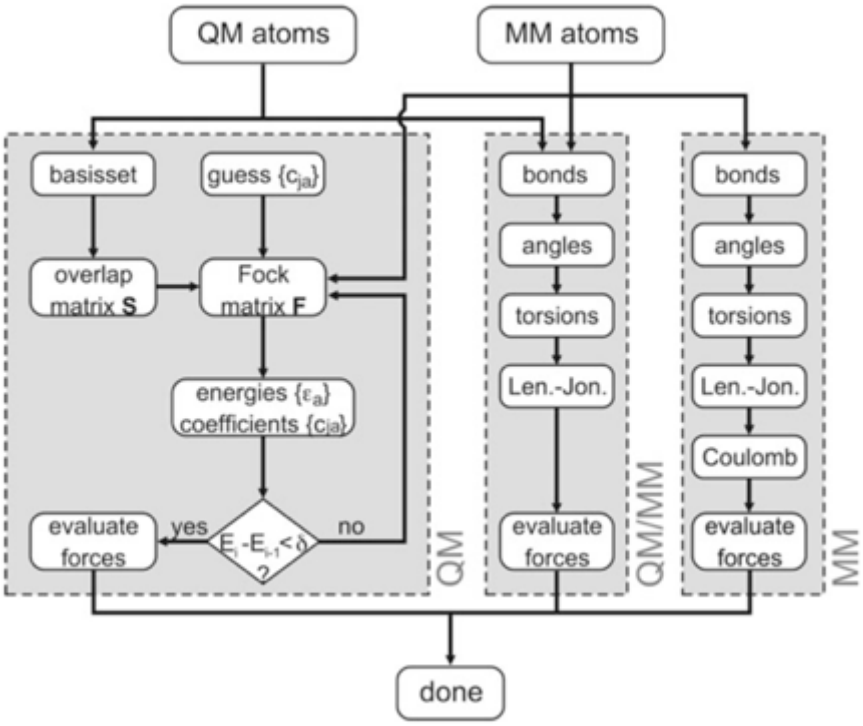


Figure 3.8: Structure parameters computed for different cations to quantify effect of optimization adapted from [6]

Additive Coupling Additive coupling involves embedding the QM region within the rest of the MM sub-system and the potential for the entire system is given by adding energies from MM, QM and coupling terms. Here, only the atoms within the MM region require a force field treatment and the interactions between the QM and MM layers need to be expressed explicitly.

$$V_{QMMM} = V_{QM}(QM) + V_{MM}(MM) + V_{QM-MM}(QM - MM) \qquad \text{(III..97)}$$

Ionic liquids have been projected to be green chemicals due to their low vapor pressure and several other environmentally benign characteristics. Although, many ionic liquids are water soluble and experimental advances have shown that they have significant toxicity and ability to bioaccumulate. Thus, their design needs to incorporate rational methods in order to successfully propagate their large scale industrial usage. Former experimental works have shown that introduction of suitable structural moiety into their cationic part can make the overall ionic liquid to be biodegradable. The mechanism put forth by pioneers in the field has been related to site-specific hydroxylation. Presently, there is a dearth in molecular level details for this mechanism and the present research works is an attempt to uncover some of those details. The development of a computational framework for understanding the biodegradability of a novel solvent forms the innovation behind this work. Such a framework has not been reported in literature related to ionic liquids and this work aims to bridge the molecular principles into macromolecular design level.

CHAPTER IV

METHODOLOGY

4.1 QM System Setup

4.1.1 Conformational Dependent Geometry Optimization

For geometry optimization of cations in the presence of FBP and FeP molecules, a detailed exploration consisting of three different conformations identified by the initial placement of the imidazolium cations with respect to the porphryins was undertaken by adopting the approach described below:

For the first set of conformations, the geometrical orientation of the imidazolium cation was considered such that the alkyl chain pointed away from the plane of the porphyrin ring. One of the hydrogen atoms connected to the 3'-methyl moiety of the cation was placed at a distance of 3 Å from the geometrical center of the four planar N-atoms of the FBP molecule. These conformations will be identified with the tail up conformations for the rest of the article. In the case of FeP complex, the distance of the hydrogen atom was measured from Fe. The second set of confomratios consisted of the alkyl chain of the cations pointing towards the porphyrin molecule. In these configurations, referred to as the tail down configurations, one of the hydrogen atoms attached to the terminal carbon was placed 3 Å from the porphyrin center. The third set was composed of the conformations in which the planes of the imidazolium ring in the cation and the porphyrin molecule were aligned in a parallel fashion such that the initial co-facial distance was 3 Å. These conformations will be denoted as interplanar conformations henceforth.

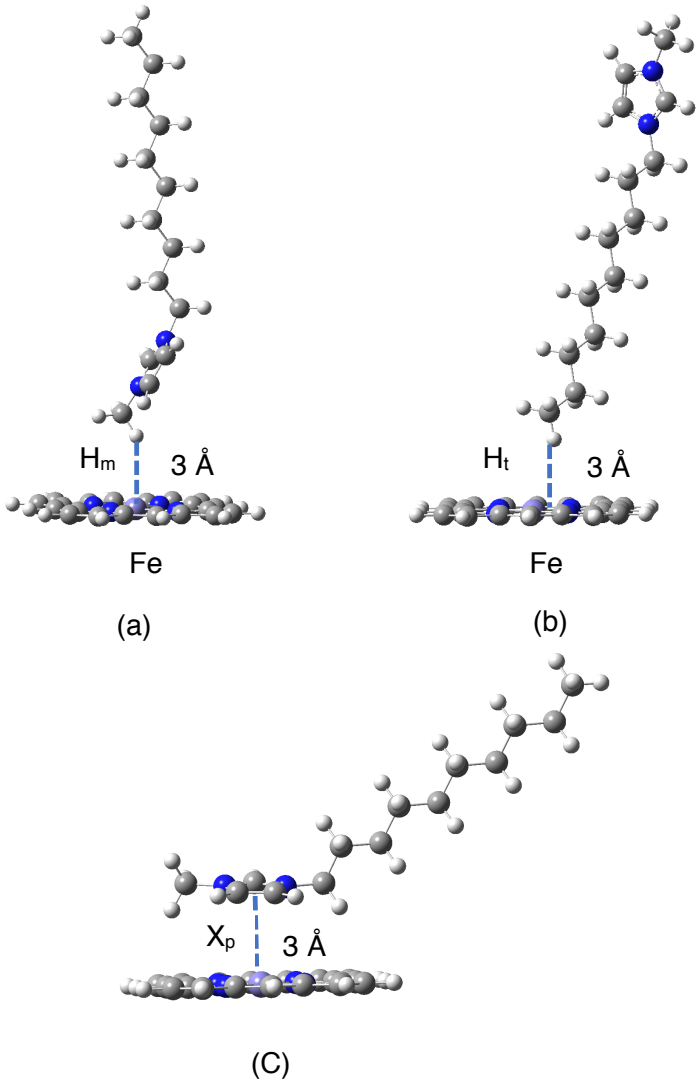


Figure 4.1: Initial structures considered for (a) Tail Up (TU) (b) Tail Down (TD) and (c) Interplanar (IP) conformations, H_m : hydrogen connected to the 3'-methyl group ; H_t: hydrogen connected to the terminal carbon of the alkyl chain ; X_p : distance between the center of the planes of porphyrin and imidazolium ring

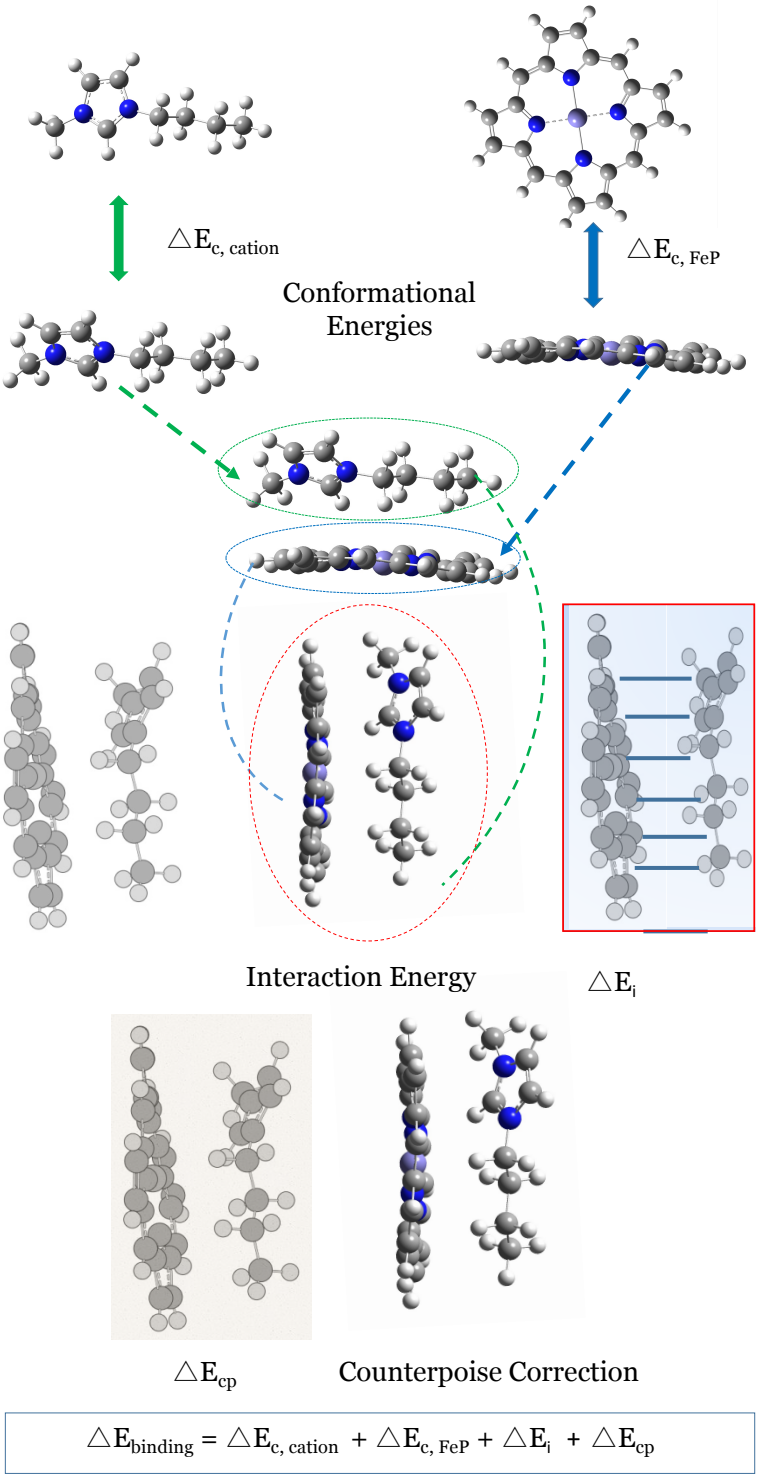


Figure 4.2: Schematic showing the difference steps in the binding process

ΔE_c is the energy change associated with the change in the conformation energies of the two chemical entities participating in the binding process; E_i is the

interaction energy between the ionic liquid and FeP or FBP, and ΔE_{cp} refers to the counterpoise correction.

4.1.2 Population Analysis

Fukui functions are significant in the understanding of local reactivity of a given atom in the scheme of the overall transfer of charge and electronic density in a chemical process. It presents a method from conceptual DFT to gauge the effect of an electron addition or removal process on a certain atom within the whole molecular framework. It is defined as the difference in the associated electronic density for the considered atom which can also be expressed locally as the difference in the charges to condense the effects to a certain point. For an atom specific treatment of the reactive attack, the condensed dual Fukui descriptor [224, 225] was evaluated for the metal at the center of the active site molecule. The function involving the addition of an electron is stated as the difference in the charges of the atom evaluated on the neutral geometry and adjusting the electron density from the charged and neutral species. On the other hand, the removal of an electron is evaluated by imposing a positive charge on the molecule and subtracting from the neutral species. The relationship for the nucleophilic and electrophilic functions are given as :

$$f_M^+ = q_M(N+1) - q_M(N) \tag{IV..1}$$

$$f_M^- = q_M(N) - q_M(N-1) \tag{IV..2}$$

where the terms f_M^+, $q_M(N+1)$, $q_M(N-1)$ and $q_M(N)$ represent the Fukui function for local electrophilicity of metal in complex, charge on metal in the charged state upon losing an electron, the same charge computed after gaining an electron and the charge on metal (M) in the neutral state. Upon combing these two descriptors, the condensed Fukui function is expressed as the difference in nucleophilic and electrophilic indices:

$$f_M^{(2)} = f_M^+ - f_M^- \tag{IV..3}$$

The computed $f_M^{(2)}$ yields a positive value for an electrophilic region and a negative value for a nucleophilic region. All the charges associated with the calculation were evaluated using the Hirshfeld scheme on the geometry of the neutral species of the optimized cation-porphyrin complex to reflect the character of transfer of the electron density.

4.2 Catalytic Cycle Free Energy Calculations

The cytochrome P-450 operates through a well established catalytic cycle mediated by its active site molecule, the heme. In the present work, this heme is modeled by considering an iron porphyrin ligated to a de-protonated cysteinate [226, 227] at the proximal end, which is abbreviated as FePCys. The resting state (S_1) of P-450 involves this FePCys having a water molecule attached at its distal end forming a hexa-coordinate Fe complex. This resting state has shown to exist in a doublet state that requires the inclusion of a suitable substrate that displaces the water molecule and renders the active site molecule penta-coordinated. At this state (S_2), the central metal atom is in a ferric configuration and its spin state transforms to a high sextet. The binding of substrate activates the heme by making it a better electrophile that triggers an electron transfer from the reductase domain of the protein and leads to the formation of a pentacoodinate quintet reduced active site (S_3). This reduced complex then undergoes binding of molecular oxygen (in triplet state) to form an oxy-ferrous compound (S_4) which again acts as an effective electron acceptor. Following the dioxygen binding, a second reduction step takes place which has been reported to be the unique rate determining step in the cycle in literature [228, 229, 230] leading to a reduced dioxygen adduct (S_5). After this second reduction step, a series of protonation steps take place to form a ferric hydroperoxide (S_6) and the ultimate active species (S_7). The formation of S_7 is also associated with the release of a water molecule. This active species plays the role of monooxygenase to the substrate and the resulting product exits the binding

pocket. The cycle is culminated by the re-entry of water molecules to the pocket after release of the oxygenated product to restore the resting state by binding to the heme at the distal site (S_1).

The FePCys complexes were subjected to energy minimization by exploring two different conformations that were described by the initial position of the imidazolium cation with respect to the heme molecule. These were tail up (TU) and tail down (TD) conformations respectively. For the first set of conformations, the geometrical orientation of the imidazolium cation was considered such that the alkyl chain pointed away from the plane of the porphyrin ring, while one of the hydrogen atoms connected to the 3-methyl moiety of the cation was placed at a distance of 3 Å from the Fe attached to the four planar N-atoms of the FePCys molecule. These conformations are identified as tail up (TU) conformations in the subsequent discussion. The second set of conformations consisted of the alkyl chain of the cations pointing towards FeP, such that one of the hydrogen atoms attached to the terminal carbon was placed 3 Å from the porphyrin center. These conformations will be labeled as tail down (TD) henceforth.
The active site species undergoes a series of steps involving reduction, oxidation and protonations during the course of the enzyme's action ultimately leading to hydroxylation. Hence, to move from one intermediate step of the cycle to another, a synchronous approach was devised in which the ionic liquid conformations from the previous step were carried over the next stage. i. e., the optimized geometry from the i^{th} step served as the initial state for energy minimization at the $(i+1)^{th}$ step. As substrates to FePCys are usually characterized by multiple spin states as observed by Shaik and co-workers [231, 232], the most favorable state was identified by conducting electronic structure calculations accounting for various multiplicities and comparing the corresponding relative energies. The geometry associated with the most favorable state was propagated to the next step.

The active site species undergoes a series of steps involving reduction, oxidation and protonations during the course of the enzyme's action ultimately leading to hydroxylation. Hence, to move from one intermediate step of the cycle to another,

a synchronous approach was devised in which the ionic liquid conformations from the previous step were carried over the next stage, i.e., the optimized geometry from the i^{th} step served as the initial state for energy minimization at the $(i+1)^{th}$ step. As substrates to FePCys are usually characterized by multiple spin states as observed by Shaik and co-workers [231], the most favorable state was identified by conducting electronic structure calculations accounting for various multiplicities and comparing the corresponding relative energies. The geometry associated with the most favorable state was propagated to the next step.

4.2.1 Free Energies

Gibbs free energies were evaluated for each step of the catalytic cycle in the gas phase and solvated environments(ε = 1). In the case of the reduction step, both substrate free and substrate-bound energy differences were considered in order to observe the effect of the binding of the $[C_nmim]^+$ cations on the heme molecule. It should be noted that the free energies reported in the work are solution phase free energy changes rather than solvation counterparts. In case of the solvation free energies, an appropriate reference needs to be accounted for in the energy calculations. For evaluating the free energy differences, a direct approach was employed in both the gas phase and dielectric solvent similar to one explained Ho and co-workers [233, 234, 235]. This approach is based on the direct calculation of the properties at both reactant and product side for the reaction within the solvated environment without introducing any external references. Also, all properties in case of embedding are evaluated for solvated geometries optimized by defining an environment having ε = 5.7. This method using a continuum solvation model has been shown to be effective in evaluating free energies in solution and has yielded thermodynamically relevant quantities for biological molecules in experimental agreement [236]. In fact, the authors mention that in the situations where solvation induced geometrical changes to the solute are significant, the direct approach is a better alternative to a thermodynamic cycle approach. All free energies reported in the work have been evaluated at 298.15 K and zero-point corrected without any scaling of the harmonic frequencies. The thermal corrections

to the free energies in solvated environment were computed using the partition functions derived from ideal gas rigid-rotor harmonic oscillator approach [237]. For the discussion related to each step, the section is divided into two subsections; Gas Phase (to discuss the outcomes from the optimization carried out in the gas phase) and Embedded (optimization carried out in the electrostatic embedding having ε = 5.7.

$$\Delta G_{1,x} = G_{S_2,x} + G_{H_2O} - G_{S_1} - G_{[C_nmim]^+} \tag{IV..4}$$

$$\Delta G_{2,x} = G_{S_3,x} - G_{S_2,x} \tag{IV..5}$$

$$\Delta G_{3,x} = G_{S_4,x} - G_{S_3,x} - G_{O_2} \tag{IV..6}$$

$$\Delta G_{4,x} = G_{S_5,x} - G_{S_4,x} \tag{IV..7}$$

$$\Delta G_{5,x} = G_{S_6,x} - G_{S_5,x} - G_{H^+} \tag{IV..8}$$

$$\Delta G_{6,x} = G_{S_7,x} - G_{S_6,x} - G_{H^+} \tag{IV..9}$$

$$\Delta G_{7,x} = G_{S_1} + G_{[C_nmim]OH} - G_{S_7,x} - G_{H_2O} \tag{IV..10}$$

where G_{S_n} are the respective free energies of the complex at the n^{th} intermediate state; the subscript 'x' denotes either substrate bound (sb) or substrate free (sf) state of the species in complex; G_{O_2}, G_{H_2O}, and G_{H^+} are the free energies of oxygen, water and a proton in the gas phase, respectively, and $G_{[C_nmim]OH}$ for the hydroxylated $[\text{C}_\text{n}\text{mim}]^+$ cation.

4.2.2 Reduction Potentials

Reduction potentials serve as an indicator of the affinity of biomolecules to reduce themselves via an electron uptake in the course of their respective catalytic cycles. In the case of heme, this uptake takes place as the cycle moves from S_2 - S_3 in the presence of substrate [238, 239]. It is well known that substrate binding causes the Fe on heme to convert from a low-spin state (S=1/2) to a high spin (S=5/2) state. [240] Therefore, the high spin state was considered for evaluating the reduction potential for the substrate free heme. To account for the protein environment, the whole system was immersed in a dielectric of 5.7 [58] that is representative of the probable electronic embedding of P-450 BM3. Using this setup, the tail up (TU) and tail down (TD) conformations were optimized and then calculations for the reduction potentials (E_{mv}) were carried out. For probing the effect of substrate binding, the relative reduction potential [241, 242] was computed with the substrate free state of FePCys as the reference. In the work, reduction potentials were calculated for both the reduction steps involved in the cycle, i.e. S_2 - S_3 and S_4 - S_5 in the presence and absence of the substrate to gauge its effect. The following relationship was used for the above said computation for understanding substrate effects.

$$\Delta E_{red} = -(\Delta G_{2,sb} - \Delta G_{2,sf})/F \tag{IV..11}$$

where ΔE_{red} is the change in the reduction potential upon binding of $[C_nmim]^+$ cation and F is the Faraday constant that is equal to 23.06 $kcalmol^{-1}V^{-1}$. Also, for understanding the effects of solvated environment on the resulting reductive ability of the active site molecule in the presence of the substrate using the following relationship.

$$\Delta E_{red}^{relative} = (\Delta G_{2,sb}^{1} - \Delta G_{2,sb}^{5.7})/F \tag{IV..12}$$

where $\Delta E_{red}^{relative}$ is the reduction potential difference between two different environments both containing the substrate, $\Delta G_{2,sb}^{1}$ and $\Delta G_{2,sb}^{5.7}$ are the free energy

changes for the reduction step arising from gas phase and embedded environments, respectively.

4.3 Molecular Docking

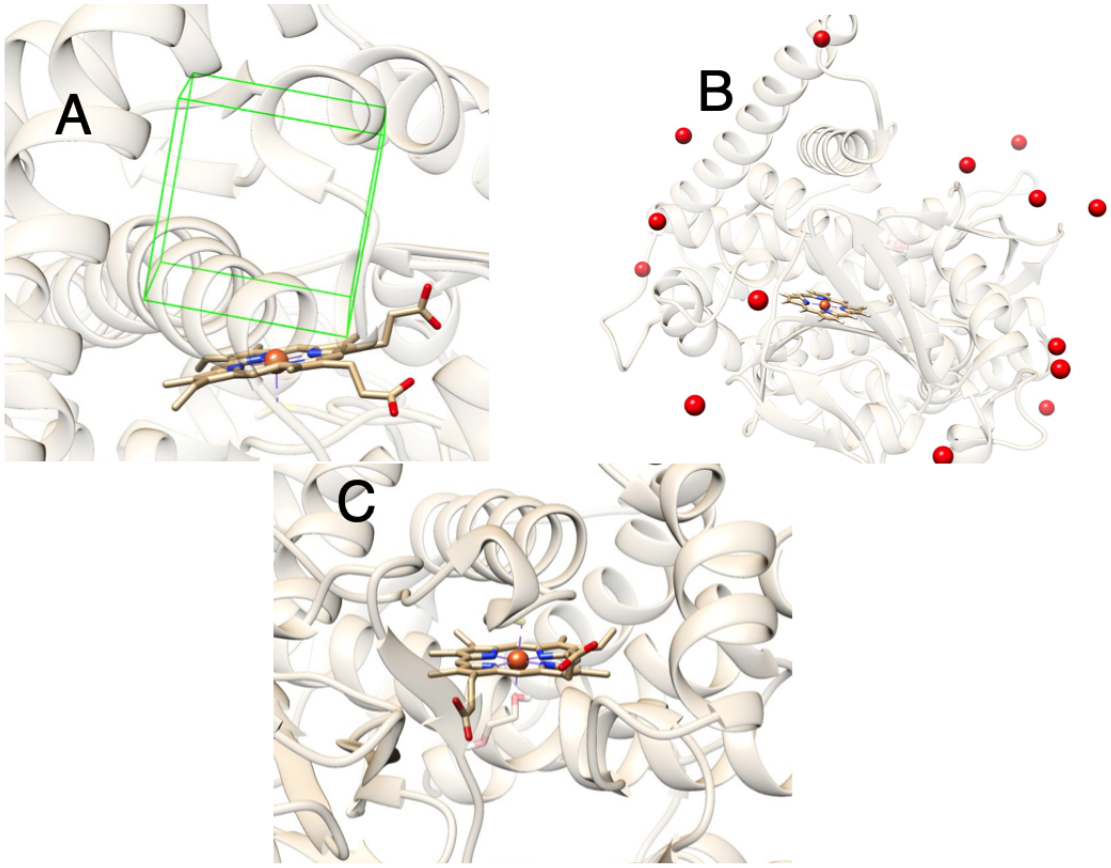


Figure 4.3: Schematic showing different elements of the receptor for docking setup (A) Search volume within binding pocket in green (B) Receptor consisting of the counter ions (Na^+) in red (C) The binding pocket surrounded by the secondary structure elements of P-450

For the docking studies performed in the present work, one or more variants of the cytochrome P450 BM3 were utilized as the macromolecular target. This was done to ensure that substrate affinity is captured for the IL cation for the protein molecule. The binding pocket of the P450 BM3 enzyme is well defined as widely

studied for its electronic and geometrical properties. Docking studies become essential to our work in the absence of any crystal structure available encapsulating the ionic liquid cation as the small molecule embedded within the protein. As mentioned in the introduction, the $[C_n mim]^+$ cation is not a native substrate for the cytochrome P450 protein and hence, docking is performed to understand favorable poses of the cation within the binding pocket and also to serve as effective initial guesses for hybrid QMMM modeling in the subsequent section. All docking calculations were performed using the Autodock Vina plugin [243] available within Chimera. Some of the reasons for choosing this software among all the others available are its easy implementation, scoring function (which is a combination of both empirical and knowledge-based methods) and its recommendation for cytochrome P-450 containing systems from various sources. For the docking process, the receptor and ligands need to be prepared in a way that is chemically adequate and reflects the electronic properties of these molecules. The receptor preparation begins with loading a crystal structure of the substrate-free form of the enzyme. The substrate-free form was chosen to fully eliminate the effects of reorganization of the binding pocket in response to the substrate. Effects such as desolvation and interaction energies depend on the lock and key mechanism that is central to the prediction of the substrate binding affinity. These effects are sensitive to the shape of the binding pocket as well as the potential interaction mechanisms of the binding pocket with the ligand. Hydrogens are then added to the system to reflect the protonation states of the individual residues in the crystal structure. In the present work, the hydrogens were added by using the web server H++ [244] by employing a neutral pH for all the protonations involved in the special residues like histidine, proline and arginine. One of the important points to be noted for this step is that the heme is attached to a de-protonated cysteine group attached to the heme molecule. In literature and theory, a de-protonated cysteine serves as the thiolate donor to the heme in P450. Due to this reason, the proton added to cysteine previously by H++ was removed before docking to make sure that the residue attached to the HEM is a CYM which stands for a deprotonated cysteine. Following the hydrogen addition, equivalent charges were added to the crystal

structure that were modeled using the Gasteiger scheme available within Chimera. As heme is a non-standard residue, total charges, protonation states and chemical state need to be specified. For the same, the overall charge of -2 reflecting the two propionate chains attached to porphyrin macromolecule was stated to the model. Also, the net charge of the whole molecule without the Fe (non-Fe) was kept at -4. The protonation states of the residues were kept unchanged as they had already been subjected to a protonation at the neutral pH. Upon completion of these steps, the structure recovered from the docking setup was used as a template for the calculations. Another key point to be noted here is in terms of the oxidation state of Fe. The binding of substrate changes the spin state of the underlying heme from doublet to sextet. To capture this, the spin of the Fe was kept at doublet and the inclusion of substrate in the hybrid modeling step would reflect the change in the electronic potential of the central metal atom. Also, counter ions were used to neutralize the standard part of the protein, i.e. without HEM. For the neutralization, simple Na^+ ions were added (Figure 4.3 to the crystal structure and it was made sure that all of these ions added to the surface so that they did not interfere with the docking of the ligand.

Upon completion of the dock setup for the receptor, the ligand was prepared for docking. For the docking process, 1-n-alkyl-3-methyl imidazolium cations of chain lengths 2,4,6,8 and 10 were used as the template. For accuracy of the docking process and proper description of the ligand geometry, these cations were initially subjected to pure QM calculations. These calculations were performed at the M06 level of theory with a basis set of 6-31g(d,p) that is an adequate configuration for addressing the electronic structure of $[C_n mim]^+$ cations. The optimized geometries of the ligands were used as the template structures for the docking. For the electrostatics, the charges of the ionic liquid cation at the QM level were used as the template for performing the docking calculations. Another important point to be noted here was that each of the ionic liquid cation structures in gas phase

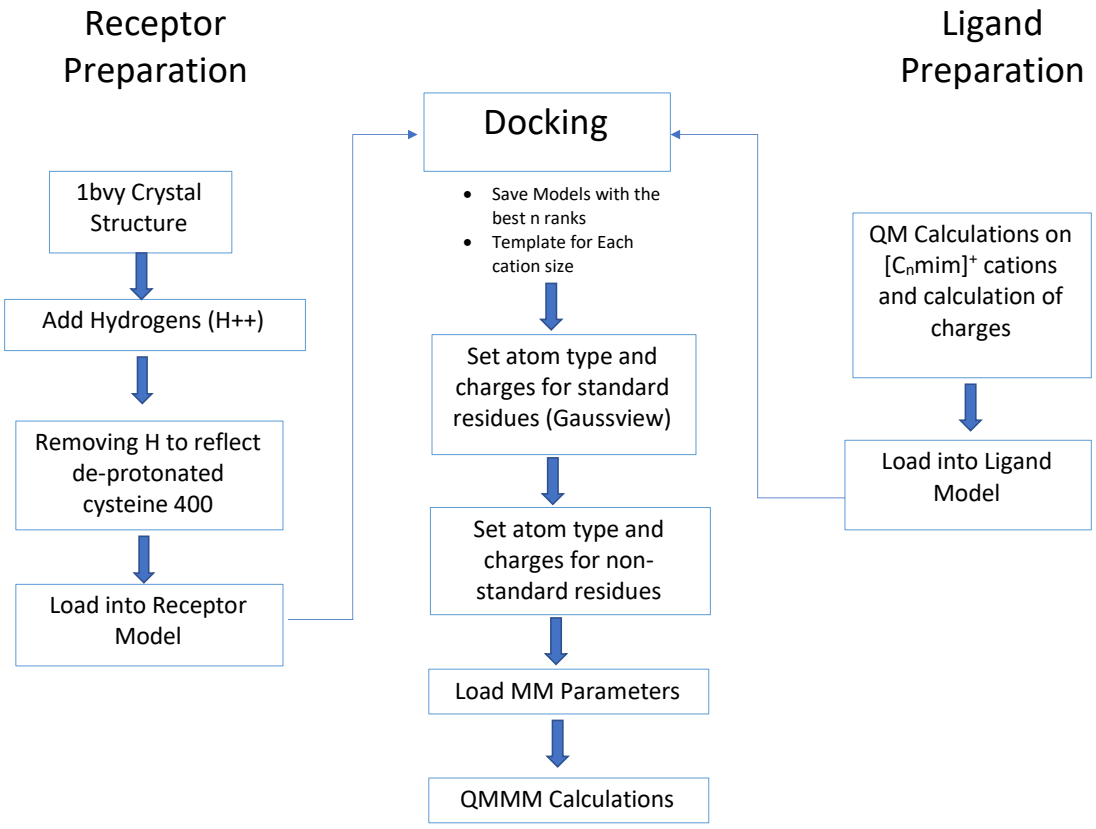


Figure 4.4: Schematic showing flowchart of the docking process and QMMM setup

were evaluated at a very high level of DFT accuracy and visually inspected for extraneous minima or chain curling effects.

Upon the completion of the ligand and receptor preparation, the docking was performed by using Autodock Vina. The three necessary elements for any docking calculation are the search volume, energy penalty and the number of modes. In the current work, for finding the search volume, a visual inspection of the binding pocket was carried out. Upon locating the center of the binding pocket, a cubic space of 10 Å was carved out to fit the ligand conformation. The dimensions were decided based on similar efforts in literature [238, 245]. To test the effect of the search space on the binding score produced from the docking calculations, the search volume was altered to 12 Å and the results would be discussed in the

section dedicated to docking within Results and Discussion section in the dissertation. No energetic penalties or extra smoothing functions were applied to the docking process in this case and the number of modes recorded were the top five or higher, whenever possible. It is also to be noted that the Autodock Vina scoring function works through a combination of empirical and knowledge based methods. Some of the conformational aspects related to the geometrical fit are dictated by a genetic algorithm that takes into account rotatable bonds within the molecule. For this reason, the number of rotatable bonds , N_{tor} were kept unchanged and no geometrical restrictions were imposed on the ligand molecules.

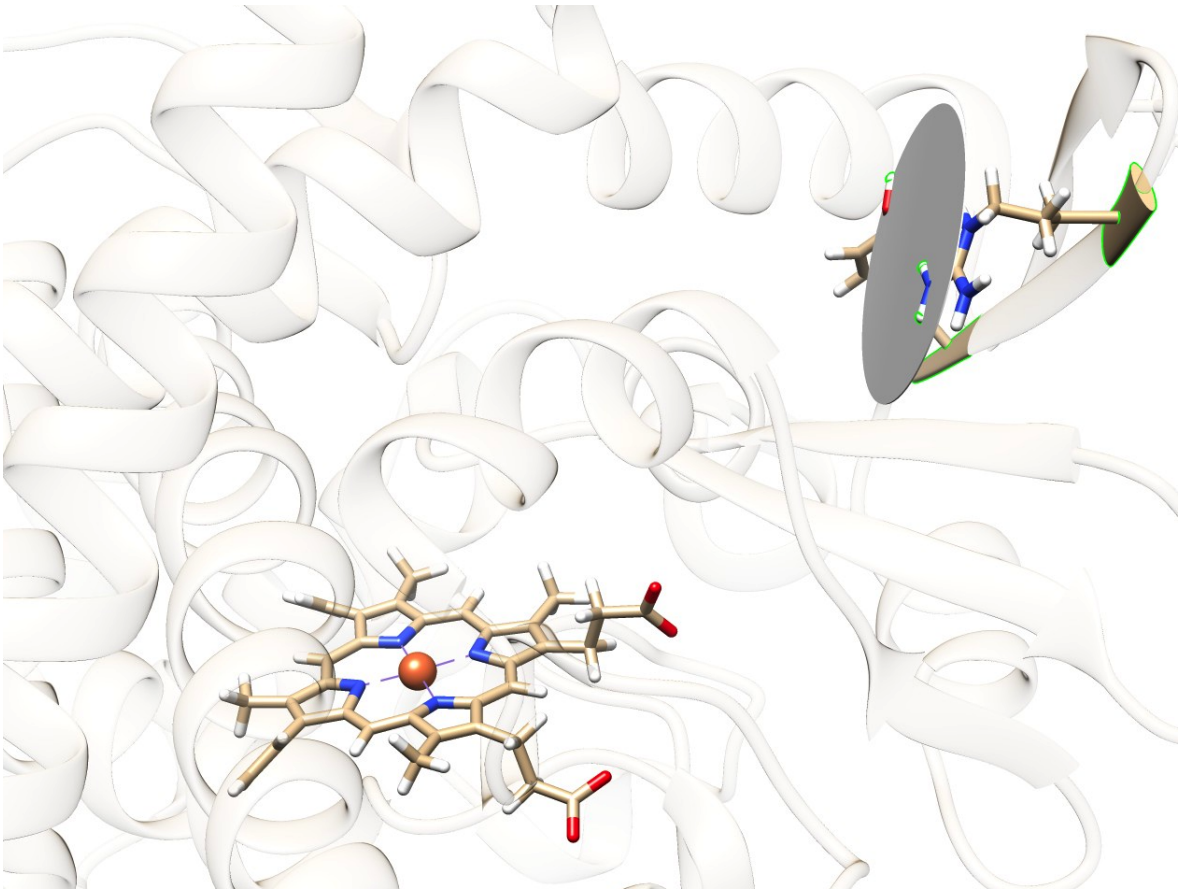


Figure 4.5: Grey disc showing the point of entry of the substrate into the binding pocket of P450 BM3

The termination and the fitness of the generations obtained as a result of the

treatment of the initial geometry through a genetic algorithm converge to different minima. Thus, it is understood that the same conformational state may not be recovered upon re-docking in the presently used scheme. To avoid discrepancies, the docking has been performed multiple times to check for the overall binding score, statistics and conformational samples. The results presented in the work are reflective of the sampling from the docking process. Another key limitation of the docking seems to be the search volume orientation. From Figure 4.5, it is clear that the entry pathway of the substrate leading to the point of association with the receptor is not exactly cubic but rather a tapered cylinder that has been mentioned in literature [246]. In the absence of such a geometrical grid, a cubic grid was assumed to search for the associated geometry and position of the ligand. Although, it has been well established through former docking experiments that a cubic search volume is adequate in providing conformations of natural substrates that are able to validate key quantities such as free energy of hydration (ΔG_{hyd}) and binding rate constants. In the figure 4.5, the grey disc represents the probable point of entry for the substrate leading the hydrophobic pocket consisting of the heme active site. The gate keeper residues, Arg47 and Tyr51 and their role has been briefly discussed in the introduction to cytochrome P-450 (Section 3.1.1.1). The docking calculations also only work well when the binding pocket or site of the receptor is well known. In the case of P-450 BM3, for each of the chains (A and B), the binding site is well defined and validated experimentally through electrostatic maps and NMR studies. Docking does not provide the mode of entry of the ligand that also plays a key role in the regioselectivity of a molecule. In the case of our model, it serves as an ingredient for creating initial structures for QMMM calculations and then calculating properties from geometry optimization at the QMMM level.

4.4 QMMM Calculations

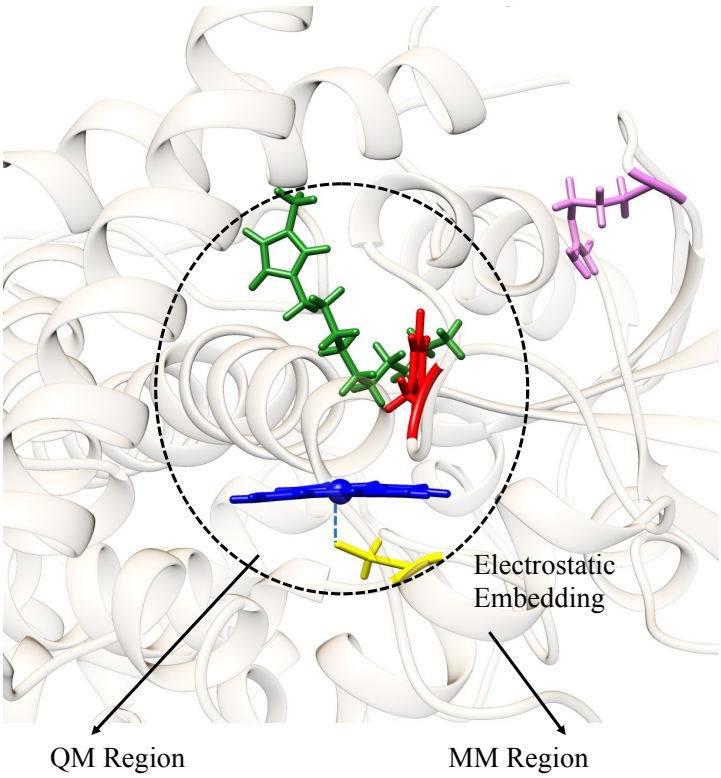


Figure 4.6: Schematic showing a typical electrostatic embedding setup

The successful docking was followed by a hybrid QMMM calculation to predict the molecular and electronic properties. In this section, the common techniques and general procedure for the setup would be discussed. Finally, the methodology for the system considered in the present work will be explained following from the docked structure. The final subsection in the Theory chapter gives an overview of the QMMM technique. Here we discuss some of the specific ideas and methods employed in the present dissertation work. Figure 4.6 shows a schematic of an electrostatically embedded active site region within the surrounding protein region. This embedding is the most common two layer partitioning in the QMMM scheme. The QMMM scheme employed in Gaussian is called ONIOM (our own n-layered integrated molecular orbital and molecular mechanics) that refers to the combination of three separate energetic elements to avoid the usage of specific coupling parameters between the QM and MM region. The most basic relationship for the subtractive coupling is given as

$$E^{ONIOM} = E^{Model,High} + E^{Real,Low} - E^{Model,Low} \qquad \text{(IV..13)}$$

where the superscripts Real and Model refer to the total system and the QM region respectively. The High and Low superscripts refer to the level of theory that the given region is treated with, i.e. quantum mechanical (QM) and molecular mechanical (MM) respectively. The most important part of the protein chemically is the one that is engaged in the formation and destruction of bonds. This bond formation part is normally treated with a higher level of theory at the QM level. The remainder of the system is then treated with molecular mechanics and the partitioning is done keeping that in mind. In the case of cytochrome P450, the active site is well defined and the molecular properties can be obtained from the electronic structure treatment of this part. As a rule, the atoms that do not participate in the chemical reactivity of the system are partitioned into the MM part. Due to the size limitation of electronic QM calculations, there is often a trade off between accuracy and speed. Hence, a decision needs to be made regarding the size of the QM region and the level of theory to be incorporated for the region. The partition of the system needs to be consistent throughout all calculations related to reactions or other processes. The lower level of theory incorporates two different components of energy, the molecular region external to the model region and the region of interaction between the model and MM region. The interaction potential is treated at the lower level of theory. In theory, the ONIOM technique tries to approach to the condition where the whole system would be treated at the model level.

The partitioning of the whole system is done by normally introducing link atoms between the regions. Some of the simple rules for introducing these atoms will be discussed in this paragraph. The number of link atoms to be introduced in the system is equal to the connections of the atoms at the low level with the ones at the model side. None of the trailing connections should have dangling bonds that might cause unnecessary polarization. The link atoms introduced should be

as far from the chemical active region as possible. This prevents them from adding in extraneous electronic effects into the model region. Normally, the error being introduced from the link atom would cancel between the two models depending on the compatibility of the levels of theory involved. The error caused by the introduction of link atoms is reduced when non-polar bonds are incised and hydrogens are added. Proper care needs to be taken when partitioning a system that has atoms with significantly different electronegativities on either side of the QM-MM boundary. Also, cutting through any other bonds than single ones disrupt the electronic structure of the model through improper charge and orbital distribution. Cyclic structures should not be meddled with in case of the partition into the QM and MM regions and should be kept intact. In terms of electronic embedding, the total overall charge of the MM region should remain constant throughout the calculation. For chemical reaction calculations such as transition states, the reaction path becomes the key independent variable for which the formal charge of the total region should remain constant.

The QMMM setup for the systems in the present work will be discussed in this paragraph. The starting geometry for the QMMM calculations was considered to be the ones obtained from docking as discussed in the previous section. Prior to the docking process, the protonation states, neutralization and incorrect links or bonds were cleaned and this reduces the numbers of steps involved in the QMMM setup. The docked geometry or docked model is passed through GaussView to fit the charges for the standard residues involved in the MM region. It is important to note that for the heme fragment and the ionic liquid, it is essential to use charges that are more accurate and obtained from an electronic fitting at the QM level. Thus, for the active site region, a simple SCF calculation is conducted to populate the atomic charges. These charges are used in the input for the overall QMMM setup. In this work, two different sets of models have been considered from the docked ones. The first model consists of just the atoms involved in the iron porphyrin with the connection to Cysteine 400 as the reactive molecule with the IL cation bound to it. The second model consisted of the heme molecule, having the

two propionate chains and the adjoined methyl groups to the central porphine core in addition to the thiolate linkage and the substrate.

The next step involved in the setup for QMMM is the inclusion of the parameters for the non-standard atoms in the MM region. These parameters were obtained from a variety of sources and they have been cited in the present work. A full collection of the parameters may be accessed through the database. The parameters for the ionic liquid cation were obtained from the work of Lopes and coworkers [247] and others [248, 249]. All QMMM calculations and results presented in the work were performed using Gaussian 09 and the associated AMBER [250] database (for standard residues, not including the non standard heme, cysteine or ionic liquid cation) available within the program. Gaussian allows the addition of explicit parameters through different keywords in the MM region. An exhaustive list of the keywords in available in the manual. Some of them have also been discussed in the Calculation Protocols section included in the dissertation. Finally, the spin and multiplicity of the different regions were decided based on the cluster models already developed in the previous section. As the bound substrate takes the whole system to a higher sextet state, the charge and multiplicity of the active site model region was taken to be +1 and 6 for the first model, which will be referred to as M^1_{QMMM} . The larger model containing the propionate chains, M^2_{QMMM}, was considered to have the charge -1 and multiplicity of 6. Both of them include the +1 charge donated by the cation in the model. The protein as a whole was considered neutral. Hence, the real system is considered to have a multiplicity of 1 and the MM calculations were performed on the model system with the same charge and multiplicity as the model. Link atoms were introduced in between the boundary of the model region and the MM region. In case of both the models, the cysteine is linked to ILE401 and ALA399 on either side. Hence, two separate link atoms are introduced at these positions that separates the deprotonated cysteine from the MM region.

CHAPTER V

RESULTS AND DISCUSSION

The chapter is divided into substituent sections dedicated to the different phases of modeling in the work. The first part of the section gives the results and discussions from the cluster modeling on Free base porphyrin (FBP) and Iron Porphyrin (FeP) containing systems. This is followed by the treatment of the free energy of the different steps encountered in the catalytic cycle of P450. Upon completion of the various cluster models, the next sections would be dedicated to docking and QMMM hybrid modeling. The final section would be dedicated to the inferences obtained from the substituting the central atom in metal porphyrins with metals other than Fe and understanding the binding mechanism with imidazolium cations.

5.1 Cluster Model1:B3LYP

5.1.1 Spin States

Before our discussion on the results obtained from the geometry optimization and subsequent population analysis, we tested the energy of the spin states of all of the initial geometries after optimization to determine the most stable one. This was done due to the fact that the 3d orbital of Fe consists of five sub-orbital lobes that are very closely related in terms of their energy. Hence, the number of unpaired electrons can vary depending upon the orbital interaction with the cationic molecule in complex. All of the different conformations discussed in the methodology section were subjected to geometry optimization. After completion

of the optimization process, their single point energies were compared to check for the relative energies between different spins for the same conformation. Figure 5.1 shows the single point energy comparisons for all the three conformations (tail up(TU), tail down(TD) and interplanar (IP)). It is clearly evident that for all of the geometries, the singlet state is the most unstable one with relative energy of 30 kcal/mol or more as compared to the triplet state for the tail up and tail down complexes. For the interplanar complexes, they are on an average about 26 kcal/mol above with the only exception being the largest cation in complex. The quintet state is also energetically unstable as compared to the triplet but closely related in case of $[C_8mim]^+$ and $[C_{10}mim]^+$ FeP containing complexes. The key conclusion that would be carried from this step is that the triplet state is the most stable one amongst the three spin states possible. Hence, all further geometry optimization and analysis were carried out with this electronic configuration for the porphyrin containing complexes. This observation from our calculations is also verified by other works of similar nature that deal with the electronic states of metal porphyrins [251, 252, 253, 254]. All population analysis for the systems were also evaluated at the triplet state only with the exception of the Fukui indexes for which the different electronic states have to be estimated at the neutral geometry.

In this first section, we discuss the results obtained from the calculations performed using the B3LYP functional and 6-31g(d,p) basis set. The population analysis will be followed by the conformational bias of IL cations with respect to the underlying porphyrin and their implications on the binding energies of the overall complex. The HOMO and LUMO energies along with the HOMO-LUMO energy gap of the optimized structures of the imidazolium cations as a function of alkyl chain length are provided in Fig. 5.2. It can be observed that both the orbital energies tend to become more positive with respect to the variation in chain length. It can also be inferred that the change in HOMO energies in going from $[C_2mim]^+$ to $[C_{10}mim]^+$ is ~44 kcal/mol while the difference in the corresponding LUMO energies is quite modest ca. 3 kcal/mol. As a consequence of this behavior, the HOMO-LUMO gap significantly reduces as the alkyl chain is progressively extended. [255, 256] Experimentally, it has been observed that the biodegradation

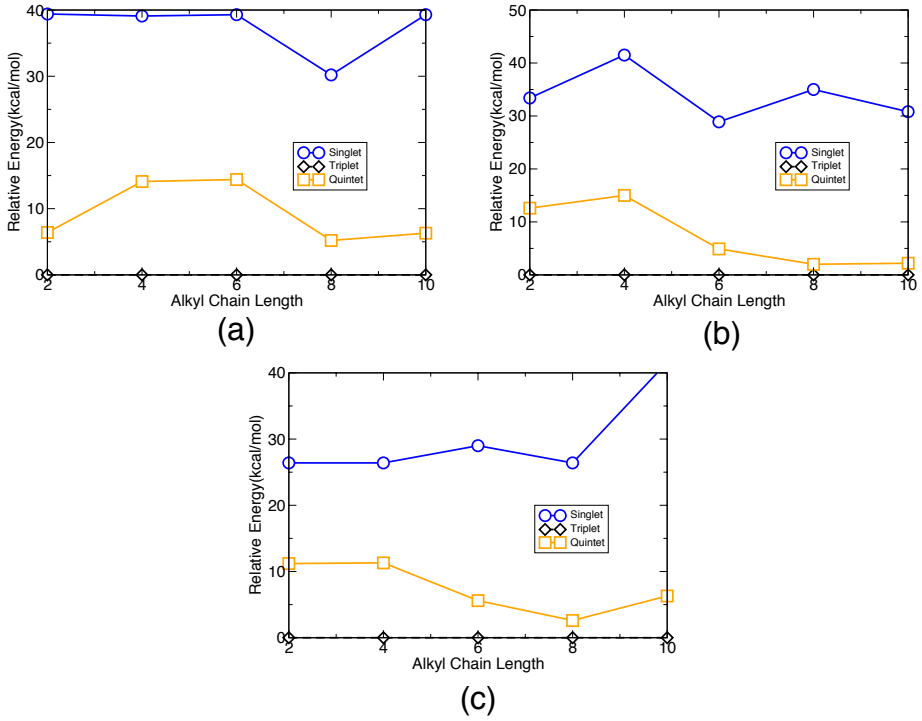


Figure 5.1: Relative Energies of different spin states for $[C_n mim]^+$FeP complexes for (a) Tail Up (b) Tail Down and (c) Interplanar Complexes

occurs to a greater extent for the imidazolium-based ionic liquids bearing longer chains [50]. The HOMO energies of the cation has been shown to bear a linear relationship with the overall biodegradability of the IL. As an increase in HOMO energy is normally associated with a greater tendency of the molecule to react, the gas phase calculations suggest a correlation between the HOMO energies and biodegradation potential of ionic liquids.

In order to provide an insight into the likely sites for reactivity, visualization of the HOMO and LUMO orbital profiles for the cations was carried out and these profiles are presented in Figs. 5.3, 5.4 and 5.5. These figures clearly show that the LUMO remains concentrated on the imidazolium ring for all the cations, which is consistent with the fact that the positive charge of the cation is primarily distributed among the ring atoms. Contrary to the LUMO profiles, the HOMO distributions are cation dependent. For example, the HOMO is located in the ring region in an orthogonal manner to their respective LUMO's for the $[C_2mim]^+$

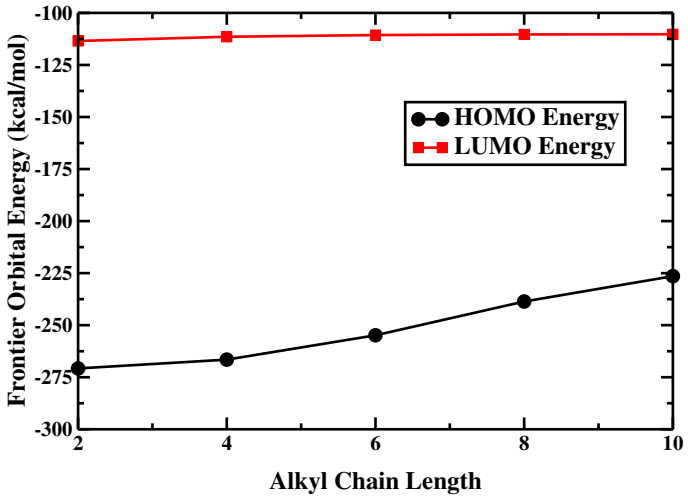
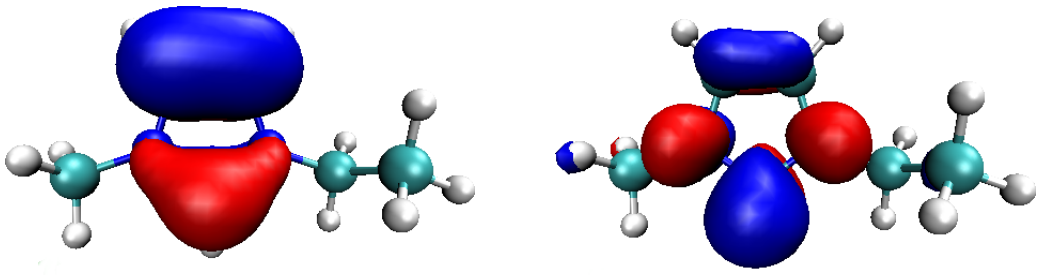


Figure 5.2: Frontier orbital energy values (in kcal/mol) for imidazolium cations in the gas phase

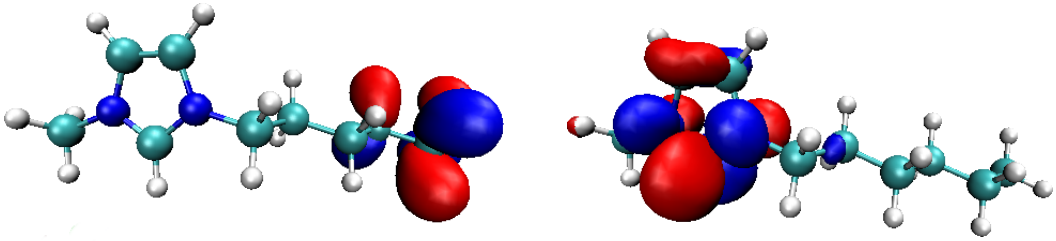
and $[C_4mim]^+$ cations. The HOMO migrates towards the hexyl chain and can be be observed to center on the terminal and sub-terminal carbon atoms for the $[C_6mim]^+$ cation. Further redistribution of the HOMO orbital takes place when the alkyl moiety is extended to $[C_8mim]^+$ and $[C_{10}mim]^+$ such that the occupied frontier orbital is delocalized along the entire alkyl chain. In the context of the biodegradability, the placement of the HOMO indicates that groups along the alkyl chain are likely candidates to participate in reactions involving ionic liquids which is in line with oxidation of the terminal or subterminal groups observed when longer alkyl chains (n $\geq$ 6) are biotransformed. The placement of these frontier orbitals on the molecule do not change even in the presence of the porphyrins however their energy levels are shifted. These effects are discussed further in the analysis included in this section.

The frontier orbitals of the porphyrin active site molecules were also evaluated using the same level of theory and basis set configuration. The key observation is that the HOMO-LUMO gap is about 67 kcal/mol for FBP and 56.5 kcal/mol for FeP. In case of FBP, in the absence of the central metal atom, both the frontier orbitals are populated by the atoms in the pyrrole rings. The hydrogen atoms attached to the pyrrole nitrogens were shown to not participate in the formation of the reactive orbitals. For the FeP, the LUMO orbital is shown to be the d_{z^2} of Fe while the HOMO is equally interspersed on the pyrrole rings. The HOMO, or



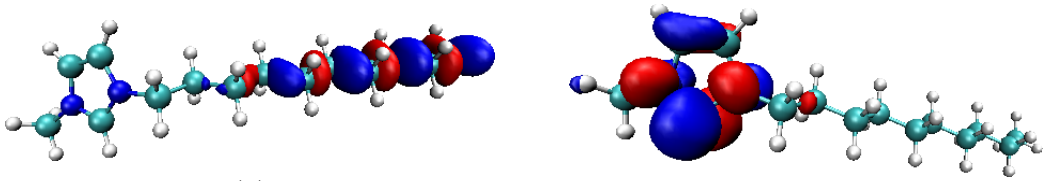
(a) (b)

Figure 5.3: (a) HOMO and (b) LUMO orbital profiles for the $[C_2mim]^+$ cation; the blue and red regions are the positive and negative isosurfaces



(a) (b)

Figure 5.4: (a) HOMO and (b) LUMO orbital profiles for the $[C_6mim]^+$ cation; the blue and red regions are the positive and negative isosurfaces



(a) (b)

Figure 5.5: (a) HOMO and (b) LUMO orbital profiles for the $[C_{10}mim]^+$ cation; the blue and red regions are the positive and negative isosurfaces

the occupied orbital having the highest energy does not show participation from Fe or the pyrrole nitrogens. The donor acceptor interactions that would arise from the complexation of the FeP with ionic liquid cations would be addressed later on the in this chapter. The tentative positioning of the electronic population of the both of the monomers in complex has been given in the adjoining schematic (Figure 5.8). These are relative placements of the individual and collective energy levels that are populated once the complex between the porphyrin and cation is formed. The observations are true for all of the cations considered for our calcu-

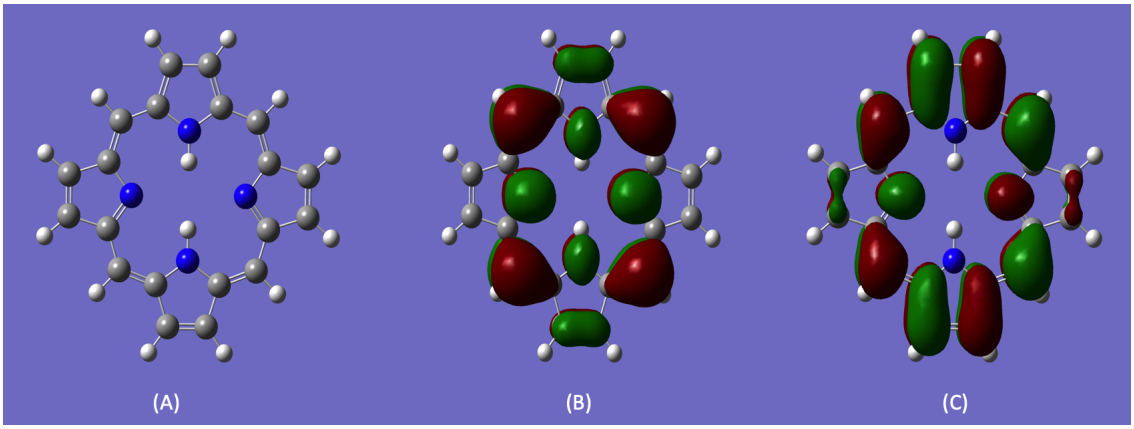


Figure 5.6: Free base porphyrin (A) Optimized Geometry (B) HOMO orbital (C) LUMO orbital

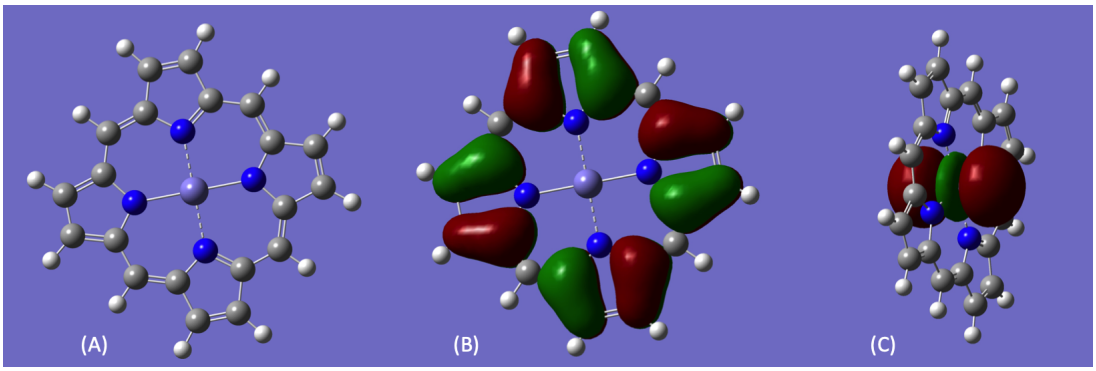


Figure 5.7: Iron porphyrin (A) Optimized Geometry (B) HOMO orbital (C) LUMO orbital

lations.

Before the discussion of the results, it is important to note that the complex HOMO and LUMO are the ones that are reported from the population analysis. They are not to be confused with the porphyrin frontier orbitals. The complex frontier orbitals might have its space occupied solely by porphyrin, cation or might be shared between the two molecules in complex. It is evident from the schematic that the presence of cation shifts the electronic levels of the active site and the cation frontier orbitals populate in between the porphyrin ones. The key conclusion drawn from the figure is that cation LUMO in isolation is more significantly affected than the cation HOMO. At a more specific level, out of the three conformations, the tail down ones reduce the HOMO-LUMO gap of the cation to the

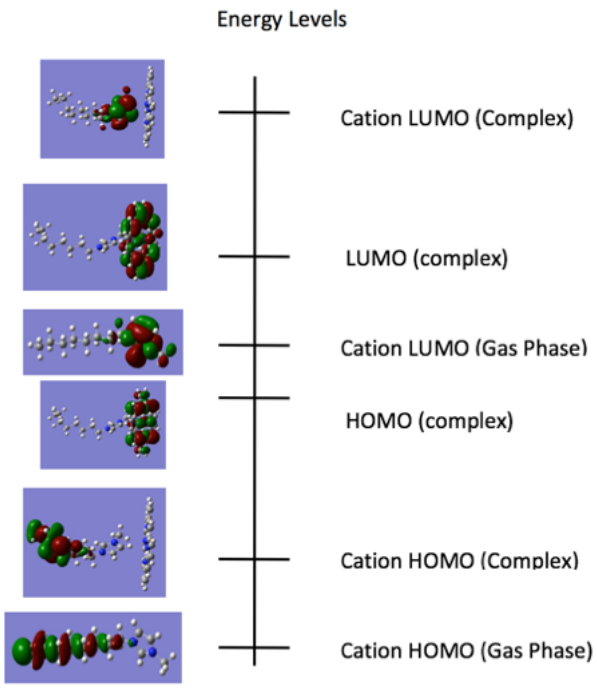


Figure 5.8: Frontier orbital population arising from the complexation of IL cation with porphyrins

greatest extent. This is tied to their structure in which the alkyl chain points towards the underlying porphyrin while in case of the others, the ring is exposed to the same.

5.1.2 Frontier Orbitals in the presence of FBP

In order to account for the change in the HOMO and LUMO energies of the $[C_n mim]^+$ cations in the presence of FBP, the HOMO of cations was identified as the highest energy orbital residing predominantly on the cation from visual inspection. The energy level was then confirmed with the density of states (DOS) plots generated using GaussSum [257]. Similarly, the cation LUMO and the corresponding energy were determined.

The change in the HOMO energies for the cation relative to the respective gas phase energies is depicted in Fig. 5.12 as a function of the alkyl chain length and the orientations. Conformations in which the alkyl chain is oriented towards FBP (tail down) exhibit HOMO energy shifts that are positive for all the cations. The

elevation in the HOMO energies is of the order of 20-30 kcal/mol. A markedly different behavior is displayed by the interplanar and tail up conformations. For these conformations, the HOMO energy upscaling is very similar to that found for the respective tail down conformation for the first two cations in the series. In fact, the HOMO changes are very similar (within ca. 1-4 kcal/mol) for all the orientations investigated for the $[C_2mim]^+$ and $[C_4mim]^+$ cations. However, unlike the tail down conformations, the difference in the HOMO energies for the larger cations with respect to the corresponding gas phase values begin to diminish from the $[C_6mim]^+$ and higher alkyl chain analogues. It is noteworthy that the HOMO energy changes for the tail up and interplanar conformations for the cations up to $[C_8mim]^+$ are almost superimposable. For the $[C_{10}mim]^+$ cation, the departure of the HOMO energy from its gas phase value is negligible for the tail up conformation while the HOMO energy level drops below its gas phase value by ~15 kcal/mol in the interplanar conformation. Upon visualizing the cation HOMO and LUMO orbitals in the complex, it can also be shown that the relative placement of these orbitals are unchanged in the presence of porphyrin. Figures 5.9 - 5.11 show the location of the cation LUMO and HOMO in optimized FBP containing TU complexes for $[C_2mim]^+$, $[C_6mim]^+$ and $[C_{10}mim]^+$ cations respectively. Similar to the location of the HOMO and LUMO orbitals in isolation, the LUMO remains concentrated on the ring while a part of the HOMO migrates away from the ring as the alkyl chain length on the side of the cation is increased from butyl to hexyl and greater.

The perturbation of the cation HOMO energies in the presence of FBP relative to the gas phase HOMO energies can be explained considering the proximity of the HOMO region to FBP. For the $[C_2mim]^+$ and $[C_4mim]^+$ cations, the gas phase HOMO is distributed on the imidazolium ring which is in direct contact with FBP. The closer positioning of the imidazolium ring in the tail up and interplanar conformations results into the HOMO energy shifts being higher than that for the tail down conformation in $[C_4mim]^+$. The presence of FBP influences the tail down conformations for all the cations in the series as the HOMO, which is located on the alkyl chain is exposed to FBP in these conformations. On the

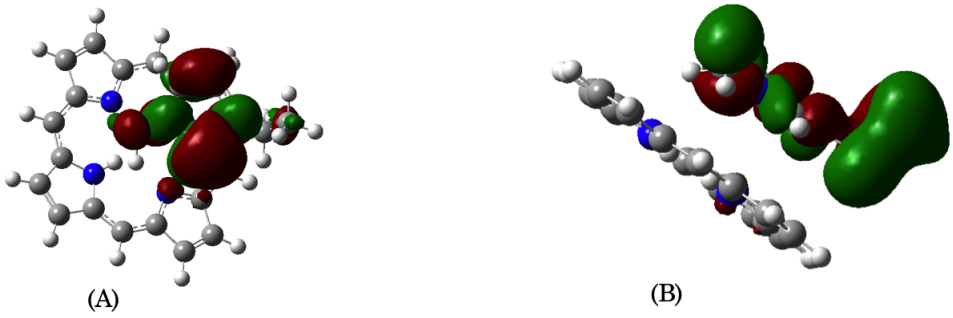


Figure 5.9: Cation (A) LUMO and (B) HOMO orbitals in optimized $[C_2mim]^+$FBP TU complex

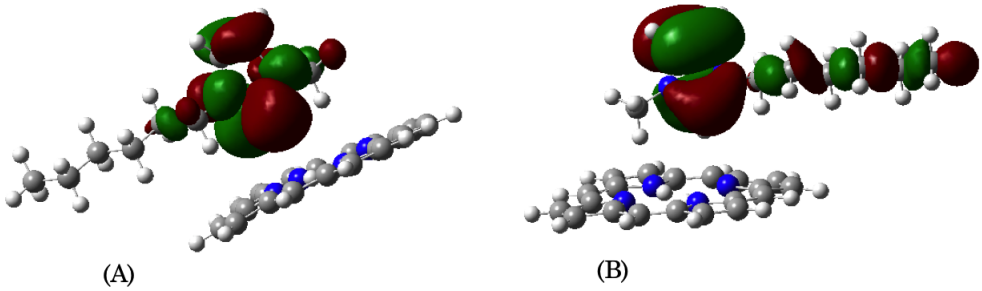


Figure 5.10: Cation (A) LUMO and (B) HOMO orbitals in optimized $[C_6mim]^+$FBP TU complex

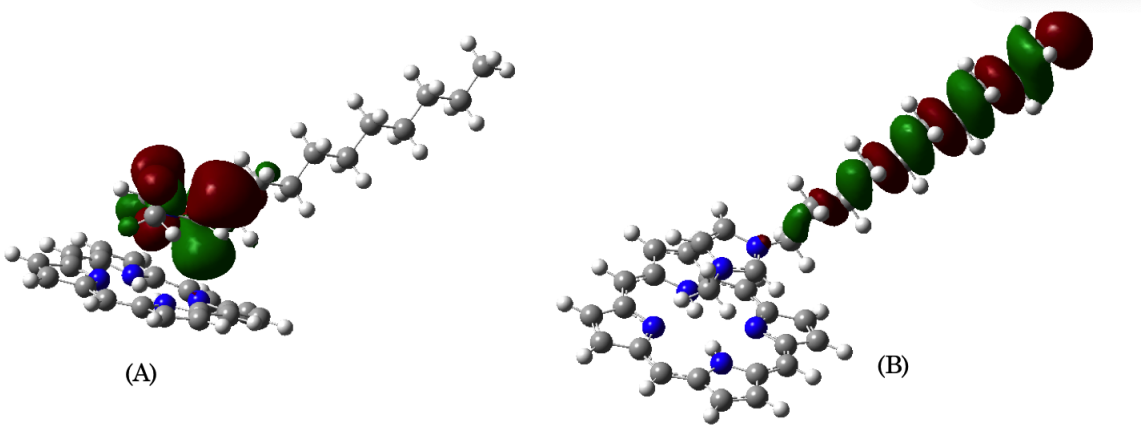
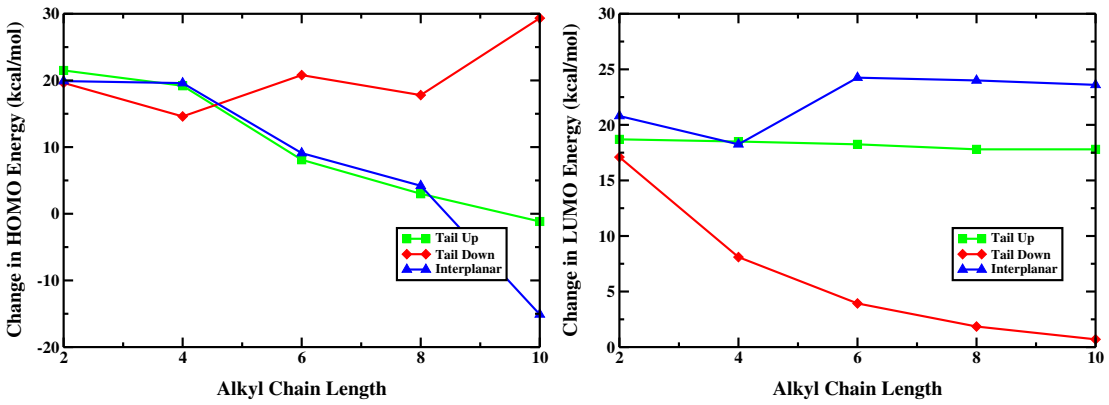


Figure 5.11: Cation (A) LUMO and (B) HOMO orbitals in optimized $[C_{10}mim]^+$FBP TU complex

contrary, a continuous movement of the HOMO location away from FBP, in the tail up conformations, is a reason that the HOMO energy shifts relative to the



(a) (b)

Figure 5.12: Comparison of changes in frontier orbital energies in the presence of free base porphyrin (a) HOMO Energies (b) LUMO Energies

gas phase values begin to drop beyond $[C_4mim]^+$. In the case of the interplanar conformations, there is a shift in the orientation of the ring with respect to FBP for $[C_6mim]^+$ and beyond such that the most acidic carbon in the ring is found to be closer FBP as opposed to C_4 and C_5 for the smaller ones. Moreover, the alkyl chain adopts a parallel orientation with respect to FBP extending beyond FBP.

The LUMO energy change with respect the gas phase datum of the cations in the presence of FBP as a function of the cation orientations are provided in Fig. 5.12b. Similar to the HOMO energies, the LUMO energies of the cations increase for all the conformations. The absolute average deviations (AADs) from the gas phase LUMO energies are 5, 19 and 25 kcal/mol for the tail down, tail up and interplanar, and conformations, respectively. Although the AAD for the tail down conformations is the least, it is worth noting that the elevation in the LUMO energies for $[C_2mim]^+$ in the tail down conformation is as high as 17.5 kcal/mol which is similar to those observed for the tail up and interplanar conformations for the same cation. The rise in the LUMO energy can be rationalized based on the observation that the ring hydrogens are electron deficient and can easily accept electrons from FBP as long as there is an interaction of FBP with these hydrogen atoms. The extent to which the LUMO energies are upscaled is highly dependent on the conformations as expected. The LUMO energy change for the

tail down conformations drop precipitously as a function of the alkyl chain length such that it is almost negligible for $[C_{10}mim]^+$. Such a variation in the LUMO energy change can be accounted based on the fact that the LUMO is positioned on the imidazolium ring in the gas phase. Thus FBP is only effective in perturbing the LUMO energy level for the tail down conformations as long as the imidazolium ring is in its close proximity. Such is the case for the tail up and interplanar conformations in that the imidazolium ring is always situated close to FBP which results in an increase in the LUMO energy change from its gas phase values. As the conformation of the imidazolium ring remains more or less similar for the tail up conformations, the LUMO energy change is predicted to be nearly constant across the homologous series. In the case of interplanar conformations, the LUMO energy changes are similar for $[C_2mim]^+$ and $[C_4mim]^+$ but increase for $[C_6mim]^+$ and are nearly constant for the longer alkyl chain analogues. The behavior of LUMO energy change is attributable to the optimized geometries for $[C_2mim]^+$ and $[C_4mim]^+$ cations, which show that the acidic hydrogen is oriented away from the FBP plane. On the hand, the acidic hydrogen interacts preferentially with FBP for the rest of the homologous series.

5.1.3 Frontier Orbitals in the presence of FeP

Regarding the frontier energy orbitals of FeP, HOMO is equally distributed among the substituent five-membered rings of the molecule, the LUMO site envelops the Fe atom. This may be expected due to the two unfilled 3d states that are closely related to each other on the energy scale.

The variation in the HOMO energies of the cations with respect to their gas phase values in the presence of FeP as a function of the alkyl chain length is presented in Fig. 5.13. The difference in the HOMO energies is observed to depend on both the conformations and identity of the cations. For example, the interplanar conformations yield the least departure from the gas phase HOMO energies the AAD for which is 3.2 kcal/mol. For the tail up and tail down conformations, the AADs are 7.1 and 8.1 kcal/mol, respectively. These changes are modest in comparison to those obtained for FBP *vide supra* underscoring the fact that the

introduction of Fe into the porphyrin molecule reduces its Lewis-base nature. A detailed analysis of Fig. 5.13 shows that the HOMO energy level is consistently higher than that obtained for the corresponding gas phase values for all the cations in the tail down conformations, except for $[C_2mim]^+$. A large jump in the HOMO energy change for $[C_4mim]^+$ in the tail down conformation is probably related to the change in the orientation of the imidazolium ring such that instead of C_5, the most acidic carbon atom C_2 points towards FeP. The continuous drop in the HOMO level shifts for the larger cations can be explained based on the increase in the distance of the terminal carbon atom from Fe.

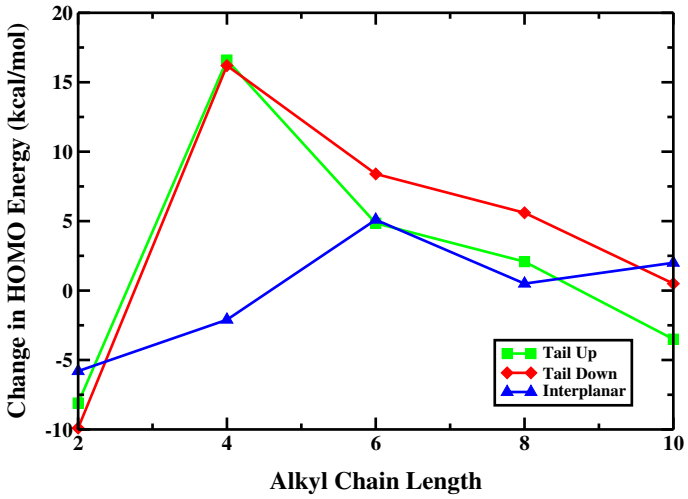


Figure 5.13: Change in the cation HOMO energy values (in kcal/mol) in the presence of FeP

For the tail up conformation, the trend in the HOMO energy difference from the gas-phase values is similar to that observed for the tail down conformations and the HOMO energy shifts are nearly identical for $[C_2mim]^+$ and $[C_4mim]^+$ cations in the tail down and tail up conformations. However, the HOMO energy shifts drop more rapidly for the tail up conformations. As for the tail down conformation, there is a large increase in the HOMO energy level for $[C_4mim]^+$, which can be attributed to the decrease in the distance between C_5 and Fe from 3.8 Å for $[C_2mim]^+$ to 3.67 Å for $[C_4mim]^+$. The increase in the distance between these atoms and the movement of HOMO location away from FeP are responsible for decrease in the HOMO energy changes for $[C_6mim]^+$ and beyond.

The interplanar conformations display maximum deviation of 5 kcal/mol in HOMO

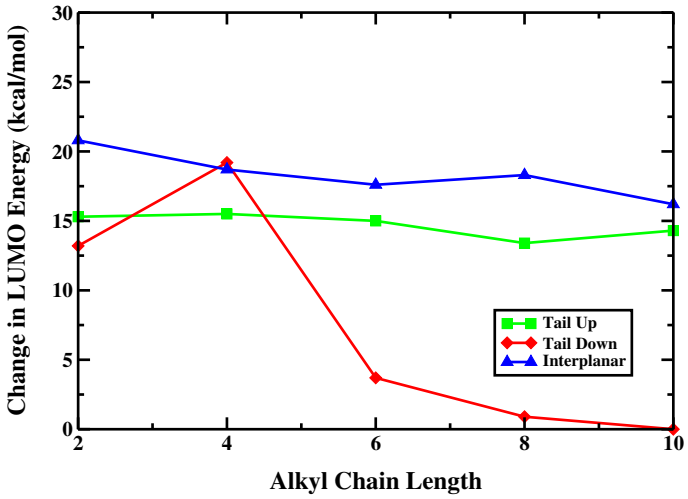


Figure 5.14: Change in the cation LUMO energy values (in kcal/mol) in the presence of FeP

energy level shifts from the corresponding gas phase values. For $[C_2mim]^+$, $[C_4mim]^+$ and $[C_8mim]^+$ the imidazolium ring is almost parallel to FeP with C_4 and C_4 pointing towards FeP, while the ring is positioned perpendicular to FeP for the other two cations with C_2 facing FeP. The preferred alignment of the alkyl chain is parallel to FeP in the complex geometries.

The change in the LUMO energies of the cation in the presence of FeP is illustrated in Fig. 5.14. Unlike the HOMO level shifts, the LUMO energies are consistently upscaled across all the conformations and the entire homologous series. The highest difference in the LUMO energy change is seen for the interplanar conformations for each of the cations. This is followed by the upscaling of the LUMO energies for the tail up conformations. For both the conformations, elevation of the LUMO energies is fairly uniform along the cation series. The most variation in the LUMO energies is obtained for the tail down conformations. The calculations suggest that the LUMO energy change increases from $[C_2mim]^+$ to $[C_4mim]^+$ beyond which there is a sharp decline for $[C_6mim]^+$. In the case of $[C_8mim]^+$ and $[C_{10}mim]^+$, gas phase LUMO energies are recovered indicating that the FeP exerts only minimum influence on the LUMO behavior of these cations. The reason for this behavior is the fact that with the LUMO is located on the imidazolium ring which progressively moves further from the FeP plane thereby reducing any potential electronic coupling with the LUMO energy state. Similarly,

for the tail up and interplanar conformations, the proximity of the imidazolium ring to FeP leads to elevation in the LUMO energies. The difference in the extent to which the LUMO energies are shifted for these two conformations is related to the orientation of the imidazolum ring with respect to FeP. For example, the interplanar conformations are characterized by the imidazolium ring that is parallel to FeP while the tail up conformations result into a perpendicular orientation of the cation ring.

As a conclusion from the above analysis concerning the HOMO and LUMO of the cation in the complex, their geometrical placement is dependent on the identity of the cation as well as the conformation presented to the porphyrin. This holds true for both FBP and FeP containing systems. The influence on the cation LUMO is more pronounced as compared to HOMO irrespective of the conformation. This shows the oxidizing ability of the porphyrin in the complex and also sheds light on the affinity of the cation frontier orbitals with respect to the complex geometry.

5.1.4 Intermediate Energy States

Intermolecular interactions between cations and FBP allow the former to reach energy states that were unattainable in the gas phase. Distinct energy states were observed for nearly all the conformations between frontier orbital levels for cation in gas phase and complex. Upon observing Figure 5.16, it is evident that for $[C_n\text{mim}]^+$FBP complexes (a-b), the number of intermediate states are numerically less than the ones containing iron porphyrin. This can be explained by noting that the central metal ion is capable of coordinating with the cation molecules in a variety of electronic configuration due to the presence of unfilled 3d-orbitals. This observation is true for all the geometries that were calculated as a result of optimization in the presence of both FBP and FeP. No consistent behaviour was observed regarding the placement of the energy levels of these hybridized states. When considering the transition of the frontier orbitals of the cations from gas phase to complex, it was observed that the highest number of hybridized states occured for tail down conformations. This again highlights the fact that interaction between the molecules is dependent on the geometrical orientations consid-

ered. The extra hybrid states observed for the tail down conformations exhibit the fact that chain interactions with the porphyrin plane promote the formation of intermediate states to a greater extent than the other conformations considered. The number of intermediate states for cationic molecules is greater than those of porphyrins. This is also explained in the schematic given in Figure 5.8 where the tentative energy levels after complexation are depicted. Upon visualizing the DOS levels(Figures 5.17 - 5.20), the apparent distribution of the electronic levels for both the monomers becomes clear. For both FBP and FeP containing systems, the separation between the HOMO and LUMO states for cations reduce irrespective of the conformation presented to them. Mathematically, the HOMO-LUMO gap of the cations reduce significantly to the order of 50-80 kcal/mol which is equivalent to about 2-3.5 eV. This change in energy is significant in driving the reduction of the cation by the porphyrin molecule. The frontier orbital energy gap is not dependent on the conformation presented to the underlying porphyrin, which is an indicator of the fact that binding energies obtained are not purely a function of the orbital exchange.

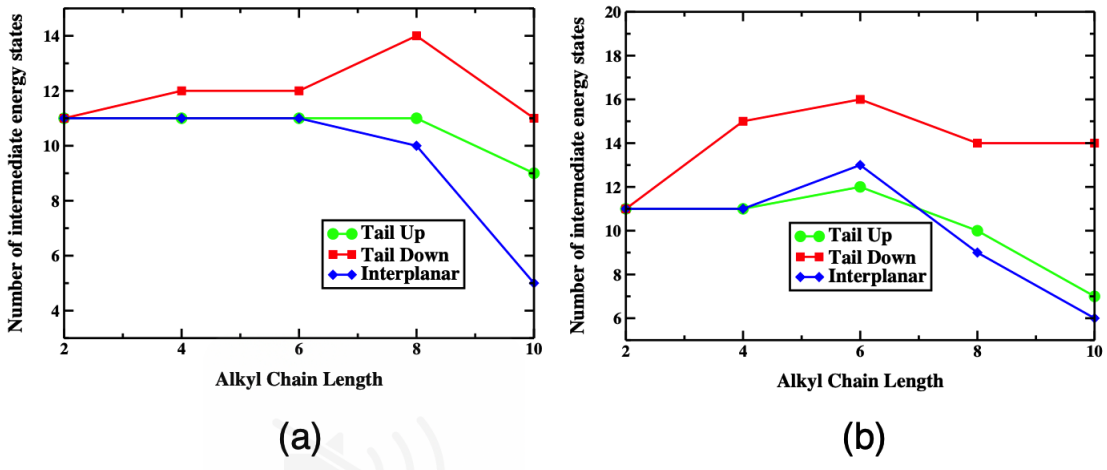


Figure 5.15: Intermediate energy states for (a) FBP (b) $[C_n mim]^+$ cations in $[C_n mim]^+$FBP complexes at B3LYP

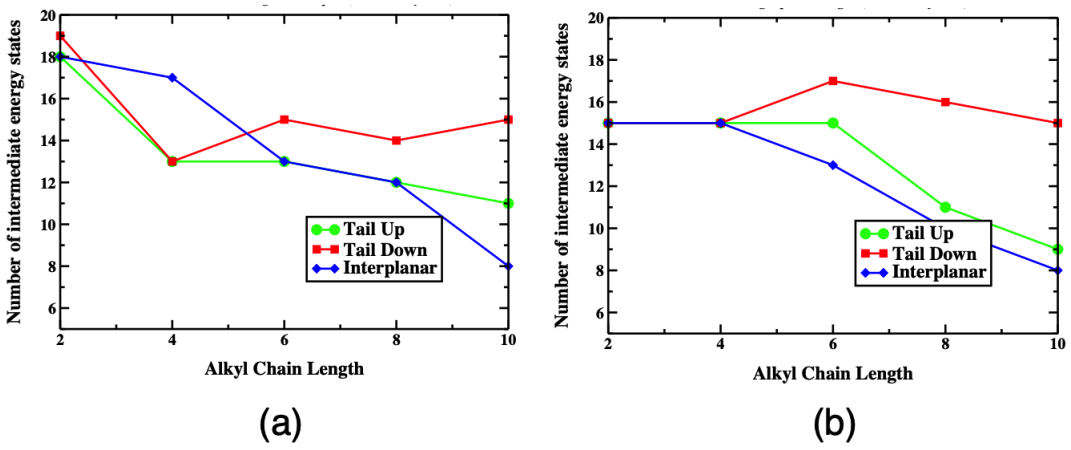


Figure 5.16: Intermediate energy states for (a) FeP (b) $[C_nmim]^+$ cations in $[C_nmim]^+$FeP complexes at B3LYP

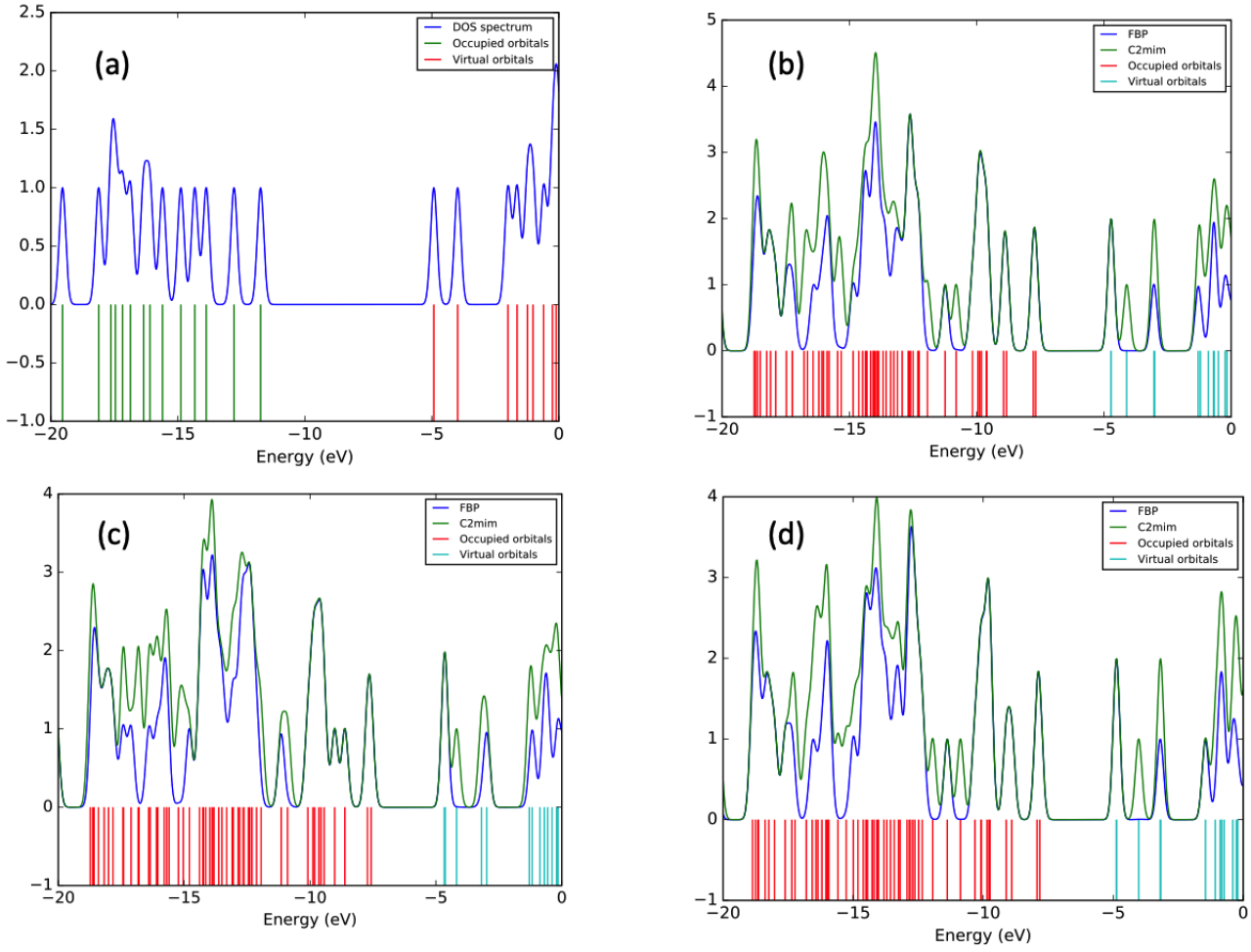


Figure 5.17: Density of States of (a) $[C_2mim]^+$ (isolated) (b) $[C_2mim]^+$FBP tail up (c) $[C_2mim]^+$FBP tail down (d) $[C_2mim]^+$FBP interplanar complexes

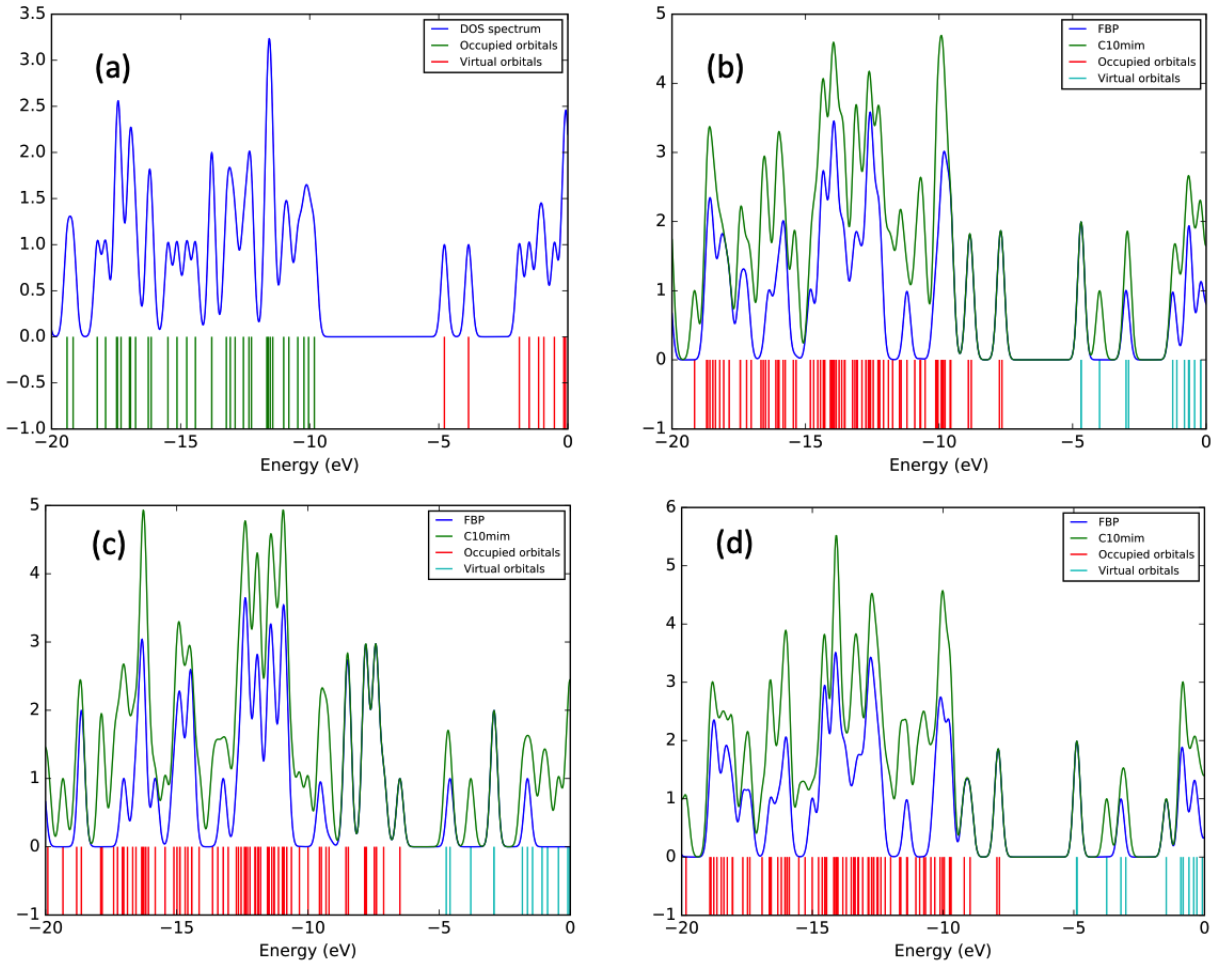


Figure 5.18: Density of States of (a) $[C_{10}mim]^+$ (b) $[C_{10}mim]^+$FBP tail up (c) $[C_{10}mim]^+$FBP tail down (d) $[C_{10}mim]^+$FBP interplanar complexes

5.1.5 Binding Energies in the presence of FBP and FeP

Binding energies for various conformations as a function of the alkyl chain were computed using equation III..82. Relative binding energies, taking into account the basis set superposition error, of the conformations with respect to the most stable conformation for a given ionic liquid cation are reported in the presence of FBP are listed in Table 5.1 and those for FeP are provided in Table 5.2.

For all the cations except $[C_4mim]^+$, the most stable conformation (relative binding energy = 0 kcal/mol) is the one in which the starting conformation of cations is interplanar. Although the tail up conformation is predicted to be the most stable for $[C_4mim]^+$, the interplanar conformation is destabilized only by 1.6

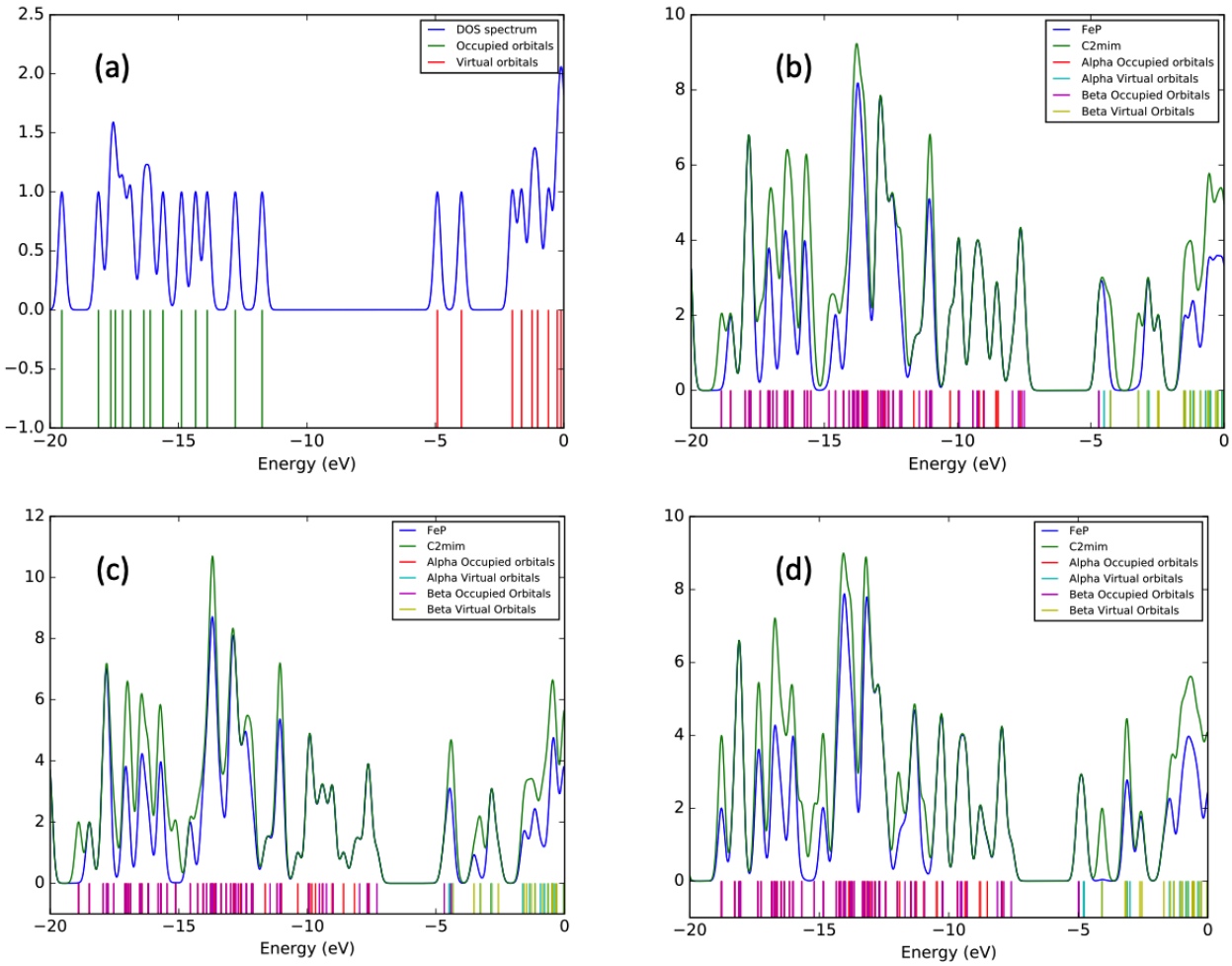


Figure 5.19: Density of States of (a) $[C_2mim]^+$ (b) $[C_2mim]^+$FeP tail up (c) $[C_2mim]^+$FeP tail down (d) $[C_2mim]^+$FeP interplanar complexes

Table 5.1: Relative binding energies (kcal/mol) of different configurations in the presence of FBP at B3LYP

Cation	Tail Up	Tail Down	Interplanar
$[C_2mim]^+$	1.5	9.2	0.0
$[C_4mim]^+$	0.0	9.9	1.6
$[C_6mim]^+$	2.5	0.9	0.0
$[C_8mim]^+$	4.1	14.7	0.0
$[C_{10}mim]^+$	4.1	14.2	0.0

kcal/mol. The tail down conformation is destabilized by 9 kcal/mol with respect to the most stable conformation for all the cations except $[C_6mim]^+$ for which all the three conformations are very similar in binding energy. The significant stabilization of the interplanar and tail up conformations over the tail down conformation

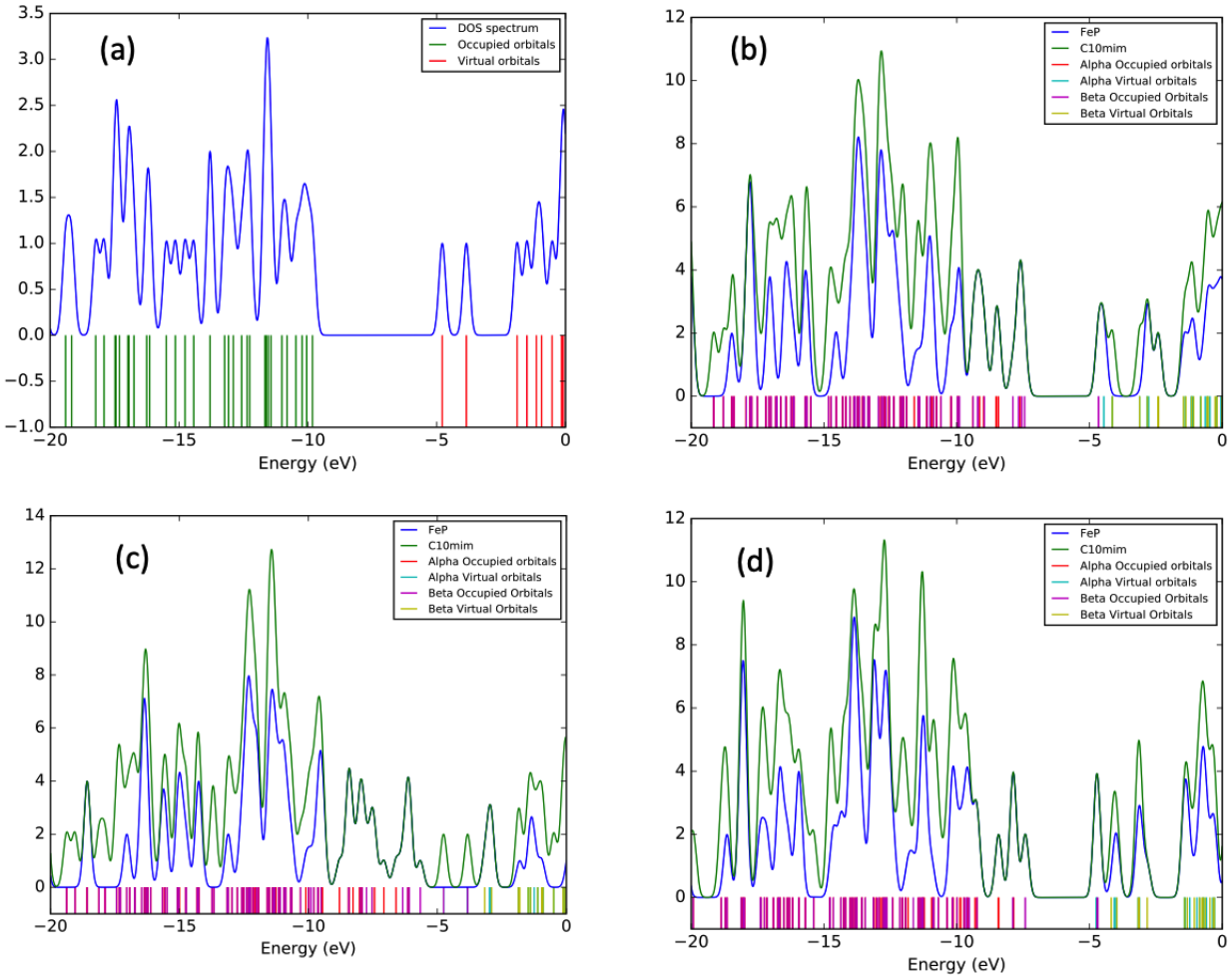


Figure 5.20: Density of States of (a) $[C_{10}mim]^+$ (b) $[C_{10}mim]^+$FeP tail up (c) $[C_{10}mim]^+$FeP tail down (d) $[C_{10}mim]^+$FeP interplanar complexes

Table 5.2: Relative binding energies of different conformations of the ionic liquid cations in the presence of FeP at B3LYP

Cation	Tail Up	Tail Down	Interplanar
$[C_2mim]^+$	4.9	4.5	0.0
$[C_4mim]^+$	0.4	1.0	0.0
$[C_6mim]^+$	2.5	9.2	0.0
$[C_8mim]^+$	1.3	9.7	0.0
$[C_{10}mim]^+$	2.1	4.1	0.0

for $[C_2mim]^+$ ($[C_4mim]^+$) is due to the optimum interaction of C_5 with the nitrogen atoms of FeP in the interplanar and tail up conformations (average distance is 3.55 Å (3.66 Å), 3.92 Å (3.94 Å)) while the average distance of C_5 is 5.32 Å (5.08 Å) for the tail down conformations. The interaction of the three conformations in

$[C_6mim]^+$ with FeP is established through the C_2 moiety which is located at an average distance of 3.95 Å, 3.94 Å, and 3.97 Å for the interplanar, tail up, and tail down conformations yielding binding energies that are within 2.5 kcal/mol. The loss in the favorable acid-base interactions in $[C_8mim]^+$ and $[C_{10}mim]^+$ leads to a large unfavorable binding energies for the tail down conformations.

In contrast to the observations for FBP, the most stable ionic liquid conformation in the presence of FeP varies with the alkyl chain length. For example, the tail up conformation is predicted to be the lowest binding conformation for $[C_2mim]^+$ and $[C_8mim]^+$ while the tail down conformation is the most stable binding mode for $[C_4mim]^+$; the calculations suggest that the interplanar conformation interacts most strongly with FeP for $[C_6mim]^+$ and $[C_{10}mim]^+$. Similar to the observation made in the FBP case, the tail down conformations are observed to be the most unstable for all the cations except $[C_4mim]^+$. Furthermore, the relative binding energy difference between the tail down conformation and the most stable conformation is relatively high (6-9 kcal/mol) for $[C_6mim]^+$, $[C_8mim]^+$, and $[C_{10}mim]^+$. In the context of ionic liquid biodegradation, hydroxylation of the alkyl chains is possibly related to weaker binding of the cation in the binding pocket of P450 enzyme. The weak binding mode is more conducive to the downstream step of bonding of molecular oxygen to FeP. This is hypothesized based on the fact that the binding of imidazolium cation is expected to occur in competition with dioxygen following the steps of binding and first reduction of the complex (Figure 2.2). The weak binding of TD complexes as compared to the other conformations is expected to assist in the dioxygen binding and aid in the functioning of the catalytic cycle.

5.1.6 Charge Transfer Analysis and Distribution

In order to explain the relative binding energy trends observed above, the extent of interaction was quantified in terms of the amount of charge transferred between the cations and either FBP or FeP. For this analysis, ChelpG charges for the cation gas phase, cation-FBP and cation-FeP compexes were determined. The resulting charges on the cations in the presence of FBP and FeP are provided in Table 5.3

and Table 5.4, respectively. It can be observed that the charge on the cation decreases from the gas phase value of +1 when cations interact with either FBP or FeP indicating that there is a net transfer of electronic density from FBP and FeP to the cations.

Table 5.3: Total charges on the imidazolium cations in the presence of FBP

Cation	Tail Up	Tail Down	Interplanar
$[C_2mim]^+$	0.77	0.75	0.75
$[C_4mim]^+$	0.78	0.78	0.79
$[C_6mim]^+$	0.78	0.72	0.73
$[C_8mim]^+$	0.78	0.94	0.73
$[C_{10}mim]^+$	0.78	0.94	0.73

Table 5.4: Total charges on the imidazolium cation after charge transfer in presence of FeP

Cation	Tail Up	Tail Down	Interplanar
$[C_2mim]^+$	0.79	0.81	0.78
$[C_4mim]^+$	0.79	0.75	0.79
$[C_6mim]^+$	0.80	0.94	0.77
$[C_8mim]^+$	0.81	0.94	0.79
$[C_{10}mim]^+$	0.80	0.96	0.80

In line with the relative binding energy trends, the extent of charge transfer is related to the conformations presented to the porphyrin ring. For example, in the case of FBP (Table 5.3) The amount of charge transfer is almost identical (~0.22e) for all the alkyl variants in the tail up conformations which can be rationalized by the fact that the atomic site C_5 is presented to FBP in all the cases. For the interplanar conformations, the net charge transfer is similar for the $[C_2mim]^+$ and $[C_4mim]^+$ cations (~0.22-0.25e), but a greater reduction in cation charge can be noted for $[C_6mim]^+$, $[C_8mim]^+$, and $[C_{10}mim]^+$. For the latter cations, the most acidic carbon C_2 is located the closest to FBP while C_5 occupies the site closest to FBP for the first two cations in the series. The C_2 position is known to be the most acidic site in the imidazolium cation and the proximity of this site to the basis

molecules such as porphyrins considered in this study enables strong binding via acid-base interactions. As opposed to the tail up and interplanar conformations, the tail down conformations are marked with a considerable variation in the extent to which charge transfer takes place. The amount of reduction in the overall charge on the cation is similar to those for the tail up and interplanar conformations for the first three cations in the series; however, there is a dramatic decrease (from ~ 0.22e to a meagre ~0.06e) for for $[C_8mim]^+$ and $[C_{10}mim]^+$ which is directly correlated to a high destabilization of those conformations. Such a change in the charge transfer is related to the fact that the electron deficient moieties (either C_5 or C_2) are positioned close to FBP for $[C_2mim]^+$, $[C_4mim]^+$, and $[C_6mim]^+$ facilitating charge transfer; however, the terminal carbon atoms in the alkyl chain which is actually an electron donating group, are exposed to FBP for $[C_8mim]^+$ and $[C_{10}mim]^+$. To give the reader an idea of the local charges on the atoms in gas phase, the atomic charges for all of the cations in the homologous series have been populated in Table 5.5.

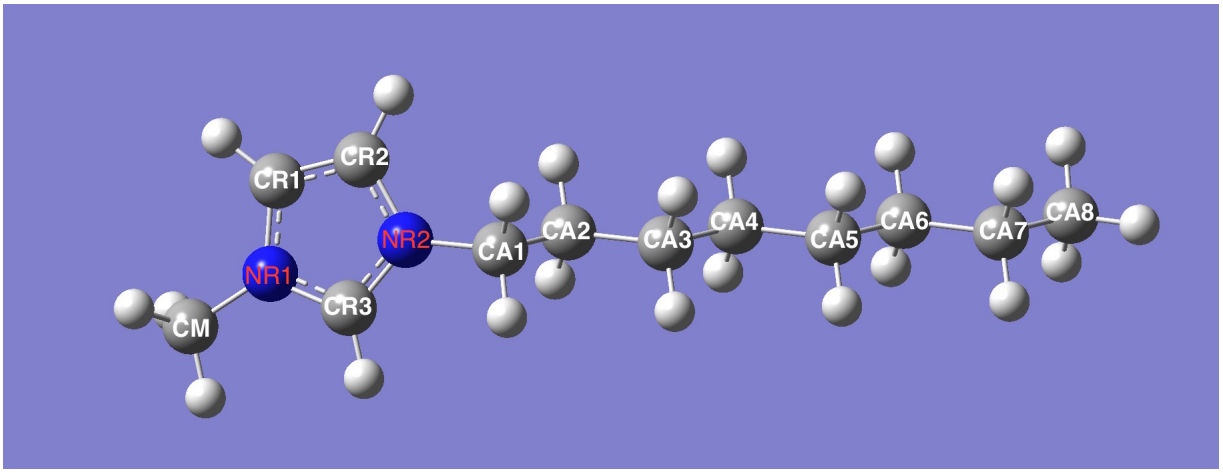


Figure 5.21: Atom numbering scheme in the charge profiles for imidazolium cations

On comparing the charges on individual atoms on the imidazolium cations with different side alkyl chain length, results show that the effect of the transfer process is highly dependent on the geometrical orientation of the cation with respect to the porphyrin molecule. This is highlighted by the fact that electrostatic potential of the overall system is highly positional in nature and depends on the radial

Table 5.5: Region Specific charge profiles for Imidazolium cations in gas phase using ChelpG scheme

Cation Atom	$[C_2mim]^+$	$[C_4mim]^+$	$[C_6mim]^+$	$[C_8mim]^+$	$[C_{10}mim]^+$
CR1	0.0943	0.081	0.090	0.083	0.082
CR2	0.150	0.002	0.007	0.0165	0.0125
CR3	0.076	0.064	0.067	0.074	0.072
CM	0.200	0.192	0.198	0.194	0.195
CA1	0.209	-0.019	0.037	0.025	0.022
CA2	0.076	0.105	0.129	0.085	0.093
CA3	*	0.138	-0.003	0.072	0.032
CA4	*	-0.042	-0.018	-0.007	-0.006
CA5	*	*	0.134	-0.024	0.104
CA6	*	*	-0.056	0.021	-0.032
CA7	*	*	*	0.102	0.112
CA8	*	*	*	-0.064	-0.043
CA9	*	*	*	*	0.083
CA10	*	*	*	*	-0.063
NR1	0.151	0.191	0.184	0.185	0.175
NR2	0.041	0.286	0.231	0.236	0.215

distance between atoms. Hence, the members of the system that are adjacent to the effect sites are more directly affected rather than the distant members. Charge distribution plots are shown for all the configurations considered. As expected, for tail up conditions, the atoms that are most affected are the methyl moiety at the vicinity of the porphyrin and the ring atoms. This detail is shown by all the alkyl variants and the degree of participation of the five membered ring is significantly higher than the alkyl chain connected to it.

For tail down conditions, the transfer effect tapers throughout the alkyl chain with the greater end at the terminal carbon and the ring is much less affected. This is in contrast to the tail up situation explained earlier. A key result to be noted here is that although the effect on the heavier atoms on the cation are pronounced, except the ring nitrogens, the net effect is distributed. Thus, it may be concluded that population of electronic transfer effects is widespread among the non-ring members. For interplanar configurations, the effect seems to distribute itself throughly along the body of the cation and all the member atoms tend to participate in the process. It may be expected as the two chemical entities (the

cation and the basic porphyrin) are co-facial in the beginning of the optimization process and then there is sliding tendency of the cation due to $\pi - \pi$ interactions between the imidazolium ring and its porphyrin counterpart.

Analysis of the same nature on FeP systems elucidated that the iron core at the nexus of the porphyrin molecule shows that the presence of Fe reduces the net charge transfer process Table 5.4. The qualitative behavior of the cation charge variations exhibited by the FBP interactions owing to geometrical dependence is reproduced with FeP complexes as well. Owing to the fact that the electrophilic iron is able to attract partially the electron cloud that is formed by the four surrounding nitrogen atoms, the LUMO of the FeP system is located at the metal centre. The interplay of the d-orbitals in iron in the energy and charge related process may be attributed to the fact that the filled and unfilled lobes are almost at the similar energy levels. This was verified by exploring the individual orbitals during gas phase calculations and calculating the energy of these orbitals which turn out to be discrete energy-wise only at 1/10000 th of an Hartree, which is insignificant in terms of our analyses. Also, in presence of FeP, different potential cationic sites show a charge difference of or more than 0.1, which makes them electrostatically important in our study.
The same attributes as shown by the FBP interactions owing to geometrical dependence were reflected by the FeP complexes. The participating sites on the porphyrin molecule seem to be dispersed throughout its plane and space of existence. Interestingly, the donor and accepting sites form distinct semicircular regions on the porphyrin plane that contribute to the charge transfer process.

The final charges on each atom on the cation was also evaluated in the complex and compared to the initial charges in isolation. The atom numbering scheme is given in Figure 5.21 for reference. The ring atoms in case of imidazolium are the three carbons (CR1-CR3) and the two nitrogens (N1 and N2). Excluding the ring atoms, the others are the methyl carbon (CM) and the alkyl chain atoms (CA). The number of chain atoms increase with the increase in the chain length of the

cation considered. First we discuss the FBP containing complexes for all the three conformations, then we follow with the FeP ones. In our discussion, the sites with the three highest charge differences will be mentioned and their placement would be mentioned. This is done to reflect the conformational dependence on the binding and the exposure of the cationic sites to porphyrin. Firstly, for tail up complexes, along the homologous series, the CM atom shows the greatest difference that is closely followed by the ring carbons, CR1 and CR3. The difference in charge is minimal at about 0.1 but shows the importance of conformational preference for the cations. For tail down complexes, the charge distribution is such that the difference is reflected for both the ring atoms and chain atoms for ethyl, butyl and hexyl containing cations. For the two greatest cations in complexation, the charge transfer to FBP is negligible. For $[C_2mim]^+$, the three most dominant sites for charge transfer are CR1, N2 and CA2. The CA3 atom participates in both $[C_4mim]^+$ and $[C_6mim]^+$ cation in complex with FBP. Finally, for interplanar complexes, the ring is totally exposed to the FBP underneath. The significant differences are shared between ring and alkyl chain atoms but primarily with the ring ones. The $[C_2mim]^+$ shows the participating sites to be N1 and CA1 while for butyl, all the sites are in the ring. Hexyl, octyl and decyl cations show the methyl carbon to be the most active among the others. Similar inferences were drawn in case of FeP containing complexes. For all of the members in the homologous series except $[C_6mim]^+$, the TU complexes show the two ring nitrogens, N1 and N2 to be most dominant in the charge transfer process. Similar to FBP containing complexes, the charge distribution for tail down complexes are interspersed between the ring and the alkyl chain members. The two smallest members in the cationic series, ethyl and butyl show significant charge transfer while for the the others, it is negligible. For the ethyl cation, the three dominant sites are CA1, CR1 and CR2 while for the $[C_4mim]^+$ containing complex, they are CR3, N1 and CA3. The interplanar complexes show charge differences in both ring and alkyl chain members. The only alkyl chain members that are shown to participate are CA1 and CA2 with all the others maintaining their total charge as in isolation. With the exception of the $[C_2mim]^+$ containing complex, all others show charge

transfers from N1 and N2 atoms in the ring.

5.1.7 Reactivity Indices

The values of the index calculated using eq. III..81 with the HOMO and LUMO values of FBP and FeP are reported in Tables 5.6 and 5.7 for the ionic liquid cations in complex with FBP and FeP, respectively. The gas phase electrophilicity index for bare FBP is less positive than that for FeP indicative of the fact that the electron transfer to FeP is more likely due to the introduction of electron deficient Fe. The presence of the ionic liquid cations leads to a dramatic increase in the electrophilicity index for both FBP and FeP, especially for the tail up and interplanar conformations. The electrophilicity index is fairly constant across the entire homologous series for these conformations. The tail down conformations for $[C_2mim]^+$ and $[C_4mim]^+$ also yield electrophilicity index that are markedly lower than the corresponding FBP and FeP values. On the other hand, the tail down conformation in $[C_8mim]^+$ and $[C_{10}mim]^+$ results in only a modest increase in the electrophilicity for FBP. Similarly, the tail down conformatio for $[C_6mim]^+$ is predicted to increase the electrophlicity for FeP only margianlly, no increase for $[C_8mim]^+$, while a decrease in the ability to acquire electrons for $[C_{10}mim]^+$. The trends obtained for the electrophilcity index are directly correlated with the amount of charge transfer to the cation. As shown in Tables 5.3 and 5.4, the greater the charge transfer, the more positive is the electrophilicity index as the transfer of electron density to the cation makes FBP and FeP electron deficient facilitating transfer of an electron.

Once the substrate is bound to FeP, the subsequent step in the P450-mediated hydroxylation is the transfer of electron from the reductase domain to the enzyme active site. Our calculations suggest that the ionic liquid conformations presented to the active site can potentially modulate this behavior. In all the cases investigated, only $[C_{10}mim]^+$ in the tail down conformation leads to a decrease in the electron acquiring ability of the active site. However, this reduction is not

Table 5.6: Parr's Electrophilicity Indices for FBP in the presence of imidazolium cations (in kcal/mol)

Cation	Gas Phase	Tail Up	Tail Down	Interplanar
$[C_2mim]^+$	64.5	149.3	148.5	157.4
$[C_4mim]^+$	64.5	148.4	145.8	151.3
$[C_6mim]^+$	64.5	147.9	159.2	158.4
$[C_8mim]^+$	64.5	147.4	92.8	158.4
$[C_{10}mim]^+$	64.5	148.2	86.7	157.9

Table 5.7: Parr's Electrophilicity Indices for FeP in the presence of imidazolium cations (in kcal/mol)

Cation	Gas Phase	Tail Up	Tail Down	Interplanar
$[C_2mim]^+$	83.2	134.5	134.9	149.3
$[C_4mim]^+$	83.2	135.2	145.1	146.8
$[C_6mim]^+$	83.2	133.4	95.4	140.6
$[C_8mim]^+$	83.2	132.5	83.6	146.5
$[C_{10}mim]^+$	83.2	133.0	76.0	142.4

pronounced. Experimentally, hydroxylation of $[C_{10}mim]^+$ has been reported indicating that electron transfer step is not impeded although there is a decrease in the propensity for FeP to accept electrons. Our calculations further imply that the transfer of electrons to the active site is probably not the cause for low biodegradability of $[C_2mim]^+$ and $[C_4mim]^+$.

Upon the culmination of our discussion of the binding of imidazolium-based cation to both FBP and FeP at the B3LYP level, it was deemed important to include the effect of dispersion into the system. This was done due to the fact that these long range interactions become important in determining the stability of the overall complex in case of biological macromolecules and complexes. Their contribution although being small for individual atomic pairs, adds up to become significant to account for the correction to the underlying wavefunction for the entire complex. The inclusion of dispersion corrections were made from two different approaches, i.e. empirical and non-empirical. Traditionally, B3LYP has been shown to not be capable in capturing the above said dispersion effects by

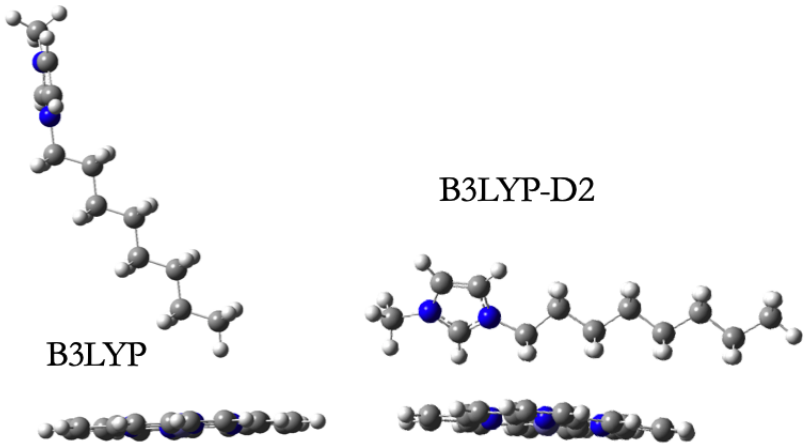


Figure 5.22: Comparison of optimized structure for $[C_8mim]^+$FeP TD complex from B3LYP and B3LYP-D2 levels of theory

itself. The interested reader is directed towards the works by Grimme et al. and co-workers [119, 120, 121] that incurred the development of various corrections to the basic DFT framework. The second theory was applied to probe the effect of a functional that is parametrized to handle dispersion interactions accurately. For this, the M06 functional [122, 126] was chosen due to its ability and performance on transition metal complexes and various other other molecules of biological interest and .

Our analysis regarding the comparison of conformations obtained from both B3LYP and B3LYP-D2 functionals clearly show that there is significant change in the optimized geometry in case of TD complexes. The slithering effect of the alkyl chain is not captured by B3LYP which renders the complex having an unfavorable relative energy as compared to the TU and IP ones. Also, it is worth noting that the conformation that is assumed by the cation in Figure 5.22 is very similar to the one that has been observed for natural substrate like N-palmitoylglycine (NPG) [9]. This is a key piece in the conformational puzzle that is center to the thesis work. It confirms the fact that inclusion of dispersion is important in uncovering physically realizable geometries shown by other substrates. Also, this gives an idea that TD conformations are important in understanding the regioselectivity

of the substrate that would eventually lead to hydroxylation.

5.2 Cluster Model1: B3LYP-D2 and M06

5.2.1 Binding Energies in the presence of FBP

Figure 5.23 presents the binding energies calculated using eq. III..82 for the ionic liquid cation-FBP complexes as a function of the alkyl chain length for the two functionals, B3LYP-D2 and M06 used in this work. All of the reported binding energies for the complexes are shown to be negative to make the binding process of the cations thermodynamically favorable. On the relative scale of stability, the first two cations ($[C_2mim]^+$ and $[C_4mim]^+$) show the tail-up geometries to be the most favorable among the three distinct conformations, within the series of cations considered for the work.

Table 5.8: Relative binding energies (kcal/mol) of different configurations in the presence of FBP at B3LYP-D2

Cation	Tail Up	Tail Down	Interplanar
$[C_2mim]^+$	0.0	1.4	0.5
$[C_4mim]^+$	0.0	0.6	0.7
$[C_6mim]^+$	4.2	2.4	0.0
$[C_8mim]^+$	3.5	7.6	0.0
$[C_{10}mim]^+$	2.2	16.3	0.0

The stability of the complexes although, gets switched to interplanar ones as the chain length is enhanced for butyl to hexyl and above. Upon considering a quantum chemical accuracy of 2 kcal/mol, it can also be said that there is a clear preference of the interplanar conformations to be most stable among all of three. The destabilization on the relative scale is shown to be dependent on the alkyl chain length as well. It can be noticed that while the energetic separation between interplanar and tail complex reduces upon increasing the cation size, the opposite is relevant for the tail down conformations. The absolute binding energies reported for the conformations are different for both of these functionals, but importantly both of them capture the relative energetics to a high degree of mutual accuracy.

Tables 5.8 and 5.9 report the above mentioned relative energies for all of the conformations with the most stable one bearing the datum as zero.

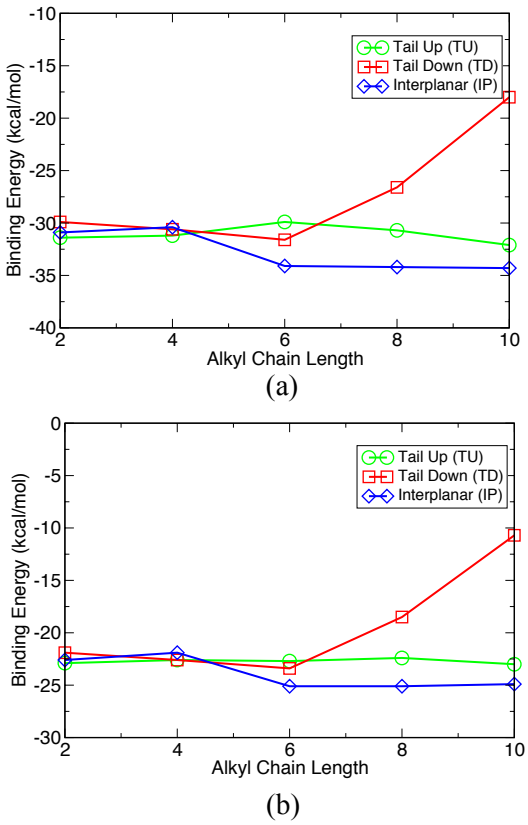


Figure 5.23: Binding Energies of FBP containing complexes at (a) B3LYP-D2 and (b) M06 levels of theory

To understand the levels of binding energies predicted at different levels of theory, the same energy from all three theories, namely, B3LYP, B3LYP-D2 and M06 were compared for both FBP and FeP containing complexes. Figure 5.24 gives the comparative trends of binding for all the three different kinds of conformations presented to FBP. it can be clearly observed that the relative energies are captured by all of the three levels of theory, but the highest favorability is presented by B3LYP-D2, which is followed by M06. Due to the absence of dispersion corrections, the pure B3LYP theory predicts binding energies that are highly unfavorable as compared to the other two theories. A similar observation is made for FeP containing complexes that will be discussed in the following subsection.

Table 5.9: Relative binding energies (kcal/mol) of different configurations in the presence of FBP at M06 level of theory

Cation	Tail Up	Tail Down	Interplanar
$[C_2mim]^+$	0.0	0.9	0.3
$[C_4mim]^+$	0.0	0.0	0.6
$[C_6mim]^+$	2.3	1.7	0.0
$[C_8mim]^+$	2.7	6.6	0.0
$[C_{10}mim]^+$	1.9	14.3	0.0

An alignment of the geometries obtained from both the levels of theory in this section, B3LYP-D2 and M06, are given in Figures A1 - A3. These are the optimized geometries obtained from both of these theories and the alignment is performed by minimizing the root mean square distance for the heavy atoms in the porphyrin macromolecule. The final geometries for $[C_2mim]^+$ and $[C_4mim]^+$ interplanar complexes are such that the cationic ring remains co-facial to FBP. For all the other cations in the homologous series, the imidazolium ring is slightly tilted to help in the exposure of the most acidic proton to the underlying FBP. Energetically, this change in orientation is reflected by yielding a more favorable binding energy for the cation as compared to the other conformations. The alkyl chain also bends towards the active site molecule to enhance its interactions from the native position with FBP.

The resulting geometries from optimization of initial tail up conformations show the rings to be in perfectly co-facial position as compared to the underlying FBP. As this interaction becomes the most dominating amongst the ones geometrically possible for the complexes, the binding energies reflect the energetic closeness between them. All of the members of the cationic series give very similar binding energies. There seems to be no clear preference of the alkyl chain orientation for these conformations, in which most of the chains are pointing away from the underlying molecule. Some of the members also show a parallel chain and such a structure was also tested for a local minima in the case of FeP and the resulting analysis pointed out that it in fact was not a local minimum but actually a global

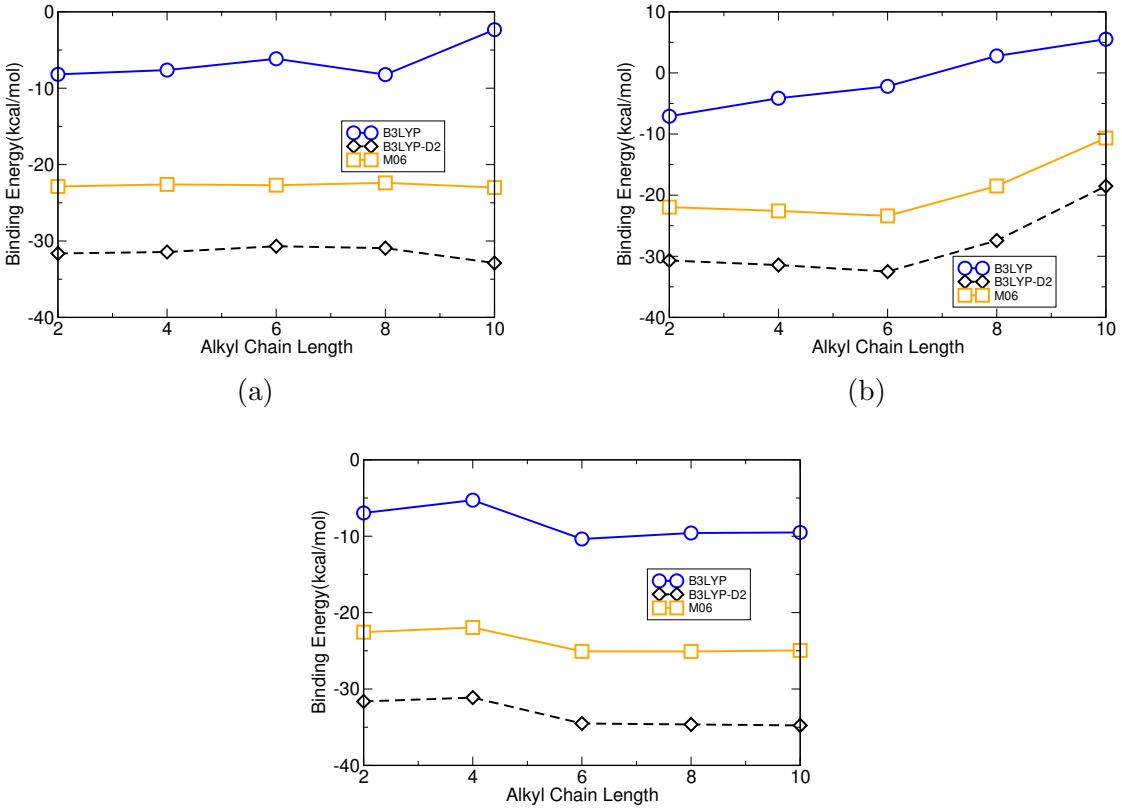


Figure 5.24: Comparison of binding energies evaluated at different levels of theory for $[C_n mim]^+$FBP (a) TU (b) TD (c) IP complexes

minimum. Such perturbative analysis is essential in case of disputes in geometries and checking for the sanctity of the binding regimes involved and their uncertainty.

The geometries from tail down structures for $[C_2mim]^+$-FBP and $[C_4mim]^+$-FBP are similar to the ones from interplanar and tail up conformations and lead to similar binding energies. The gradual increase and exposure of the alkyl chain to the porphyrin causes a shift in the ring from the central domain of the porphyrin and significantly changes the binding regime to be less favorable. This dramatic destabilization is reflected in the relative energies of the $[C_8mim]^+$ and $[C_{10}mim]^+$ tail down conformations as compared to the others. From our analysis, it appears that the strong binding between the cation and FBP is directly correlated with the interaction of the imidazolium ring with porphyrin ring.

Figure 4.2 in the methodology section gives the various energetic contributions

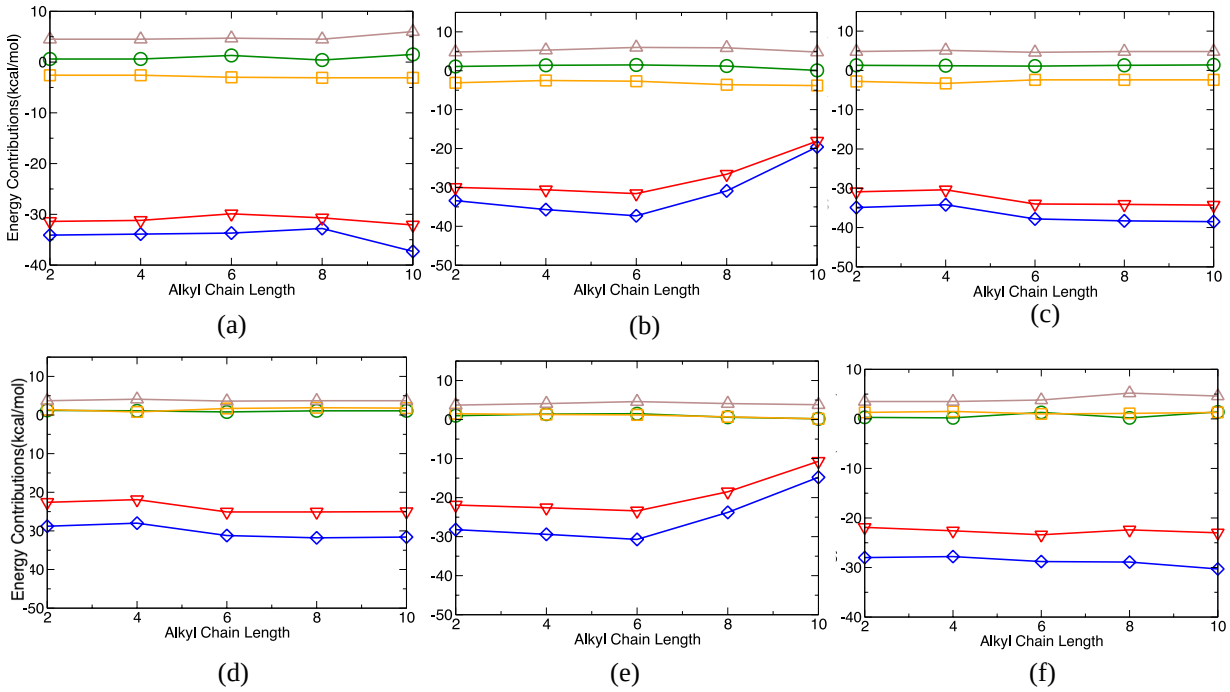


Figure 5.25: Binding energy contributions for $[C_n mim]^+$FBP complexes at B3LYP-D2 (a) TU (b) TD (c) IP and M06 theories(d) TU (e) TD (f) IP (Legends: $\Delta E_{c,cation}$ ○, $\Delta E_{c,FBP}$ □,ΔE_i ◇,$\Delta E_{cp,complex}$ △, $\Delta E_{binding}$ ▽)

to the binding energies for a typical process. These contributions were depicted for both the FBP and FeP containing complex to check which of them is the key indicator of the overall binding strength of the complex. Figure 5.25 (a-f), reflect that the interaction energy between a cation and FBP makes the predominant contribution to the binding energy of the overall complex.

The straining of the cations is relevant as the conformational rearrangement is shown to have a positive energy change as compared to their gas phase energies in isolation. Its counterpart is true for the porphyrin molecules where the induced conformational changes in FBP result in an energy gain when B3LYP-D2 functional is employed. The counterpoise correction is small but positive and reflects the over stability of the wavefunction when the two monomers come into close contact and assume their preferred geometries in the complex. The interaction energy is the only piece among all of the others that shows the resulting dependence

on the geometries. These interactions energies are shown to be only marginally affected in case of the tail up and interplanar geometries, as is indicated by the optimized geometries. A significant effect of the interaction energy on the overall binding energy was noted only for the conformations which were optimized starting from a tail-down geometry. In line with the binding energies discussed above, the interaction energies rise when the alkyl chain is octyl or decyl, which can be correlated to the structural differences *vide supra*.

The same energetic contributions from the M06 level of theory also reflect the conclusions drawn from B3LYP-D2 in the paragraph above. On an absolute scale, the M06 functional predicts less favorable binding energies as compared to B3LYP-D2 but the relative energetics are captured in a similar manner. Figures 5.25 (d-f) give these contributions for the cations in the series and show that the only exception to the former D2 containing level is the conformational energy associated with the deformation of FBP is positive. It disfavors the binding process and results in the overall binding energy to be less favorable. It is to be noted, however, that the contribution is rather small *ca.* 1 kcal/mol.

5.2.2 Binding Energies in the Presence of FeP

The binding energies obtained from the cation-FeP complexes are given in Figure 5.26. It is clearly observed that the association of FeP with the ionic liquid cation is thermodynamically favorable irrespective of the chain length and conformation presented to FeP. The relative energetics of the process is given in Tables 5.10 and 5.11. In terms of their stability for $[C_2mim]^+$, $[C_4mim]^+$, and $[C_6mim]^+$, the tail up and tail down conformations are indistinguishable from the interplanar conformations if the quantum chemical accuracy of 2 kcal/mol is considered. From the relative binding energies, it is clear that except for the $[C_6mim]^+$, all the others give the interplanar conformations to be the most stable among the others. For longer alkyl chain lengths having most propensity for biodegradability in the series, the tail down conformations are relatively destabilized as compared

to the interplanar ones. The $[C_8mim]^+$ and $[C_{10}mim]^+$ show relative energies of 5.4 and 2.6 for B3LYP-D2 while 6.2 and 3.2 for M06 functional. Also, upon comparing the relative energies from FBP and FeP, it is evident that the presence of the metal atom in the core of the porphyrin leads to an enhanced stabilization of the tail facing conformations.

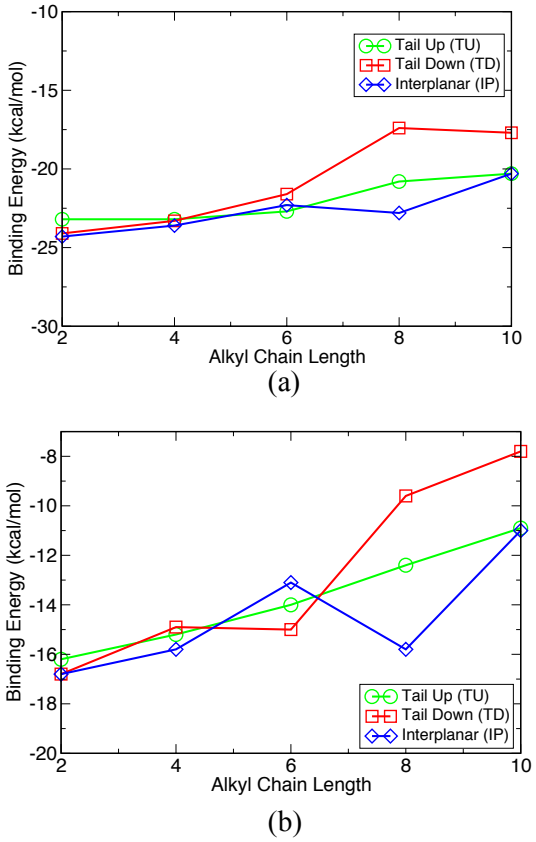


Figure 5.26: Binding Energies of FeP containing complexes at (a) B3LYP-D2 and (b) M06 levels of theory

Figures A4 - A6 give the alignments of the optimized geometries from the two functionals for all of the cationic sizes bound to FeP in complex. For the interplanar complexes, the orientation between the imidazolium ring and the FeP remain co-facial even after optimization. The position of the alkyl chain for these complexes is parallel for the cations $[C_2mim]^+$, $[C_4mim]^+$, and $[C_8mim]^+$ while it points away for the others. For the tail up conformations, the cofacial arragement is retained for all of the cations except in the case of $[C_6mim]^+$. Similar to FBP

containing complexes, the relative destabilization occurs due to the offset of the ring moving from the central domain of the porphyrin. This displacement makes the overall complex of octyl and decyl cations significantly unstable as compared to others. Taken together, these observations suggest that the binding conformations involving the imidazolium ring and FeP in the co-facial arrangement lead to enhanced stabilization of the complex. On the other hand, exposure of the alkyl chain exerts a destabilizing effect. The conformational preferences appear to be only marginally affected by the choice of the functional. The geometrical features of the $[C_n mim]^+$ cations before and after optimization for the TU and TD conformations are given in the adjoining figures (Figure 5.27 and 5.28).

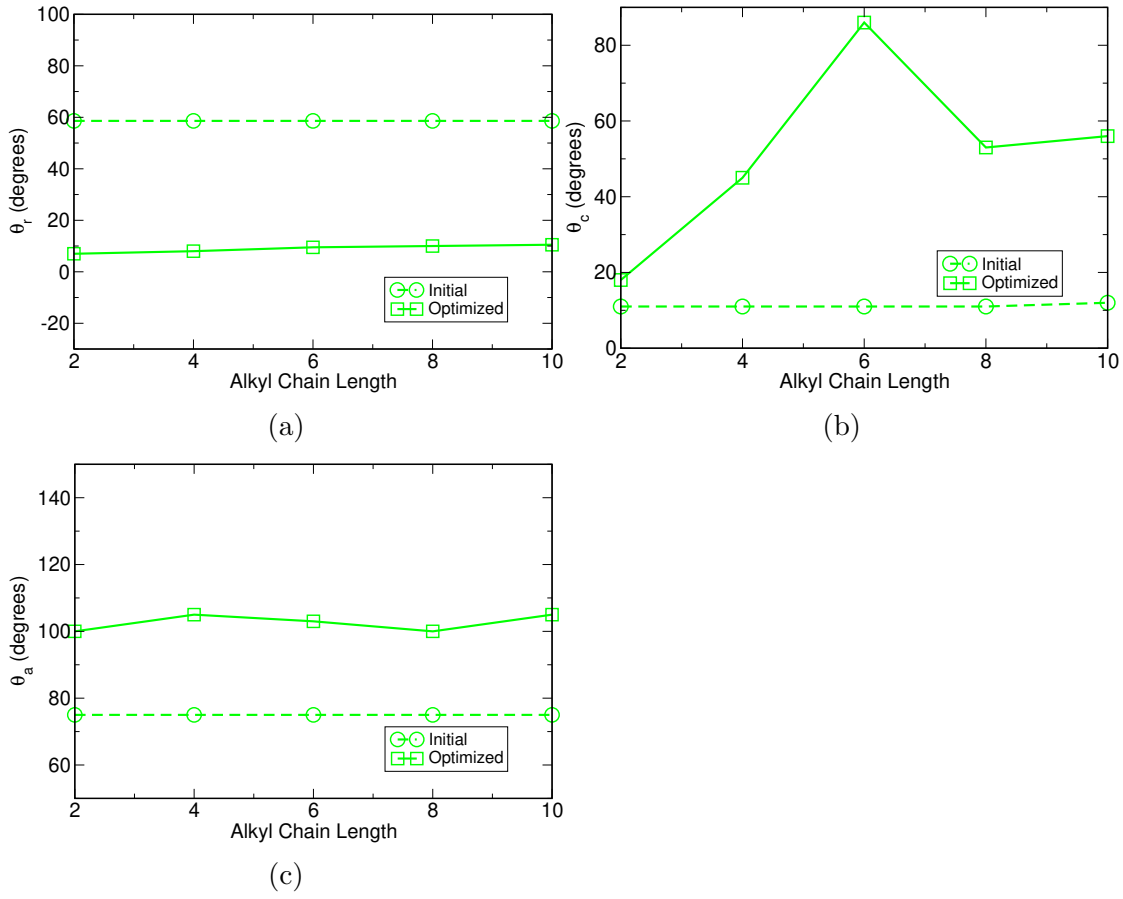


Figure 5.27: (a) Ring orientation (θ_r) (b) Chain Orientation (θ_c) (c) Azimuthal angle (θ_a) for $[C_n mim]^+$ tail up complexes optimized at M06/6-31g(d,p)

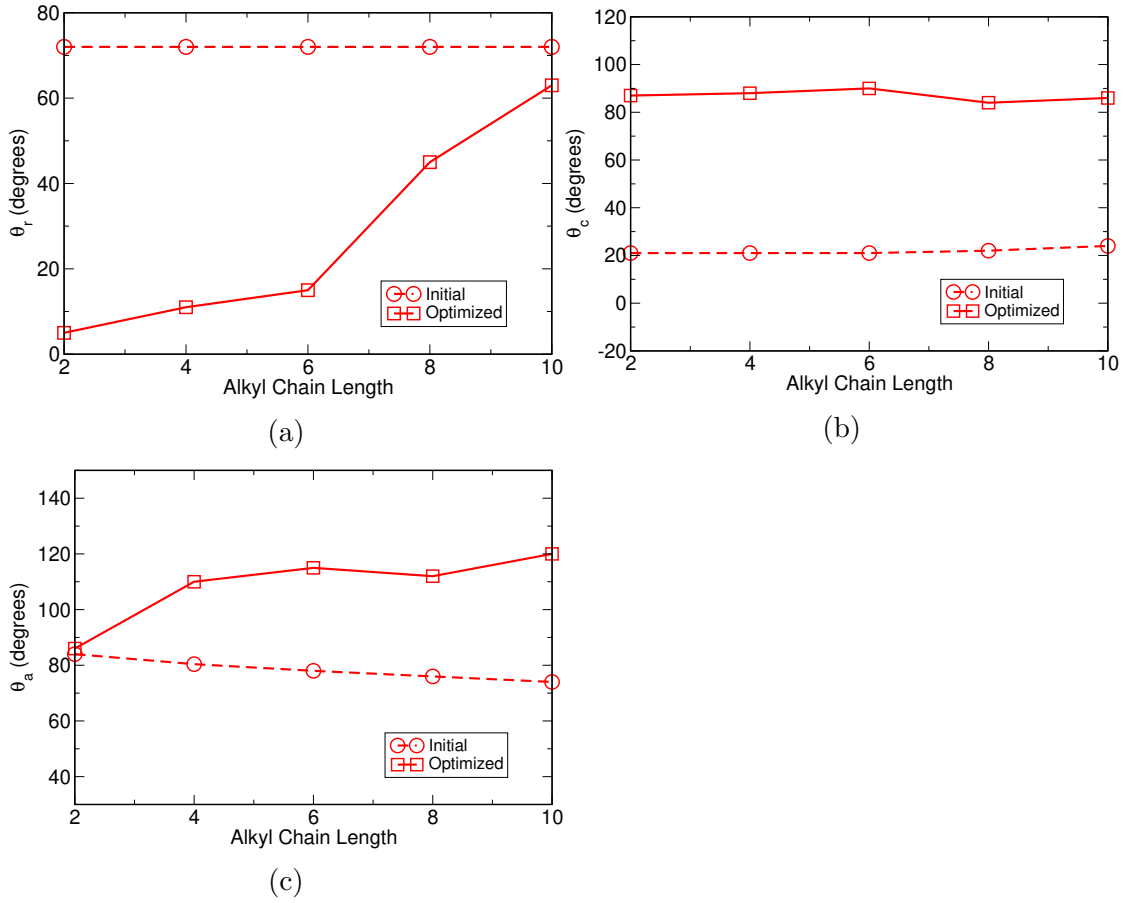


Figure 5.28: (a) Ring orientation (θ_r) (b) Chain Orientation (θ_c) (c) Azimuthal angle (θ_a) for $[C_n mim]^+$ tail down complexes optimized at M06/6-31g(d,p)

A comparative analysis on the binding energies obtained for FeP containing complexes for all the three levels of theory considered in the present work in Figure 5.29. Similar to FBP complexes, the FeP ones also show that the B3LYP-D2 theory predicts binding energies to be most favorable. This is followed by M06 and interestingly, the binding of some of the complexes are shown to be positive by B3LYP. The conclusion drawn for FBP is echoed for the metal-containing complexes where the dispersion correction makes the complexes highly stable as compared to the B3LYP optimized ones. Both B3LYP-D2 and M06 are shown to be capable of predicting the relative effects between conformations of the cation with respect to FeP. Thus, both explicit and implicit dispersion in these two levels of theory are adequate in describing the conformation of imidazolium-based cations

Table 5.10: Relative binding energies of different conformations of the ionic liquid cations in the presence of FeP at B3LYP-D2 (in kcal/mol)

Cation	Tail Up	Tail Down	Interplanar
$[C_2mim]^+$	0.8	0.3	0.0
$[C_4mim]^+$	0.4	0.3	0.0
$[C_6mim]^+$	0.0	0.6	0.4
$[C_8mim]^+$	2.0	5.4	0.0
$[C_{10}mim]^+$	0.0	2.6	0.0

Table 5.11: Relative binding energies of different conformations of the ionic liquid cations in the presence of FeP at M06 (in kcal/mol)

Cation	Tail Up	Tail Down	Interplanar
$[C_2mim]^+$	0.7	0.0	0.0
$[C_4mim]^+$	0.6	0.9	0.0
$[C_6mim]^+$	0.0	1.0	0.9
$[C_8mim]^+$	3.4	6.2	0.0
$[C_{10}mim]^+$	0.0	3.2	0.0

in complexation with FeP.

Figures 5.30 (a-c) for B3LYP-D2 and Figures 5.30 (d-f) for the M06 functional give the various energetic contributions to the overall binding energies of the cation-FeP complexes. Similar to the FBP containing complexes, the interaction energy between the monomers participating in the complex, i.e. the IL cation and the FeP tend to dominate the binding energy profile. For almost all the cases, the interaction energy tracks the trend uncovered in the overall binding energy. The other contributions arising from monomer deformation and counterpoise corrections destabilize the binding process and make the energy of association less favorable. As mentioned earlier for the non metallic systems, both the functionals predict the relative energetics to high accuracy although their treatment of the binding at an absolute scale differ significantly. The contributions other than interaction energies are shown to be not dependent on the identity of the cation, the functional employed for the analysis and the initial geometries.

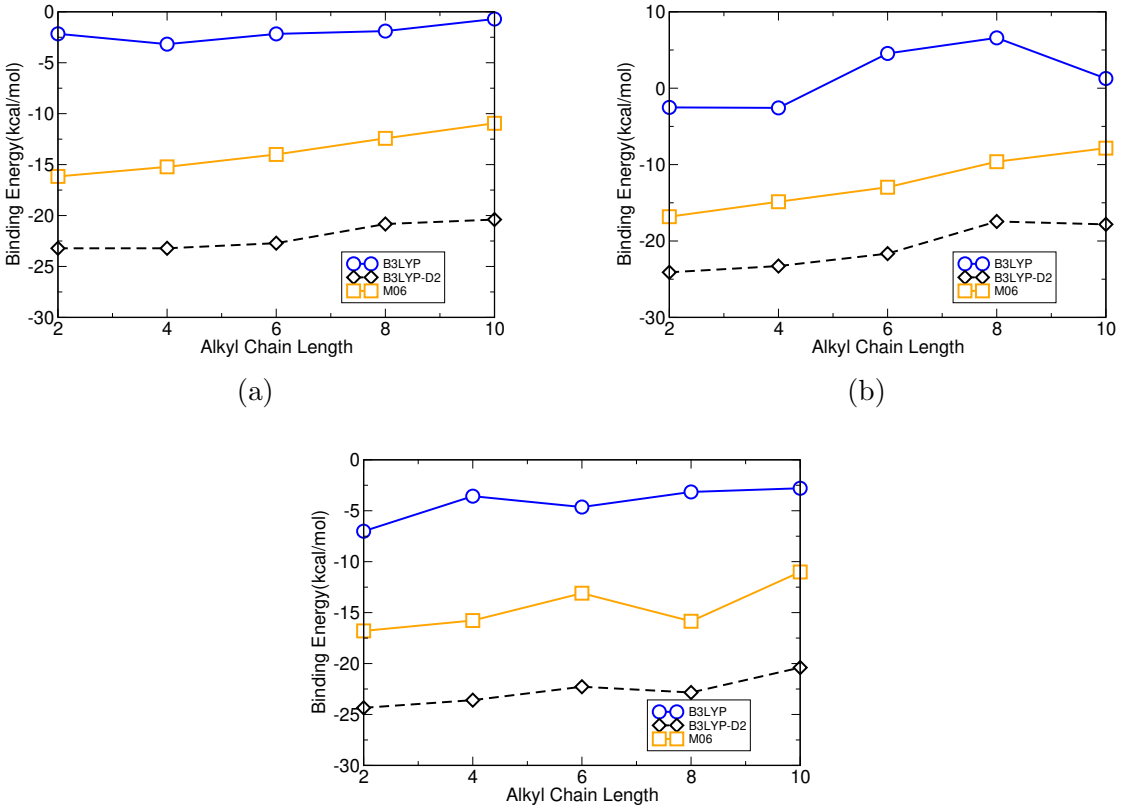


Figure 5.29: Comparison of binding energies evaluated at different levels of theory for $[C_n mim]^+$FeP (a) TU (b) TD (c) IP complexes

5.2.3 Electrophilicity Index

The binding process is followed by reduction, which involves the addition of an electron to the central active species in P450. This electron addition changes the oxidation state of the Fe involved in the porphyrin and hence, the influence of the cation on this ability was captured by conceptual DFT. The electrophilicity index was used to evaluate the same and the results from the calculations have been presented in the Figures 5.31 and 5.32 for FBP and FeP containing molecules respectively. In the absence of cations, the reactivity index of the porphyrin molecules was evaluated so that the effect of the cation maybe compared. For FBP, these values were computed to be 54 and 51.5 kcal/mol with B3LYP-D2 and M06 functionals, respectively. It is inferred from the Figure 5.31 that the reactivity index is significantly enhanced in the presence of cations where the resulting value seems to be about 110 kcal/mol higher for both tail up and interplanar conformations.

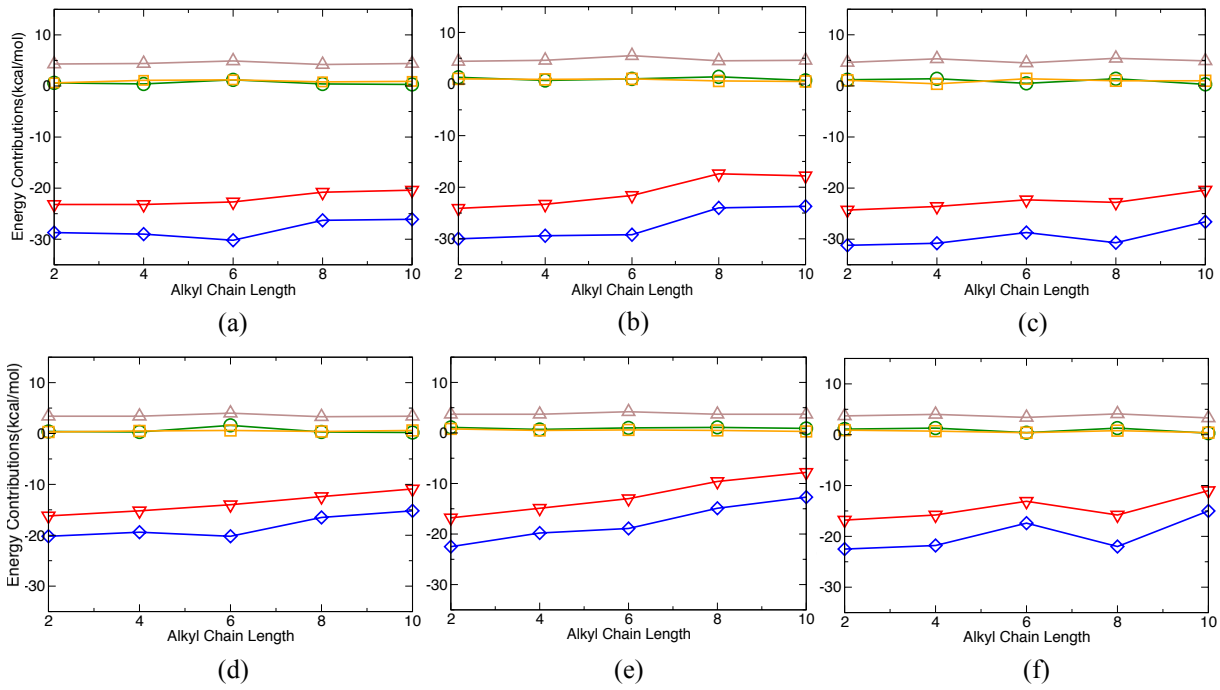


Figure 5.30: Binding energy contributions for $[C_n mim]^+$FeP complexes at B3LYP-D2 (a) TU (b) TD (c) IP and M06 theories(d) TU (e) TD (f) IP (Legends: $\Delta E_{c,cation}$ ○,$\Delta E_{c,FBP}$ □,ΔE_i ◇,$\Delta E_{cp,complex}$ △, $\Delta E_{binding}$ ▽)

This rise in the electron accepting ability is independent of the size of the cation and is caused by the presence of the positive charge concentrated on the ring and its exposure to FBP. For the tail down conformations, the electrophilicity index steadily becomes less favorable as compared to others as the alkyl chain length is enhanced from hexyl to one higher. The drop in electrophilicity index being significant still does not change the fact that it is considerably elevated keeping the gas phase one as reference. The reduction in the electron accepting ability of FBP molecule for the two cations can be rationalized based on the fact that the positively charged imidazolium ring is displaced away from FBP; the exposure of the alkyl chain to FBP molecule also serves to decrease the electrophilic index as the alkyl moiety is known to possess electron donating ability.

For FeP, the electrophilicity index was computed to be 49.5 and 48 kcal/mol for the B3LYP-D2 and M06 functionals, respectively. The presence of the metal causes the overall reductive ability of the porphyrin to be estimated as lower compared to the free base porphyrin due to lower LUMO energy values. The binding of the imidazolium cations to FeP also have a very significant effect in elevating their reducing ability as demonstrated in Figure 5.32. The rise in the electrophilicity is shown to be almost independent of the cationic size in case of the tail up and interplanar complexes. There is a significant decrease in the reactivity index for FeP as the cation size is changed from having 2 carbon atoms to 10, but that change is not as pronounced as in the case of FBP. The two key conclusions inferred from the analysis are firstly, the presence of the cation in any conformation is capable of elevating the reductive ability of underlying porphyrin. This elevation is more marked in case of FBP than FeP on an absolute scale. Secondly, for both tail up and interplanar conformations, the elevation seems to be insensitive to the change in the size of cation but for tail down geometries, the increase in the size reduces this ability.

5.3 Cluster Model 1: Ring Cationic Systems

5.3.1 Binding Energies Across Homologous Series

A detailed exploration of the conformations of imidazolium-based cations were considered in the previous section. Here we will only consider the tail up and tail down conformations and compare them with the other three systems. It must be noted that all the computational approaches obtained to perform the analysis in the previous section were repeated here and kept consistent for these systems. In the previous section dedicated to the binding of imidazolium-based cations to FBP and FeP, it was established that the interaction energies play the dominant role in establishing the binding energies of the systems and that the tail down

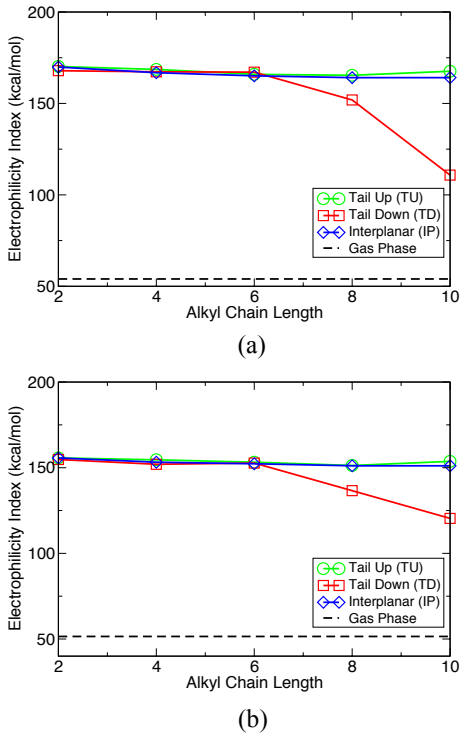


Figure 5.31: Electrophilicity Index of FBP in gas phase and complex for $[C_n mim]^+$FBP systems (a) B3LYP-D2 (b) M06

conformations become relatively unstable as compared to the other conformations as the alkyl chain length on the cation is progressively enhanced.

The binding profile of all the cations in complex is given in Figure 5.34. Firstly, the process of binding is negative for all the systems considered that shows thermodynamic favorability for the cation binding to the FeP in all conformations. For imidazolium cations (Figure 5.34(a)), these binding energies become increasingly unfavorable or less negative as the alkyl chain length is enhanced. This shows that the complexation with the larger cations is less favorable due to the greater degree of freedom presented to the active site. Conformationally, except for the $[C_2mim]^+$, all of the other cations show that the TU conformation is more stable as compared to the TD one. The ranges of the binding energies are between the bounds of -16.8 and -10.9 kcal/mol for TU systems, while the same is between

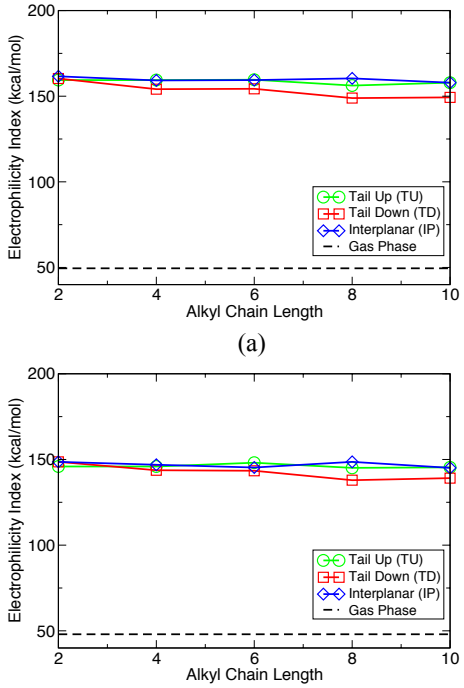


Figure 5.32: Electrophilicity Index of FeP in gas phase and complex for [$C_n$mim]-$^+$FeP systems (a) B3LYP-D2 (b) M06

-16.8 and -7.8 for the TD conformations. Upon comparing the relative energy differences, it is also evident that the destabilization of the TD geometries increase with the size of cation with the two greatest cations having relative energies of 2.8 and 3.1 kcal/mol as compared to their TU partners (Table 5.12).

The conformational preference for the pyridinium containing systems is also retained and is shown in Figure 5.34(b). The computed binding energies clearly indicate that similar to imidazolium systems, the binding energies becomes less favorable as the chain length is increased. This is shown by the range of the binding energies obtained for this class of cations while the relative destabilization of the TD geometries is also clearly expressed as the chain length is enhanced (Table 5.13) from C_2-C_{10}. The TU conformations show preferential binding as compared to the TD ones implying the effect of these conformations on the overall

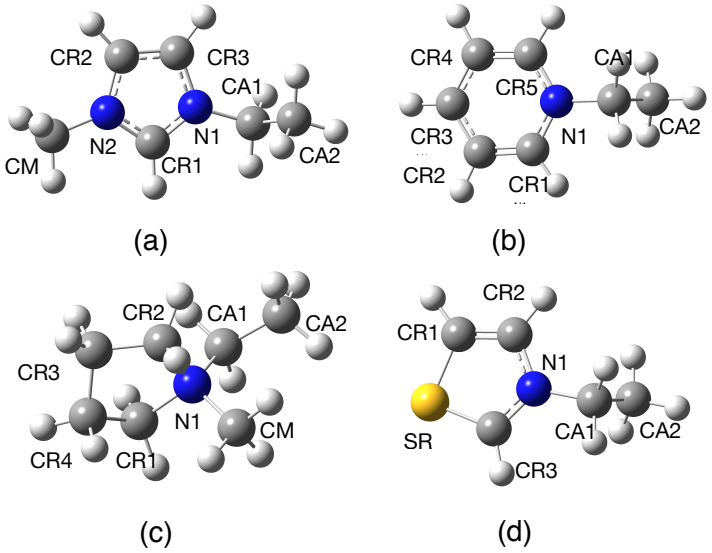


Figure 5.33: (a) Iron Porphyrin (FeP) and (b) 1-ethyl-3-methylimidazolium $[C_2mim]^+$ (c) 1-ethylpyridinium $[C_2py]^+$ (d) 1-ethyl-1-methylpyrrolidinium $[C_npyrr]^+$ (d) 1-ethylthiazolium $[C_2th]^+$ cations [Atom Color coding: H (White), C (Grey), N (Blue), S (yellow), Fe (purple) The longer alkyl chain cations are labeled as CA3, CA4, etc.

binding energies.

The next aromatic system explored was the thiazolium containing systems having a sulphur containing thiazole in their head group. This group is shown to yield a very different binding regime as compared to the to the imidazolium and pyridinium ones. The initial trend of the cations binding progressively unfavorably as the chain length is increased is maintained but the conformational trend is switched. In the case of thiazolium containing systems, the TU conformations become less favorable as compared to the TD ones and the relative instability is highly significant. Upon closer inspection, it can also be seen that the binding energies for the TU conformations ranges from -10 to -4 kcal/mol while for the TD conformations the same is -16 to -10 kcal/mol. Table 5.14 gives the relative energies of the two conformations and the TD binding is more favorable as compared to the TU ones and the maximum destabilization is about 7.6 kcal/mol. This

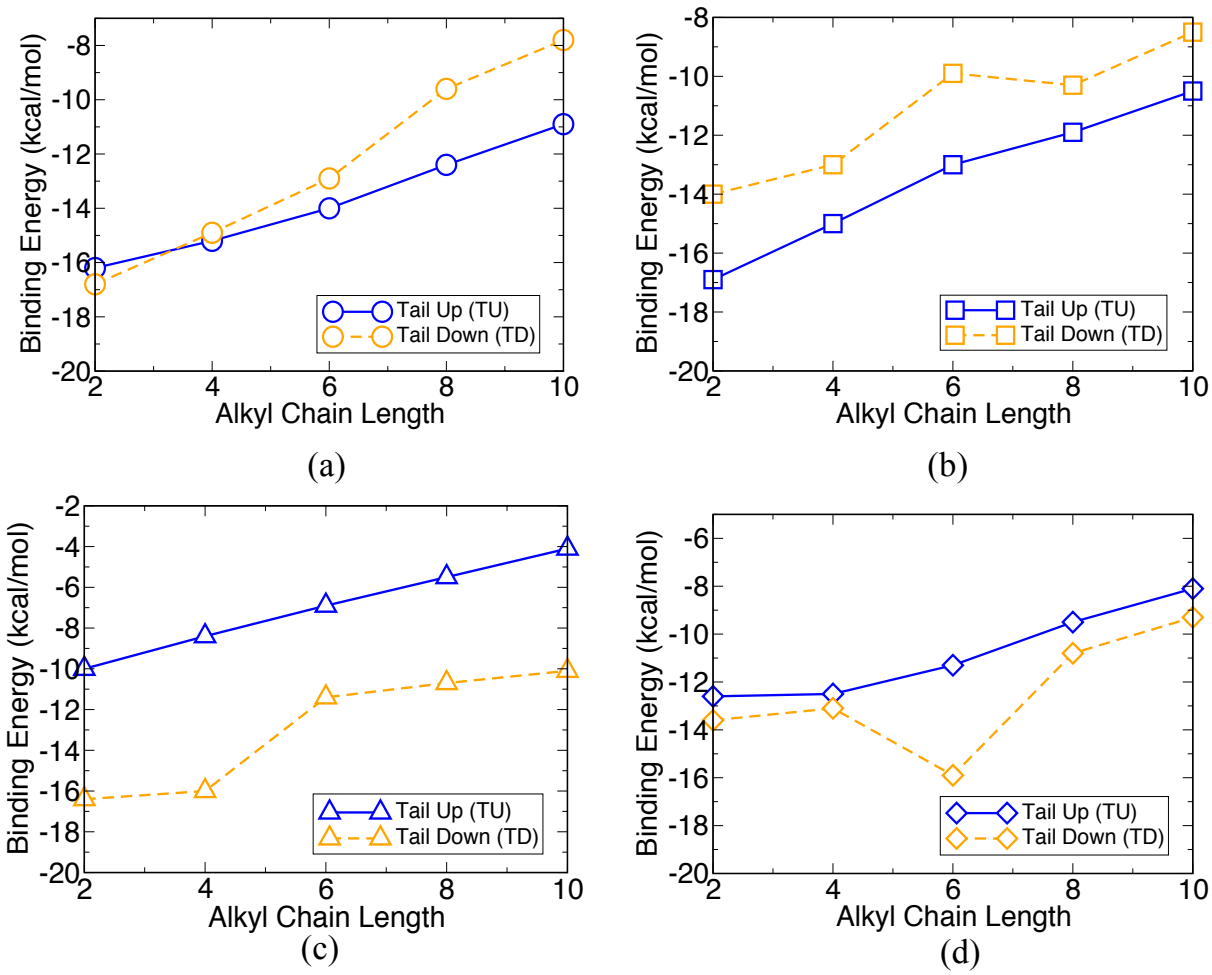


Figure 5.34: Binding energies of complexes considered (a) FeP[$C_n$mim]$^+$ (b) FeP[$C_n$py]$^+$ (c) FeP[C_nth]$^+$ (d) FeP[$C_n$pyrr]$^+$. The lines are provided only as a guide to the eye.

relative separation between the conformations is the most pronounced among all of the classes of ring containing cations considered.

The final ring containing system was the 1-methyl-1-alkylpyrrolidinium cations that lack aromaticity in the head group. It was shown to have binding energy profile similar to the thiazolium cations with the TU conformations being less stable as compared to the TD conformations, The initial conformations and their effect on the pyrrolidinium cations is shown to be minimal as compared to the other class of cations with the exception of the [C_6pyrr]$^+$ for which the TU geometries

Table 5.12: Relative binding energies (kcal/mol) of different conformations for imidazolium-based cations

Cation	Tail Up	Tail Down
$[C_2mim]^+$	0.7	0.0
$[C_4mim]^+$	0.0	0.4
$[C_6mim]^+$	0.0	1.0
$[C_8mim]^+$	0.0	2.8
$[C_{10}mim]^+$	0.0	3.1

Table 5.13: Relative binding energies (kcal/mol) of different conformations for pyridinium-based cations

Cation	Tail Up	Tail Down
$[C_2py]^+$	0.0	2.9
$[C_4py]^+$	0.0	1.9
$[C_6py]^+$	0.0	2.3
$[C_8py]^+$	0.0	1.6
$[C_{10}py]^+$	0.0	2.1

is at an energy 4.6 kcal/mol above the TD one (Table 5.15).

The effect of the aromaticity is shown in the binding profiles given for the cation. The presence of aromaticity is particularly influential in driving the relative energies between conformations with the TD ones emerging as unstable for imidazolium and pyridinium containing systems as the chain size is enhanced. On the other hand, the pyrrolidinium and thiazolium containing systems show the reverse trend in terms of stability of conformations in which the TD ones are more stable as compared to the ring facing geometries. Out of the classes considered,

Table 5.14: Relative binding energies (kcal/mol) of different conformations for thiazolium-based cations

Cation	Tail Up	Tail Down
$[C_2th]^+$	6.4	0.0
$[C_4th]^+$	7.6	0.0
$[C_6th]^+$	4.4	0.0
$[C_8th]^+$	5.2	0.0
$[C_{10}th]^+$	6.0	0.0

Table 5.15: Relative binding energies (kcal/mol) of different conformations for pyrrolidinium-based cations

Cation	Tail Up	Tail Down
$[C_2pyrr]^+$	1.0	0.0
$[C_4pyrr]^+$	0.6	0.0
$[C_6pyrr]^+$	4.6	0.0
$[C_8pyrr]^+$	1.4	0.0
$[C_{10}pyrr]^+$	1.1	0.0

the cyclic pyrrolidinium cations show the least segregation between the different conformations.

The binding affinities to the porphyrin is discussed collectively for all of the cations considered in the work. The comparative analysis of all the bound species with FeP is presented in the Figure 5.35. The two subfigures indicate the collective trends established for TU and TD conformations. From Figure 5.35(a), the relative trend for the association for the cations in complexation with FeP can be expressed as imidazolium $\approx$ pyridinium $<$ pyrrolidinium $<$ thiazolium for a given size of alkyl chain length. To explain the trends, the geometrical make up of the individual cations was analyzed and their interactions with FeP probed. The relative closeness of the energies between imidazolim and pyridinium cations can be explained because of the similar natured planar rings that advocate $\pi - \pi$ interactions (Figure A7 - A8) between the cation and the FeP. The participation of the alkyl chain is minimal as the size of the overall cation is increased.

The comparison for the binding energy profiles for the TD geometries is given in Figure 5.35(b) for all of the cations. Like TU geometries, the binding regimes obtained cannot be differentiated based on the identity of the ring composition as is participation is minimal. It is also worth noting that upon increasing the alkyl chain length of the system to octyl and decyl, the binding energies converge to a value of 2kcal/mol irrespective of the nature of the cation presented to the FeP [258]. This observation can explained on the basis of lower participation of

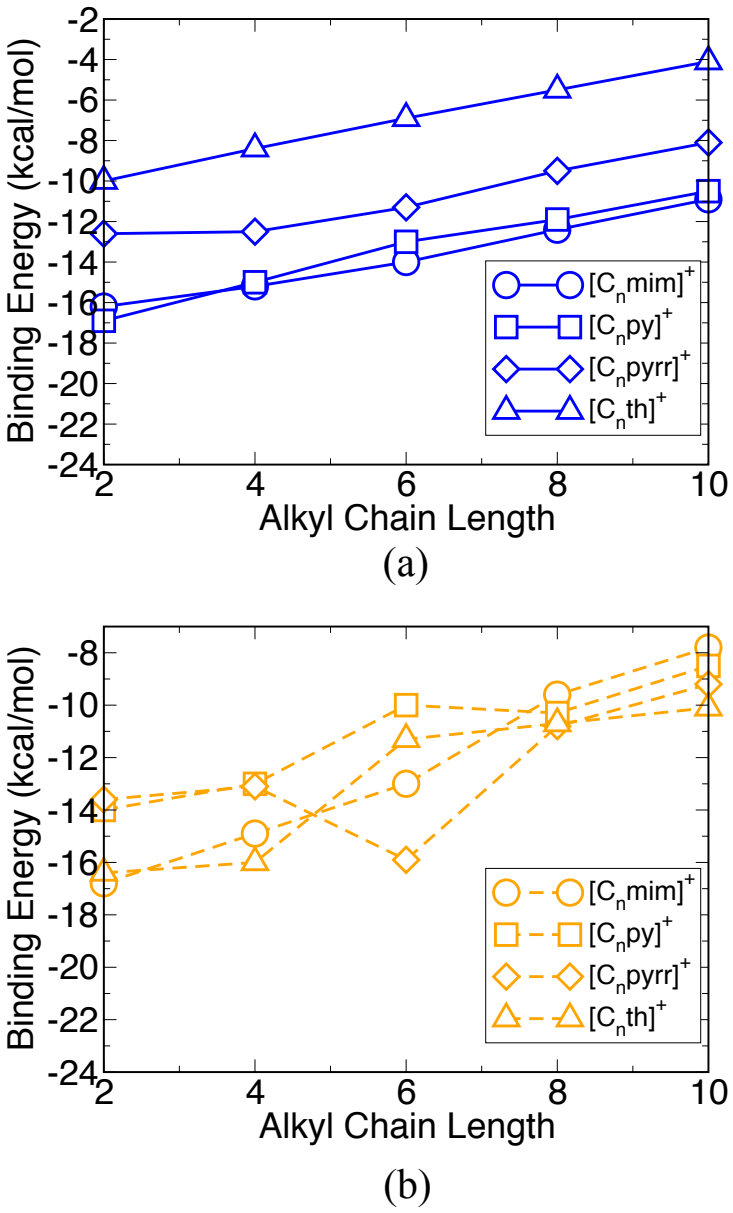


Figure 5.35: Comparison of binding energies from all complexes optimized from (a) Tail Up (b) Tail Down geometries. The solid and dotted lines serve only as a guide to the eye.

the ring entity with the underlying FeP and its minimal contribution to the overall binding energy of the complex. In case of all of the TD geometries, the chain is positioned in a parallel fashion as compared to the FeP which was also observed earlier for the B3LYP-D2 geometries of imidazolium-based cations. A collection of

the snapshots of these structures are given in the Figures A7 - A10. The reader is referred to figures 1-5 for the TU snapshots and 6-10 for the TD snapshots for each of the individual ring cationic systems.

As the complexation of FeP gives rise to different modes of interactions for different conformations, it was deemed necessary to look at the orbital analysis of the complex. To this end, the NBO analysis was performed at two separate levels. In the first one, the overall stabilization obtained from second order perturbations between the two monomers. FeP and cation, were summed and the individual structural components were observed. In the next stage, an atom specific treatment was conducted for the same analysis to populate the arising orbital interactions and their placement of these interactions in terms of the different cationic positions on the system. From a broader perspective, π-π interactions are the dominating ones for the TU conformations arising from the imidazolium and pyridinium containing systems of the TU nature, while the thiazolium and pyrrolidinium show a different electrostatic coupling. The stability of the thiazolium containing TU systems are the least on a comparative scale due to the fact in their case, the majority of the interactions are of the $\sigma - \pi$ nature giving rise to T-shaped complexes. This has also been tested by performing calculations on representative interplanar conformations of thiazolium and observing a retention in the geometries. The exposure of the thiazolium system having the sulphur atom at an upright position keeps the overall complex upright as well.

5.3.2 NBO Analysis

The first part of the NBO analysis was conducted for observing the individual contributions from the different donor groups on the cation and the porphyrin molecules. The possible donor-acceptor subsystems can be constructed from the atomic sets, namely, (a) Fe (b) Porphyrin (c) ring and (d) alkyl chain. These are used to populate the contributions and are referred to Fe, Por, Ring and Chain, respectively in the stabilization plots explained in this subsection. For dividing

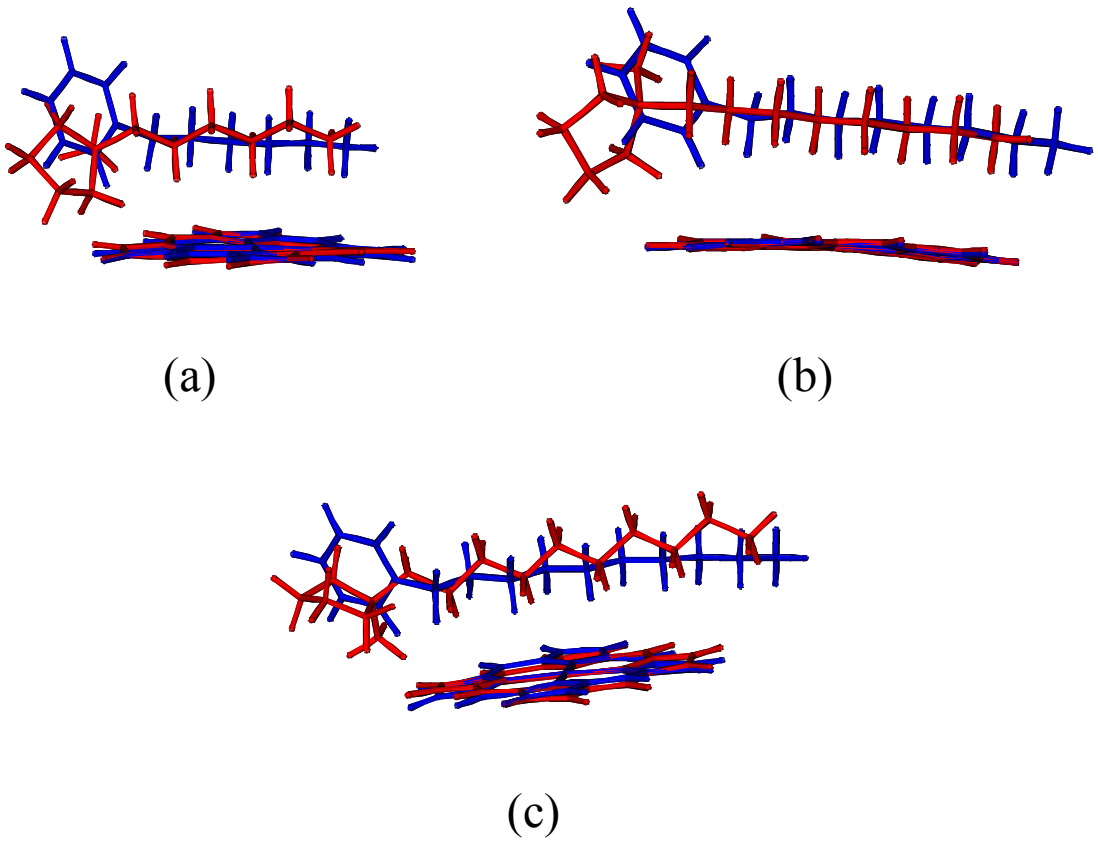


Figure 5.36: Alignment of optimized tail down complexes containing pyridinium (blue) and pyrrolidinium (red) cations with alkyl side chain lengths of (a) Six (b) Eight (c) Ten

the atoms into these groups for the cations, the atoms belonging to the ring as well as the methyl group was included in the 'Ring' subsystem. The other atoms belonging to the alkyl chain was labelled as 'Chain' subsystem. The individual donor-acceptor energies have been expressed for the TU and TD geometries and presented in the Figures 5.37 and 5.38. The first figure gives the contributions for the TU and TD from imidazolium and pyridinium containing systems while the second one expresses the same for thiazolium and pyrrolidinium systems. In addition to the sets of atoms contributing to the donor-acceptor interactions, the individual contributions of the atoms to to each other is also expressed. For this,

a heatmap consisting of the different orbital interactions possible with the cationic positions have been explained later in this subsection.

Now each of the subfigures for the cationic groups will be explained. In line with the geometry of the TU conformations, the dominant donor-acceptor stabilization is between the the Ring and Fe among the others. The magnitude of these contributions from Ring and Fe on the cation and porphyrin side is independent of the size of the cation, except for the smallest cation ($[C_2mim]^+$). The stabilization due to the orbital interactions arise from the geometrical placement of the structural components for imidazolium cations (Figure 5.37(b)). In case of the TD conformations, the Chain and Fe components within the whole system contribute more towards the stabilization. The strongest interaction for both the imidazolium and pyridinium systems are reflective of the TD geometries in which the chain interactions are dominant. Some of the TD geometries, in which the the chain is shown to at least interact with one of the pyrrole nitrogens in FeP.

Figure 5.37(c-d) shows the behavior of pyridinium containing systems where a clear distinction is observed in terms of the individual elements contributing to the overall stability of the complex. For the TU conformations, it can be easily observed that the cationic contribution is dominated by the Ring elements while the contrary can be said for the TD conformations. From the point of view of the overall donating ability, the Ring-Fe and Ring-Por contributions are shown to be comparable to each as the chain length on the cation is increased(Figure 5.37(c)). Out of the possible interactions between the FeP and IL cation, these modes are most dominant in asserting the non-Lewis energetic stabilization to the overall wavefunction. In case of the tail facing TD conformations, the Fe and Por are the most predominant donor groups. The ring interactions with the FeP are strong for the homologous series upto $[C_6mim]^+$ while extending the chain makes the chain contributions to be more dominant. The stabilization due to Fe-Ring diminishes to almost negligible values for these two cations. The contribution arising from the chain, namely, Por-Chain also rises with the increase in the chain length. This second order perturbation becomes the driving factor for the stability of the

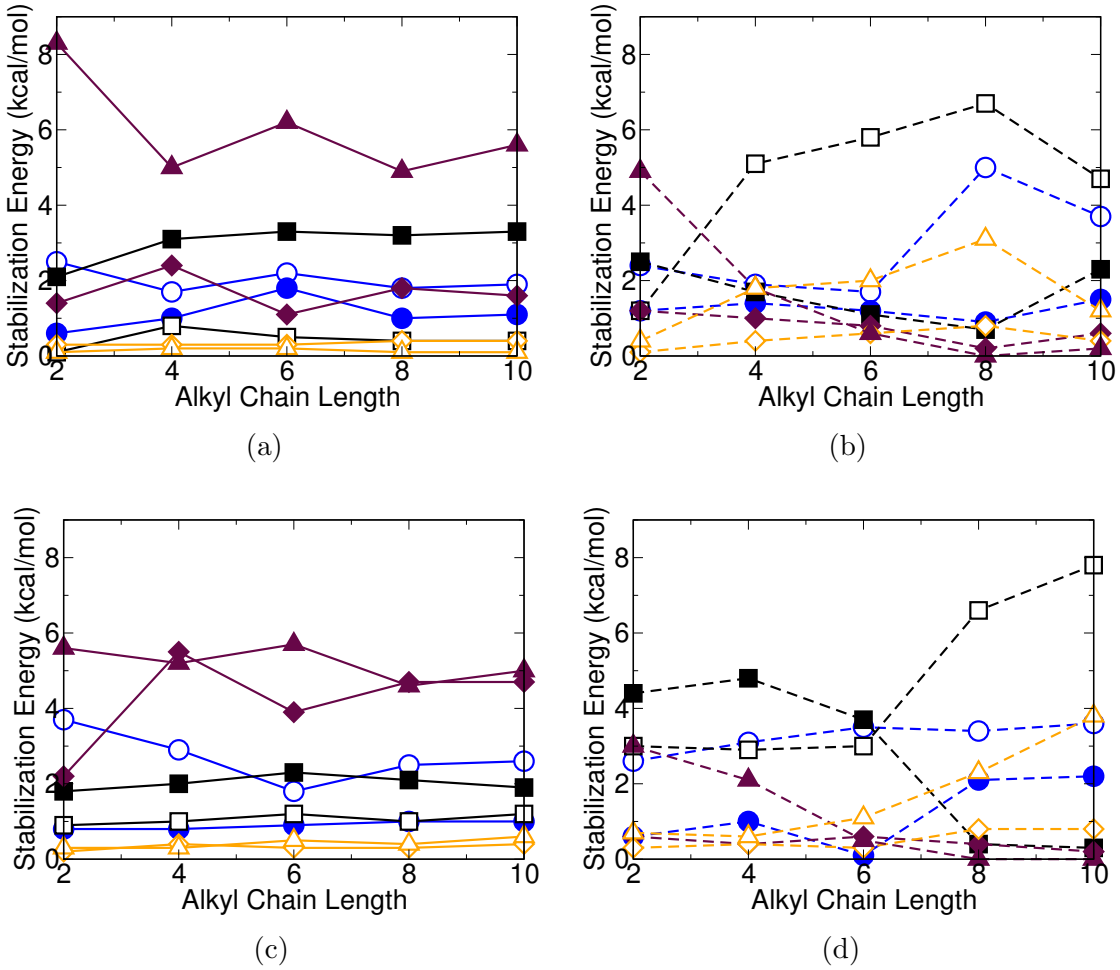


Figure 5.37: Second order perturbation energy (E_{ij}^2) for different donor-acceptor pairs for imidazolium (a) TU (b) TD and pyridinium (c) TU (d) TD complexes. (Legends: Por-Ring ●, Por-Chain ○ , Fe-Ring ■ , Fe-Chain □ , Ring-Por ♦ , Chain-Por ◇ , Ring-Fe ▲ , Chain-Fe △) Lines are include only as a guide to the eye.

complex and reflects the optimized geometries for the complexes(Figure A8).

The same analysis was also conducted for the thiazolium and pyrrolidinium containing cations to analyze the effect of the geometries exposed to underlying FeP. These donor-acceptor contributions have been laid out in Figures5.38 (a-d). First, we will discuss the implications of the placement of those interactions for thiazolium containing systems that will be followed by the pyrrolidinium containing systems. In case of the thiazolium TU geometries, the chain interactions are absent perfectly correlating with the optimized structures. In each of these structures the chain points away from the Fe center and the conclusion drawn from the perturbative analysis is expected. On the other hand, for TD conformations, in

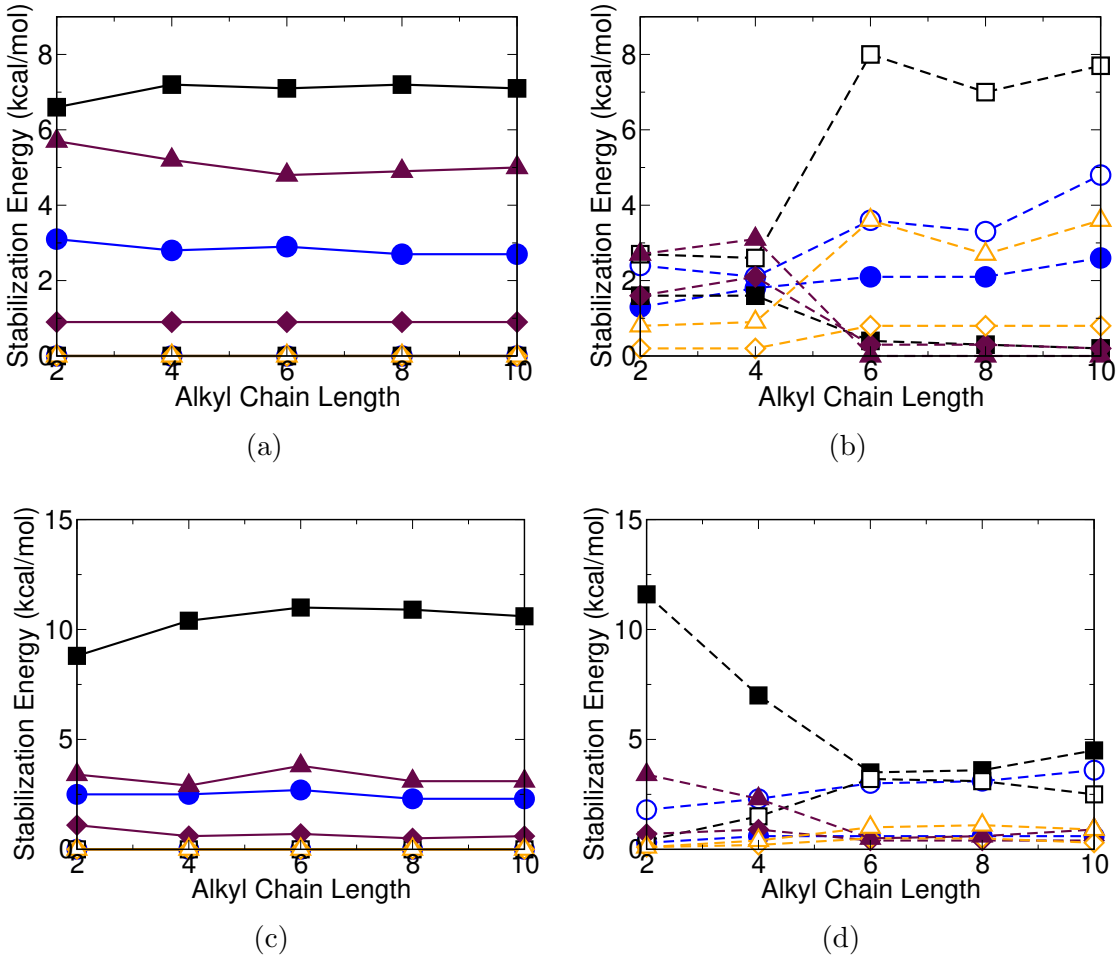


Figure 5.38: Second order perturbation energy (E_{ij}^2) for different donor-acceptor pairs for thiazolium (a) TU (b) TD and pyrrolidinium (c) TU (d) TD complexes. (Legends: Por-Ring ●, Por-Chain ○ , Fe-Ring ■ , Fe-Chain □ , Ring-Por ♦ , Chain-Por ◇ , Ring-Fe ▲ , Chain-Fe △). Lines serve as a guide to the eye.

addition to the contributions from the ring, the Chain-Fe and Chain-Por energies are non-zero adding to the stabilization of the overall complex. This result might be tied to the cause of the thiazolium-based TD geometries being more stable as compared to the TU ones. In the case of the other two aromatic systems, the converse being true, a closer inspection reveals that for a given donor-acceptor interaction are similar for $[C_2th]^+$ and $[C_4th]^+$ cations. Upon adding to the alkyl chain length, the Chain containing contributions increase in magnitude, as shown by $[C_6th]^+$ to $[C_{10}th]^+$. Similar to the imidazolium and pyridinium containing systems, the enhanced participation of the chain is attributable to the geometry obtained for the TD conformation considering it as the initial structure.

Finally for the non-aromatic pyrrolidinium cations, the Figures 5.38(c) and (d) provide the donor-acceptor stabilization energy decomposition with respect to the alkyl chain length of the cation in complex. In case of the TU complexes, only the contributions from the ring entities, i.e. the Ring-Fe , Fe-Ring and Por-Ring are the dominant while the others are zero. On the other hand, for TD geometries, the contributions from the Ring entities deiminish as the chain length is increased. While in the same vein, the contribution from Fe-chain becomes significant with the increase in the size of the cation.

The primary interaction for the TU conformations for imidazolium cations is through the two-center bonded interactions involving the bonding orbitals of the cation and antibonding lone-pair orbitals on FeP - BD-LP* (Figure A20). The donor-acceptor interactions are focused on the ring atoms. For example, CR2 and CR3 atoms in $[C_2mim]^+$ show an interaction ~ 3 kcal/mol, which is greater than the ones evaluated for all the other cationic sites. The bond-centric interactions concentrated on CR1 and N2 appear to be the most stabilizing for longer alkyl chain imidazolium cations. The bond-centric interactions are spread across the entire ring for the TD conformation(Figure A21) of $[C_2mim]^+$, while CR1 and CR3 on the ring participate in such interactions in $[C_4mim]^+$. However, from $[C_6mim]^+$, the contributions from the alkyl chain atoms, in particular, CA1, CA2, and CA3 become noticeable with nominal contributions from the ring atoms. The ring being the primary site for interactions for TU conformations and the chain for TD conformations can be directly correlated to the respective optimized geometries.

Similar to imidazolium cation, it seems that almost all of the interactions are skewed towards bond centric modes (BD-LP*) in the case of the pyridinium-FeP complexes showing that atoms prefer to interact with FeP by sharing their electronic density with neighboring members. Pyridinium containing systems in TU geometries (Figure A22) show placement of bond centric (BD-LP*) type inter-

actions for various atoms within the ring. Apart from the complex containing $[C_6py]^+$, the ring carbon CR5 is shown to have a more dominant role in stabilization as compared to other members. Similar to imidazolium containing systems, none of the alkyl chain carbon atoms show any significant interactions which can be tied to the fact that the chain is positioned away from FeP. In the case of TD complexes (Figure A23), the significant stabilizing interactions occur on the ring and alkyl chain atoms for the first two cations in the homologous series $[C_2py]^+$ and $[C_4py]^+$. However, progressive displacement of the ring from the Fe center causes the focal point of stabilization interactions shift towards chain atoms when octyl and decyl chain containing cations are considered. Both of these complexes show concentrations at sites CA3 and onwards with $[C_{10}py]^+$ exhibiting notable interaction at even the CA5 position. These TD complexes are rather energetically unstable as compared to the TU ones and the chain cation sites are shown to be directly exposed to the Fe center in the optimized geometries.

For thiazolium complexes, the TU conformations (Figure A24) show interactions concentrated at the sulphur atom (SR). In the optimized geometries, this sulphur atom points directly towards the porphyrin with the rest of the cation attached to it away from the receptor. For these complexes, the LP orbital of sulphur is shown to be interacting with the antibonding orbital of Fe (LP-LP*) irrespective of the cation size. These lone pair interactions serve as the major factor in the overall stabilization in the donor-acceptor framework with all the energy values in the TU heatmaps for SR atom in LP-LP* interactions to be more than 3.5 kcal/mol. Particularly, this mode exceeds the contributions from CR-LP* interactions and notably, BD-LP* mode appears to be inconsequential unlike that in imidazolium and pyridinium complexes. This is expected due to the fact that in the case of thiazolium containing TU complexes, the electronegative sulphur points towards Fe resulting in the LP-type interactions. TD optimized complexes (Figure A23), on the other hand, do not show any energetic contributions from sulphur that are more than 0.5 kcal/mol. For the smaller ethyl and butyl thiazolium cations, the energetic contributions are scattered across the ring and chain sites with highest

weightage to BD-LP* type interactions. Also, for TD complexes at hexyl or above, the contribution of sulphur to the stabilization is minimal and the alkyl carbon atoms (CA3, CA4 and CA5) prefer to interact through the bond-centric mode (BD-LP*). The relatively lower magnitude of stabilizing interactions in the case of TD complexes highlights the crucial role of FeP as donor rather than the cation.

For the non-aromatic system under consideration, i.e. pyrrolidinium (Figures A26 and A27), the pattern of stabilization follows suit with the other aromatic systems. In TU systems, the significant interactions are consistent with the geometry obtained from optimization and all of the contributions are shown by members of the cyclic ring. Among these, except for the $[C_2pyr]^+$, a consistent pattern consisting of the CR2, CR4 and CM1 atoms emerges (Figure A26) with highest perturbation energies at about 2.5 kcal/mol. Upon investigating TD complexes, the energies become more significant for tail facing atoms as the alkyl chain attached to the ring is enhanced. Although, the sites showing the stabilization change, the mode of interaction for the atoms in the case of these conformations still remain the same (BD-LP*) which points to the preference of alkyl chain atoms to interact with the lone pair antibonding orbitals of iron in the optimized complex. Thus, it might be concluded that the carbon and nitrogen atoms belonging to the ring and chain prefer to interact in a bond-centric scheme as compared to others while sulphur prefers to engage in lone pair type interactions. Upon carefully investigating the binding of thiazolium TU complexes, these lone pair interactions could be responsible for keeping the optimized geometries to be upright while ring heads for other planar aromatic members converge to favorable $\pi - \pi$ interactions.
For iron porphyrin as the donor molecule, similar analysis as described in the former section was performed. All classes of cation-FeP complex for both TU and TD conformations were considered and cationic sites were populated according to their mode of interaction with the FeP molecule. Various stabilizing interactions and their magnitudes for imidazolium containing complexes have been shown in Figures A28 - A29. For TU complexes (Figure A28), the stabilization seems to be distributed among the ring and alkyl sites throughout the homologous se-

ries. Although, the dominant interactions donating to the Rydberg antibonding orbital atoms of the cation (LP*-RY*) are concentrated towards the ring atoms CR1-CR3. In the case of FeP as donor, the intensity of these interactions are comparably lower than for cation as electron density transferring agent. The TD complexes on the other hand show pronounced spots of interaction for their antibonding bond centric orbitals from the LP* orbital of the underlying Fe atom for alkyl chain carbon atoms. Except $[C_2mim]^+$ system, for all other TD complexes, the dominant interactions are shown for CA1-CA3 carbons with the magnitude for CA1 of $[C_8mim]^+$ reaching up to as high as 4 kcal/mol. This is in line with our discussion of overall stabilization for TD geometries in which the contribution from the FeP side overshoots the cationic contribution significantly for all systems except $[C_2mim]^+$.

For pyridinium containing TU complexes (Figure A30), most of the interactions follow suit with those in the imidazolium cations. Stabilizing interactions are skewed towards atoms in the ring with just the exception of CA1. Among all possible modes, three types of interactions appear to be significant: BD-LP*, LP*-BD* and LP*-RY*. In addition to the LP*-RY* mode observed for the $[C_nmim]^+$ systems, even bond centric interactions are shown to be possible in case of pyridinium systems for TU geometries. The contribution of FeP to overall stabilization for TD geometries is greater than the cation. This is shown in Figure A31 where multiple atoms belonging to the cation get stabilized by porphyrin. In terms of conformation effects, both ring and alkyl chain atoms are shown to participate in donor-acceptor perturbations but as the size of the associated $[C_npy]^+$ cation is increased, the collective effect on the alkyl atoms (CA) dominates over the ring atoms. This is caused by donation from the LP* orbital from the Fe atom towards bond centric acceptor orbitals in a LP*-BD* framework. This LP*-BD* mode seems to be the most important one among the possible interacting modes for TD $C_npy]^+$FeP geometries.

Owing to the geometrical similarity of all the thiazolium TU complexes after geom-

etry optimization, a consistent pattern of interaction emerges for these geometries (Figure A32). The most dominant accepting atom is the SR atom which is closest to the porphyrin molecule. It is shown to be favorably interacting in three different modes, of which the RY* accepting antibonding orbitals compose the maximum weightage. Significant contributions are also shown by the ring atoms CR1 and CR3, both of which report a BD* mode at the acceptor end in combination with the sulphur atom. Figure A33 shows that interactions in the case of TD conformations are distributed among ring and alkyl chain atoms. As the size of $[C_nth]^+$ cation is increased from 4-6 and higher, the chain atoms CA2-CA5 are shown to participate in stabilization in a bond centric framework (LP*-BD*).

In pyrrolidinium TU complexes (Figure A34), the mode of interaction most favorable is the LP*-BD* which refers to a bond centric antibonding orbital at the acceptor side. This mode is shown to be shared mostly by atoms CR2, CR4 and CM1 among others present in the cation. Only exception to this is shown by $[C_2pyrr]^+$ in which CR1 is shown to be active. The most dominant interaction is reported for the CR2 atom in the cyclic ring that is closest to the Fe center in the optimized geometries. For TD geometries, the distribution is interspersed among ring and alkyl chain atoms on the pyrrolidinium cations. in their case, CR2, CM1 and CA2 atoms are shown to be most conducive to participation during the stabilization. These atoms also prefer to interact in the LP*-BD* mode irrespective of the conformation which seems to be the most important mode for TD complexes for all classes of cations.

5.3.3 Electrophilicity Index

Electron affinity indexes calculated from the full population analysis of the optimized cation-FeP complexes are provided in Figure 5.39. The electrophilicity index for FeP in the absence of any cation is 48 kcal/mol. Figures 5.39 clearly demonstrates that the presence of cations considered in this work are capable of significantly enhancing the electrophilic nature of FeP as indicated by the increase

in the electrophilicity index by almost a factor of three from the corresponding gas phase value. The elevation in the electrophilicity index for TU conformations is a function of the chemistry of the cation headgroup presented to FeP; however, the dependence on the alkyl chain length is not as strong. For the imidazolium and pyridinium containing systems, the electrophilicity index varies between 145-150 kcal/mol. The electron uptake ability of FeP is somewhat reduced when pyrrolidinium or thiazolium cations bind to FeP. It appears that the electrophilicity index obtained for TU conformations are correlated to the corresponding binding energies (Figure 5.35(a)), the lower the binding energies, the less of an electrophile is FeP. When TD conformations are considered, electrophilciity indexes cannot be as easily discriminated as those observed for TU conformations (except for those containing hexyl chain), suggesting that the identity of the headgroup in cations plays only a minor role in modulating the electrophilic properties of FeP. This can be rationalized on the basis that alkyl chain is exposed to FeP in TD conformations, whereas the ring is a dominant interaction site in the TU conformations. Beyond the hexyl chain, the electron affinity of the FeP molecule reduces and levels off for longer alkyl chains. Interestingly, this behavior also correlates with the binding energies that are within 2 kcal/mol for all the octyl and decyl containing cations in TD conformations.

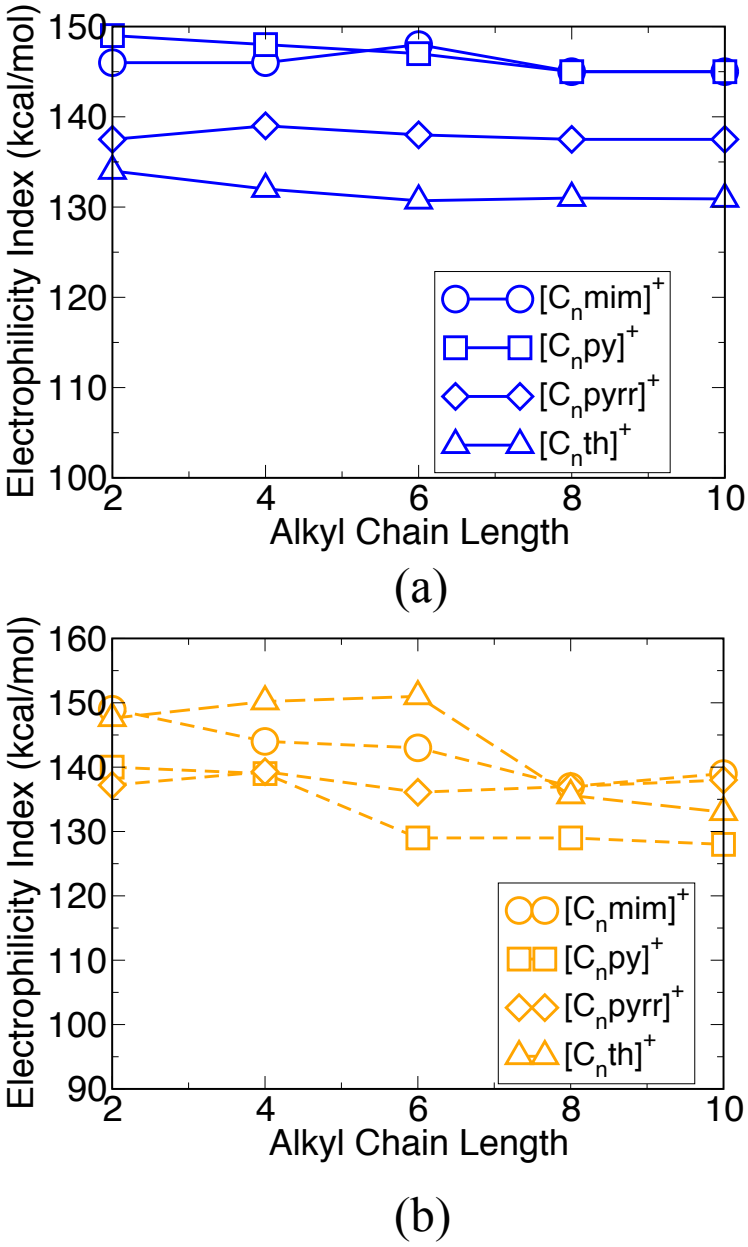


Figure 5.39: Electrophilicity indexes of FeP in optimized complexes from initial (a) Tail Up and (b) Tail Down position. Lines are included only as a guide to the eye.

5.4 Cluster Model 1: Straight Cationic Systems

In the next part of the work, ionic liquid cations having no ring or cyclic moiety were studied to observe their binding affinity in presence of FeP. For this,

the same computational setup as the previous cations was utilised. The three cations considered were namely, 1-trimethyl,1-N-alkylammonium, 1-trimethyl,1-N-alkylphosphonium and 1-dimethyl,1-N-alkylsulphonium respectively. The alkyl chain length of the systems were increased similar to the imidazolium-based cations and other systems by enhancing the carbon atoms at the chain from ethyl-decyl progressively.

Figure 5.40(a-c) give the binding profile of the straight ionic liquid cations as a function of their alkyl chain length. The geometries associated with these complexes are also give in Figures A11 - A13. The TU geometries for ammonium and phosphonium cations show the same behavior as the aromatic systems with the chain pointing away from FeP while the central nitrogen heteroatom is held by the crowding of the three attached alkyl groups pointing towards porphyrin. A geometrical comparison with pyrrolidinium containing cations provide the insight that in the presence of three methyl groups facing the porphyrin, the overall cation cannot place itself in a parallel position and the chain cannot reposition itself. On the other hand, sulphur being divalent is attached to three alkyl groups, out of which two are methyl while the third is varied across the homologous series. As a results, sulphur containing complexes are shown to access a conformation in which the chain attains a parallel mode with respect to porphyrin. it can be concluded that geometrical freedom provided at the point of association for the cation might well dictate the path of optimization as well as the final conformation that is obtained.

On the other hand the TD complexes show a very familiar behavior in which the chain seems to slither and the head of the cation is progressively pushed away from the center of the FeP. This observation holds true for all of the cations involved in the complexation with no exceptions shown from any of the groups. The porphyrin responds to the increase in the alkyl chain length by adopting a non-planar domed mode in which the unit vector of the porphyrin points upwards while the rest of the macromolecule buckles. A brief discussion of the non-planar

modes has been provided in the section dedicated to the metal porphyrins where the central Fe was substituted with other elements in the same row.

The binding energetics from the three cations show a very interesting behavior and only one of the cation families reflects the inherent conformational bias. Ammonium based cations bind with an energy of -14 to -8.7 kcal/mol for the TU complexes and -15.4 to -12.2 kcal/mol for the TD complexes. More importantly, the C_4, C_6 and C_8 are the ones that show relative change in the energy in the binding energies comparable to chemical accuracy. The TU conformations are unstable as compared to the TD ones similar to the former thiazolium and pyrrolidinium systems. Phosphonium based cations do not show any preference for binding and the maximum relative energy between the conformation is a mere 0.5 kcal/mol which shows that the presence of phosphonium in the central group makes the conformations equally stable. Finally, the sulphonium containing systems also show a similar behavior as phosphonium only with the exception of the C_2 containing cation. It is also worth noting that the effect of the presence of sulphur is not reflected in case of the sulphonium cations in the binding poses obtained from optimization, whereas in the case of thiazolium, the geometries reflected a drastic difference in the binding energies. Upon comparing the geometries of sulphonium containing systems, the only stark difference observed in the groups approaching the underlying porphyrin. In the case of TU geometries it is one of the methyl groups attached to the sulphur while in the case of TD ones, the sulphur is exposed to the underlying porphyrin. None of the two conformations allow for the contact of sulphur with the pyrrole nitrogen or the Fe atom.

In order to understand the dependence of the binding mechanism on the ring entity provided to the FeP, a combined binding of a variety of cations was populated and their binding energies evaluated as in the case of the former systems. All of these systems were constructed with a decyl chain attached to the ring head and optimized from a tail down position. These cations have been shown in Figure 5.41(a) and their associated binding energies bearing the index of the

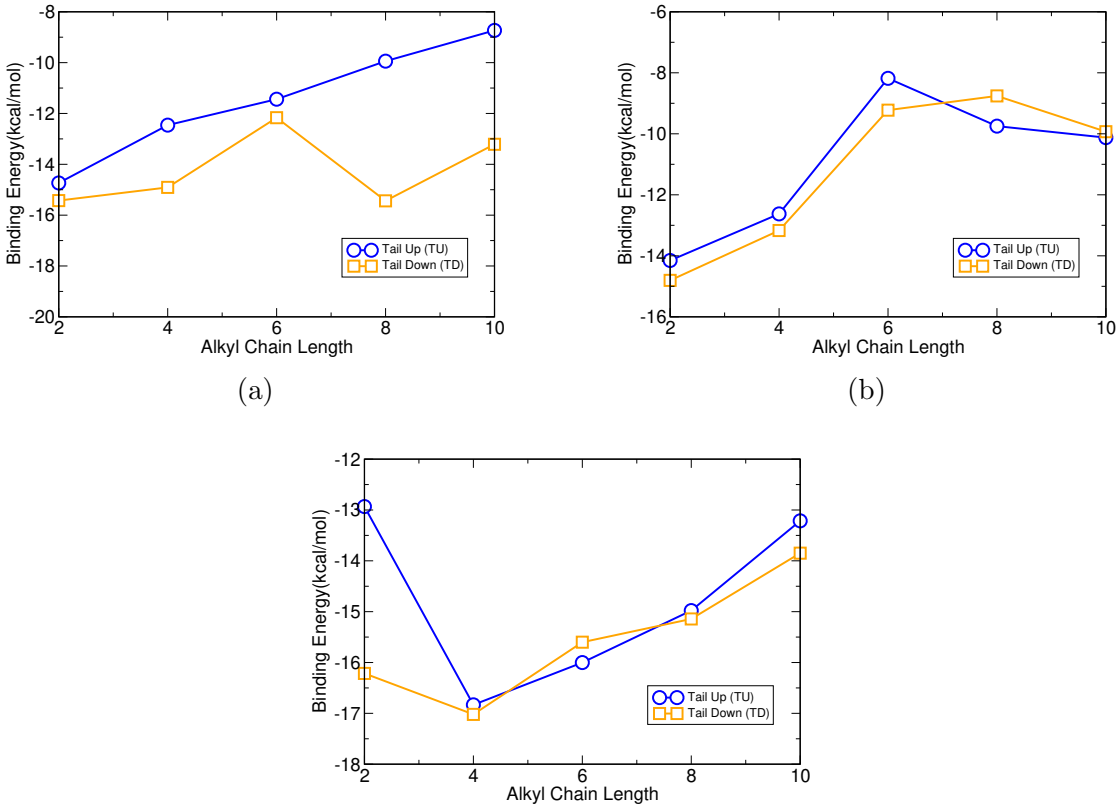
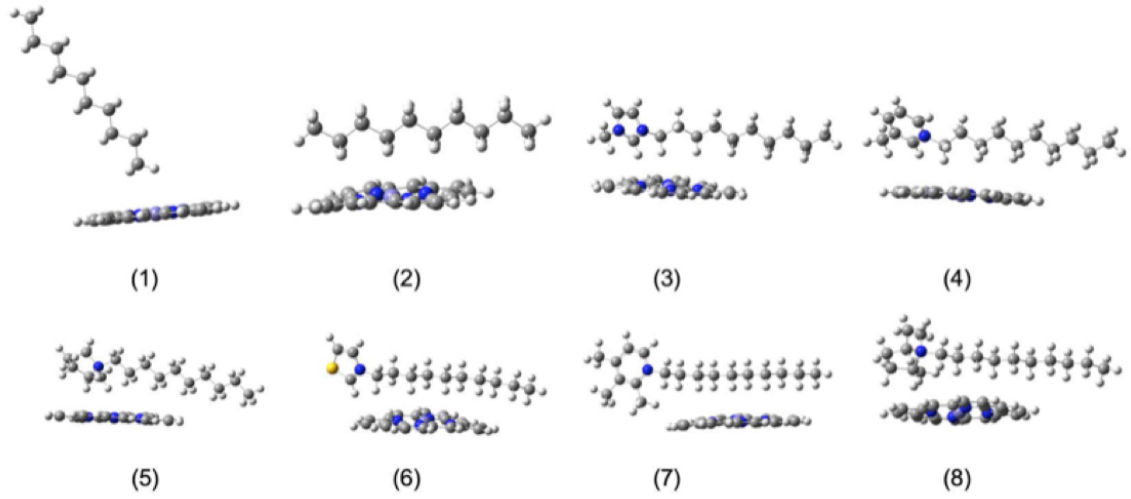


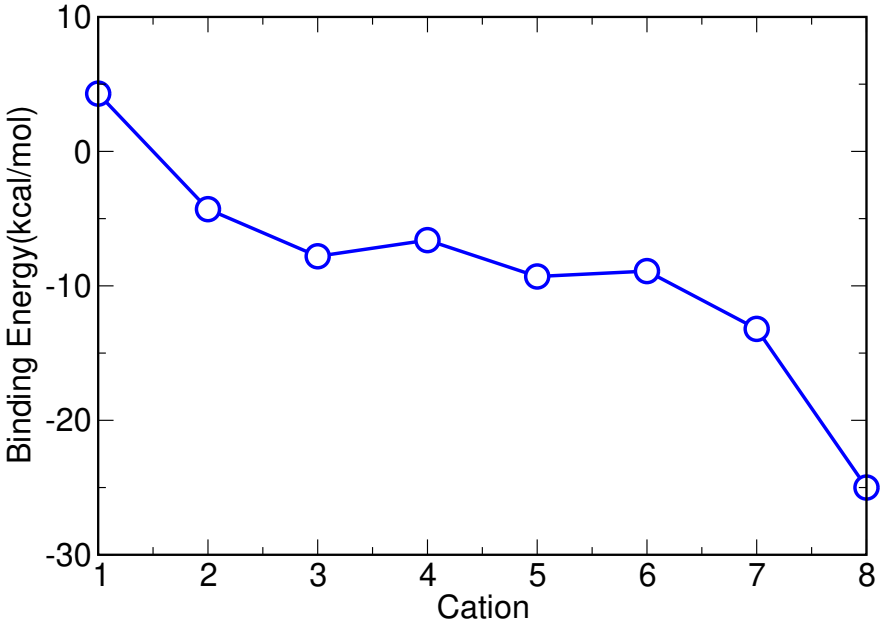
Figure 5.40: Binding energy profiles for (a) Ammonium (b) Phosphonium and (c) Sulphonium cations in complex with FeP

substrate are given in Figure 5.41(b). Very interesting trends emerge from the analysis where an almost sigmoidal behavior of the binding energy curve is visible. From the left, the most unfavorable binding occurs for a geometry in which the alkyl chain remains upright followed by favorability gained for a geometry having the C_{10} chain slithered as compared to FeP. From here onwards, moving to the right, the positions 3-6 are populated by the systems that we dealt within the section dedicated to ring cationic systems. As explained earlier, the substrates ranked 3-6 all converge within the energetic window of 2 kcal/mol. Upon moving towards the right and adding to the bulkiness of the cationic group, the binding occurs at a very favorable level with the rank 8 substrate yielding a binding energy of about -25 kcal/mol. Thus, from this brief analysis, it is concluded that although the size of the cation is increased to decyl and the tail down position is imposed on the porphyrin, the effect of the ring still remains pronounced and influences

the binding energy profile significantly. Exposure of bulky ring entities to the FeP will cause in strong association between the two monomers, which might not be favorable in introducing the dioxygen moiety into the system in the subsequent steps in the catalytic cycle.



(a)



(b)

Figure 5.41: (a) Cation molecules having various ring entities with C_{10} chain length optimized from a tail down position and (b) associated binding energies with FeP

5.5 DFT Modeling of the Catalytic Cycle

5.5.1 Substrate Free

The free energy calculation procedure and methods have been discussed formerly in the section dedicated to the methodology employed in the work. Before the association of the cations were studied, the substrate free Gibbs free energy were evaluated for each of the steps in the gas and embedded phase. Figure 5.42 gives the free energy profile of the gas phase calculations. The reader is also referred to snapshots of the representative species of the P450 catalytic cycle along with it. The changes in the free energies were added in a cumulative manner to reflect the overall change. It is clearly observed that the protonation steps in both the gas and embedded phase (Figure 5.43) are the ones that drive the total energy to a highly favorable state. Among the free energies evaluated for each step, the changes in $S_3 - S_4$ and $S_4 - S_5$ seem to be thermodynamically unfavorable for the geometries in gas phase. Experimental evidence shows that the oxygen addition step is kinetically driven and that the second reduction step is the ate limiting one in case of the P450 catalytic cycle.

However, the same quantities evaluated for the embedded geometries show that the oxygen association is unfavorable while the second reduction step is favorable in this case. This is the key difference in the treatment of the catalytic cycle from both of these environments. On an overall scale, the free energy obtained from the gas phase is almost 150 kcal/mol more favorable as compared to the embedded phase. This will also be reflected in our calculations including the ionic liquid cation substrate in the presence of the cysteinated porphyrin and related chemical species for cytochrome P450.

5.5.2 Binding Step (S_1-S_2)

All of the geometries optimized in the gas phase for the binding step have been given in A36. For the geometries optimized from an initial TU position(1-5), the

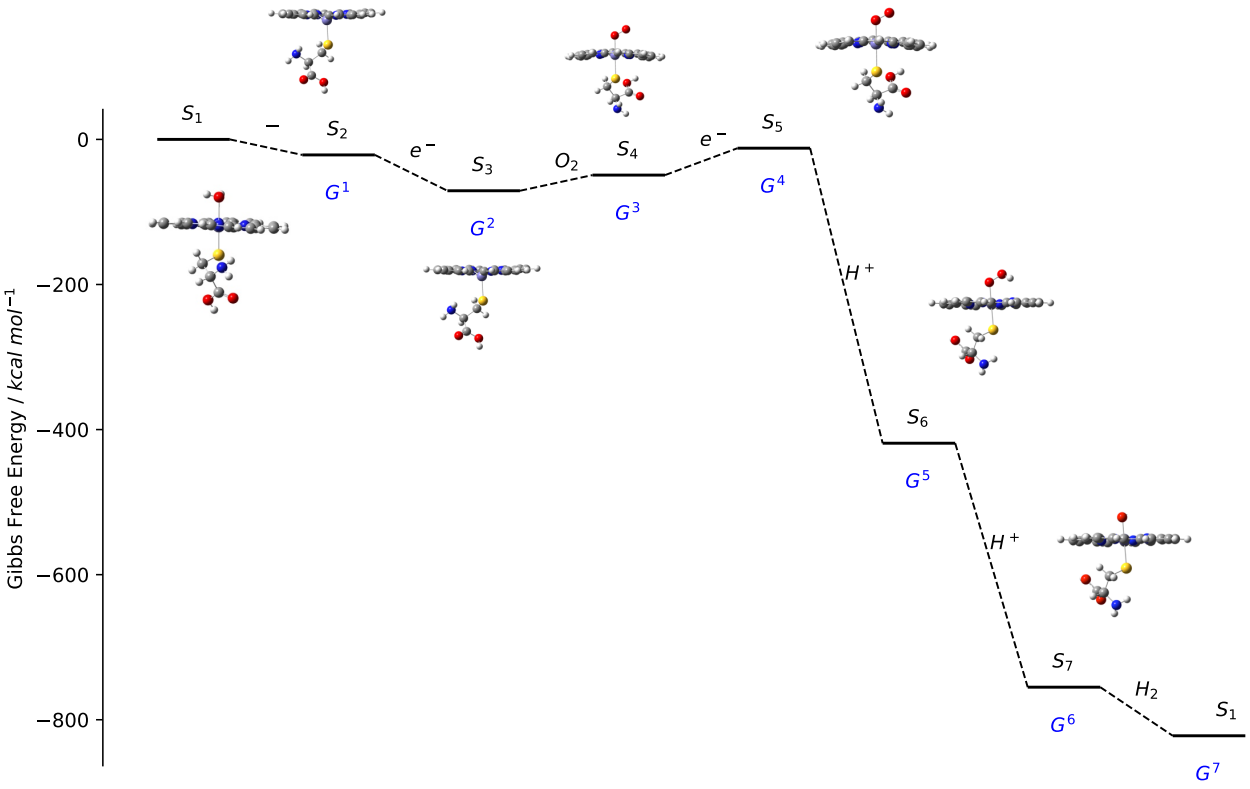


Figure 5.42: Electrophilicity indexes of FeP in optimized complexes from initial (a) Tail Up and (b) Tail Down position. Lines are included only as a guide to the eye.

cations assume a conformation in which the positively charged imidazolium ring lies parallel to the underlying porphyrin molecule. This observation holds true for all of the members of the homologous series. The TD geometries (6-10) minimize to conformations in which the tail assumes a parallel arrangement with respect to the porphyrin while the ring progressively deviates from the central position of the porphyrin. Such an observation was also noted for $[C_n\text{mim}]^+$ cations bound to FeP which was considered in our previous work [259]. On comparing the TD geometries from both calculations, it is noted that for the chain length sizes 6-10, for hexyl containing systems, the ring does not remain parallel to porphyrin as for-

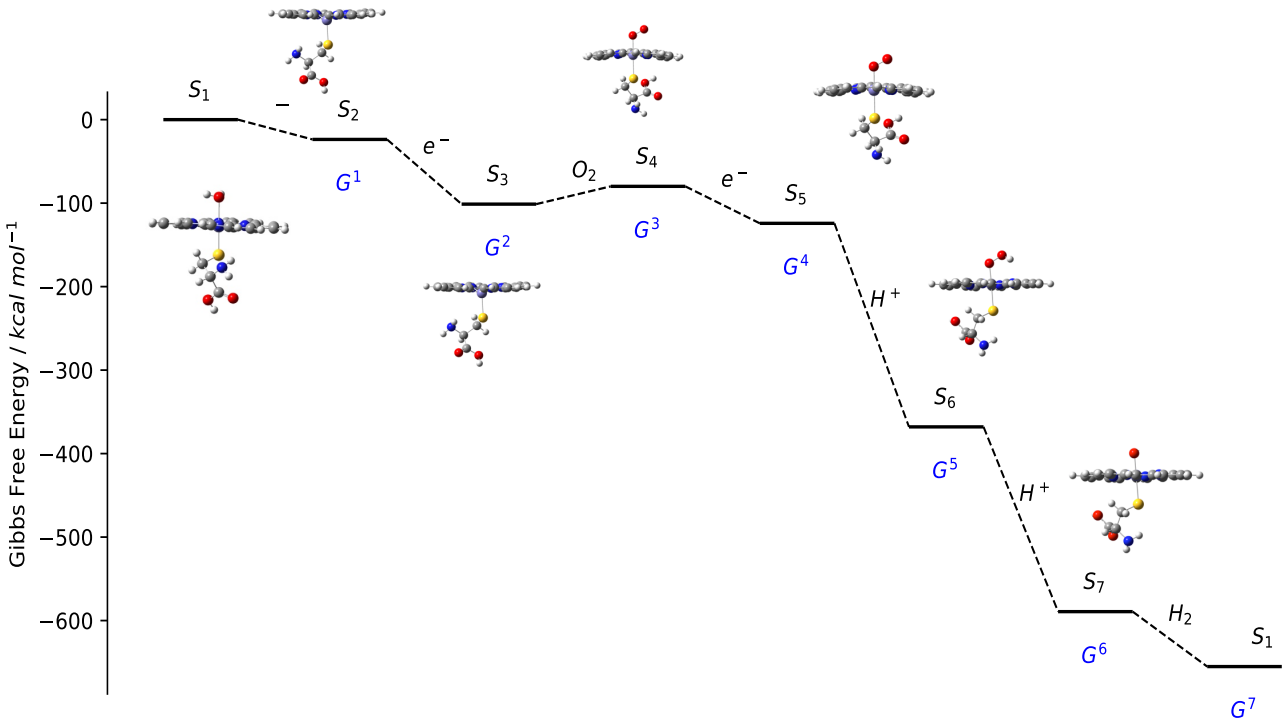
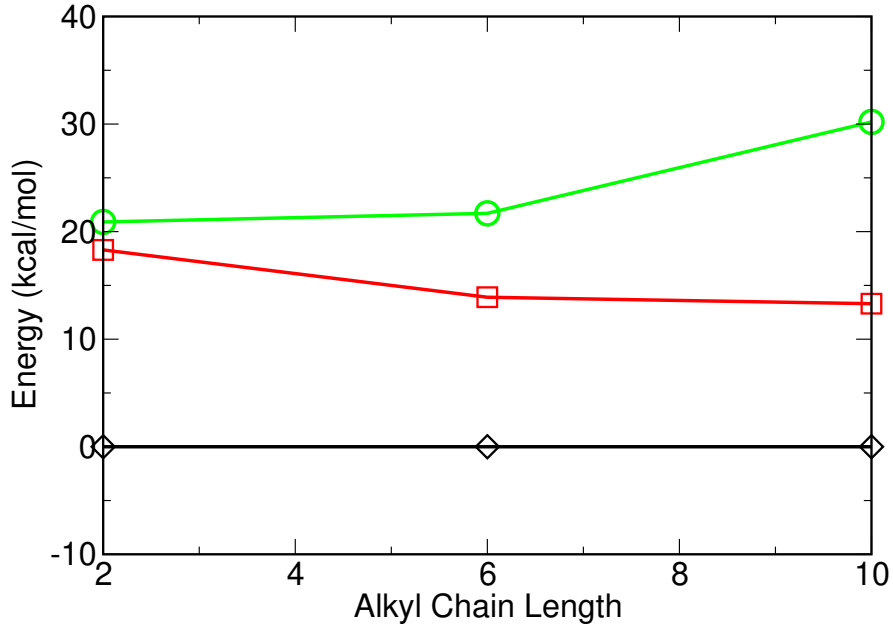
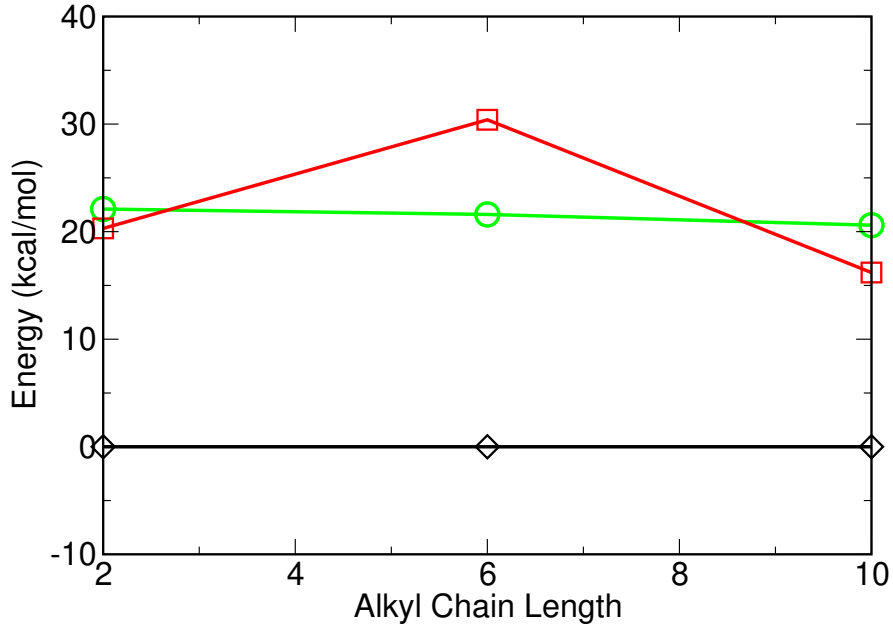


Figure 5.43: Electrophilicity indexes of FeP in optimized complexes from initial (a) Tail Up and (b) Tail Down position. Lines are included only as a guide to the eye.

merly with FeP. Also, for the decyl TD system, the ring is significantly displaced as compared to the $[C_{10}mim]^{+}$FeP complex which is reflected in the interaction energies computed for the complex. It must also be noted that the spin state (Figure 5.44) considered here is sextet which was determined by comparing the relative energies of the possible spins for the S_2 complex.



(a)



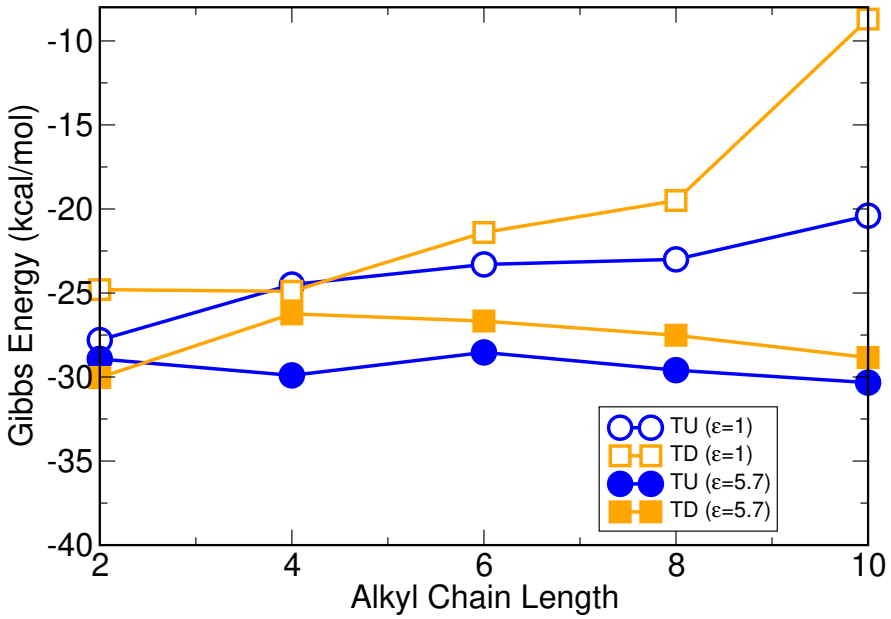
(b)

Figure 5.44: Relative energies of different spin states for S_2 (a) Tail Up and (b) Tail Down complexes optimized in ε = 1; Legend : Doublet ○ , Quartet □, Sextet ◇

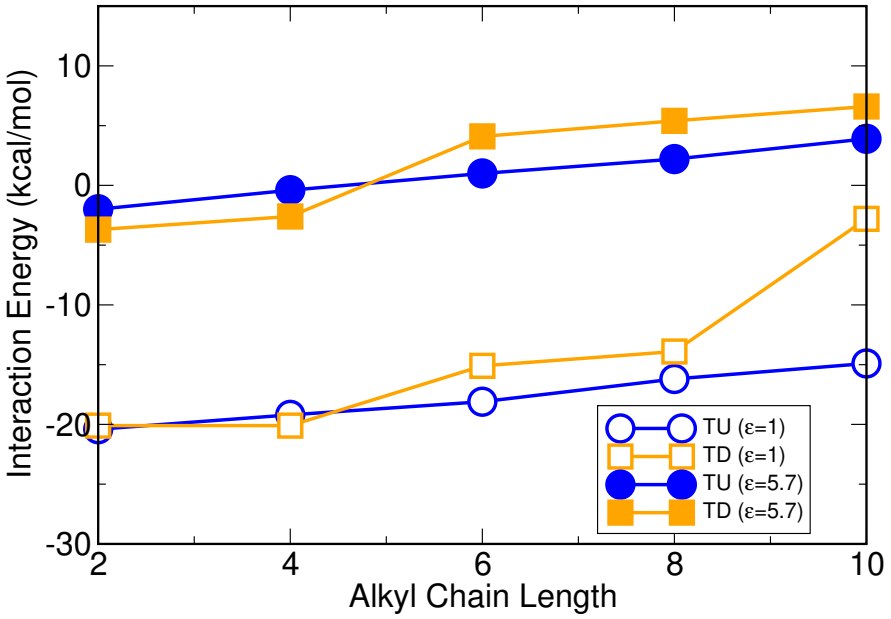
The Gibbs free energies of formation of complexes formed by association of $[C_n mim]^+$ cations to the substituted iron porphyrin molecule (FePCys) have been given in Figure 5.45(a). Upon observing the trend, it is clear that for both sets of conformations, the free energies become gradually less favorable as the alkyl chain

length is increased. Upon deeper inspection, it is also realized that except for the butyl chain containing system, all others show a conformational dependence that is quantum mechanically significant. For ethyl and chain sizes higher than butyl, the TD conformations yield a free energy that 2kcal/mol or higher than the TU ones. To probe into this trend, their interaction energies were computed by taking the difference in the energies of the supermolecule $[C_n mim]^+$FePCys and its constituent members in complex. Upon inspecting interaction energies (Figure 5.45(b)), it is clearly observed that they echo the trend in free energies. The relative interaction energies become energetically significant as the alkyl chain length is enhanced beyond four.

The effect of the interaction on the active site molecule is also captured by measuring the bond length between the central Fe atom and attached S during the transition from the resting state to the bound state (S_1-S_2). The Fe-S bond lengths in the resting state at doublet was calculated to be 2.21 Å which is reported to increase in the presence of the substrate and underlying spin change to sextet [8]. Table 5.16 gives the above mentioned bond length in S_2 complexes in the presence of substrate. It is clearly seen that in the gas phase, the interaction strength of the Fe-S bond reduces upon binding of $[C_n mim]^+$ irrespective of the chain length and conformation presented to the FePCys molecule.



(a)



(b)

Figure 5.45: (a) Gibbs energy associated with formation of S_2 complexes (b) Interaction energies for $[C_n\text{mim}]^+$FePCys complexes

From a thermodynamic point of view, the free energies were further partitioned into their enthalpic and entropic components. Figure 5.46(a) gives the thermodynamic quantities evaluated for these contributions in the gas phase. Upon evaluating the entropic component at the standard conditions considered in the work, none of them seem to be energetically significant to the step leading to the for-

Table 5.16: Fe-S bond lengths in the presence of different conformations of $[C_n mim]^+$ cations

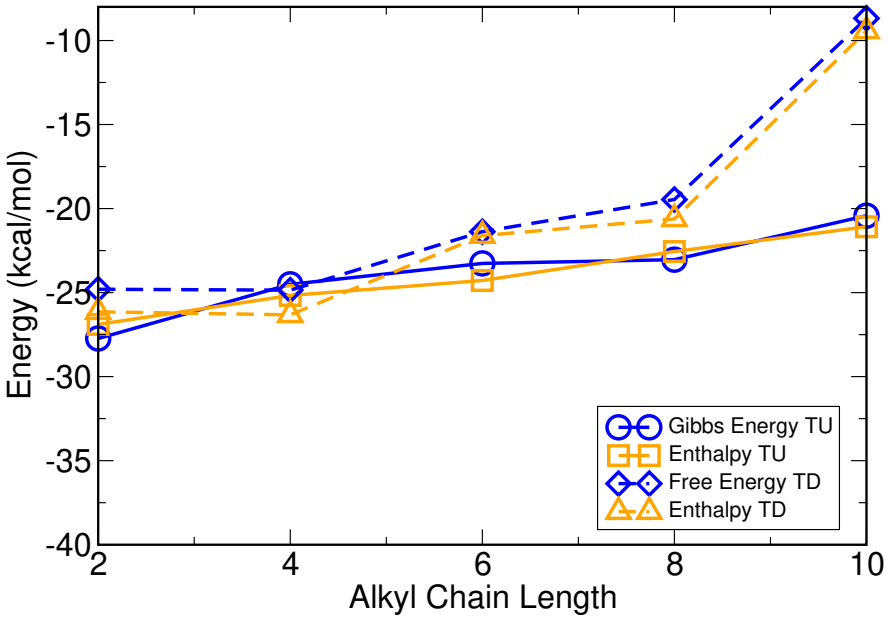
Cation	ϵ = 1		ϵ = 5.7	
	TU	TD	TU	TD
$[C_2 mim]^+$	2.27	2.28	2.31	2.31
$[C_4 mim]^+$	2.28	2.28	2.30	2.32
$[C_6 mim]^+$	2.28	2.28	2.30	2.32
$[C_8 mim]^+$	2.28	2.28	2.30	2.32
$[C_{10} mim]^+$	2.28	2.30	2.31	2.32

Table 5.17: Entropic contribution to the formation of S_2 complex for gas and embedded phase geometries (kcal/mol)

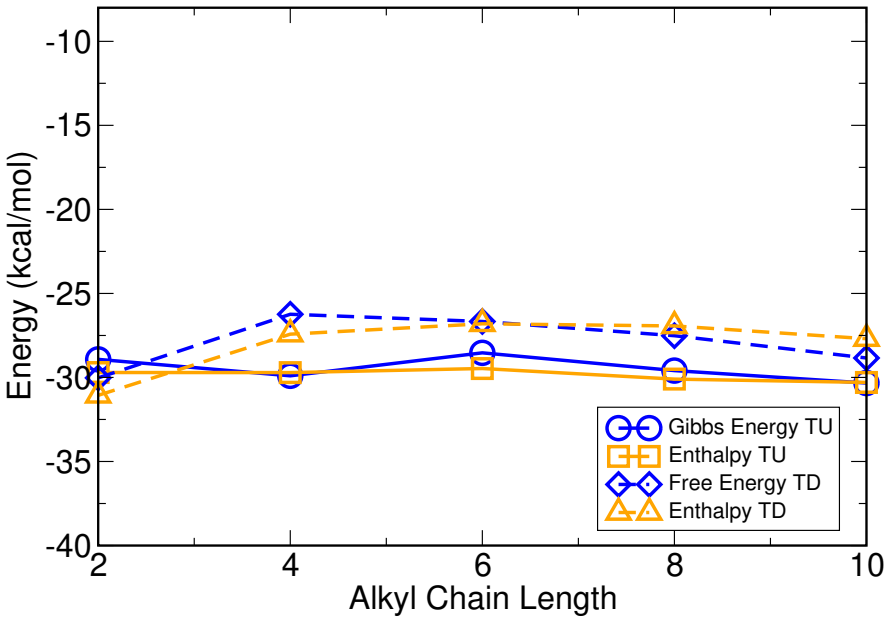
Cation	ϵ = 1		ϵ = 5.7	
	TU	TD	TU	TD
$[C_2 mim]^+$	0.8	−1.4	−0.8	−1.0
$[C_4 mim]^+$	−0.7	−1.5	0.2	−1.2
$[C_6 mim]^+$	−1.0	−0.2	−0.9	−0.1
$[C_8 mim]^+$	0.5	−1.2	−0.5	0.6
$[C_{10} mim]^+$	−0.7	−0.8	1.5	1.1

mation of the S_2 complex. Table 5.17 gives these contributions and the maximum energy in case of gas phase geometries is -1.5 kcal/mol which is energetically not significant.

Figure A37 gives the optimized geometries obtained from the embedded phase for the binding step. For TU geometries(1-5), the cations minimize to a geometry in which the ring remains upright rather than being parallel as in the case of gas phase. The alkyl chain points away from the underlying cysteinated porphyrin similar to the FeP containing systems investigated in the previous work. This holds true for all the cation sizes considered (2-10) and for all of them, the C4 carbon is observed to be the closest to the Fe atom of FePCys. The TD geometries (6-10), show two salient observations that are unique as compared to their gas phase counterparts; the first being the absence of the slithered behavior of the chain w.r.t FePCys and the second being the site exposed to underlying metal center. With



(a)



(b)

Figure 5.46: Thermodynamics for S_2 complexes optimized in (a) gas phase (b) ε = 5.7 for TU and TD conformations

the exception of the complex containing $[C_n\text{mim}]^+$, for all the other cations, the terminal carbon is exposed to the Fe center. In case of the gas phase structures, none of these terminal sites are available. For systems containing chain length 4-10, the ring is faraway from the purview of the FePCys molecule and hence, its interaction with the active site molecule is miniscule.

Table 5.18: Free energies associated with formation of S_2 evaluated by different cation entry environments

Cation	ϵ = 78.0		ϵ = 5.7	
	TU	TD	TU	TD
$[C_2mim]^+$	−28.9	−30.0	−41.0	−42.1
$[C_4mim]^+$	−29.9	−26.2	−41.5	−37.9
$[C_6mim]^+$	−28.5	−26.7	−41.2	−38.3
$[C_8mim]^+$	−29.6	−27.5	−40.8	−38.8
$[C_{10}mim]^+$	−30.3	−28.8	−39.4	−37.9

The Gibbs free energies for the step leading to the formation of S_2 complexes is given in Figure 5.45a. From a conformational standpoint, except for the butyl containing system, none of the cations show significant energetic deviation between TU and TD geometries. The separation between TU and TD conformations, in case of the complexation of $[C_4mim]^+$ is 3.6 kcal/mol. Also, a clear trend in terms of alkyl chain lengths is not noticed for the embedded geometries. For all of the members of the homologous series and conformations, the free energies energies are more favorable as compared to their gas phase counterparts. For these calculations, the cation was considered to be entering the binding pocket through an aqueous medium. A comparison of the binding energies for the step depending on cation entry is given in Table 5.18.

The interaction energies, however, for most of these geometries, are not favorable. Only the complexes containing $[C_2mim]^+$ are an exception to this behavior. This reflects the fact that the geometries optimized in the embedded medium are significantly not favorable as compared to the ones in gas phase. Also, the parallel interaction for TU geometries as opposed to an upright ring in embedded medium serves to stabilize the overall complex. For TD geometries, the interaction between the chain placed parallel to the FePCys plane containing the pyrrole rings also acts to stabilize the tail facing complexes, which is absent in the embedded medium.

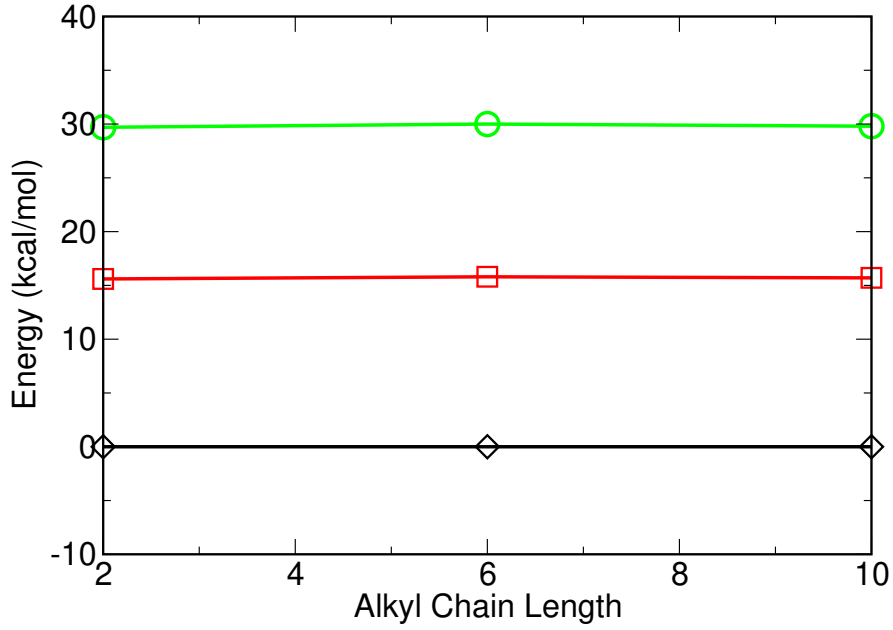
Figure 5.46b shows the enthalpic contribution to the free energy computed for the embedded geometries. It can be clearly observed that the entropic contributions are not energetically significant in the formation of these complexes even in

the presence of embedding. The entropic contributions are given in Table 5.17 among which the maximum energy is given by $[C_4mim]^+$ TD conformation with -1.2 kcal/mol.

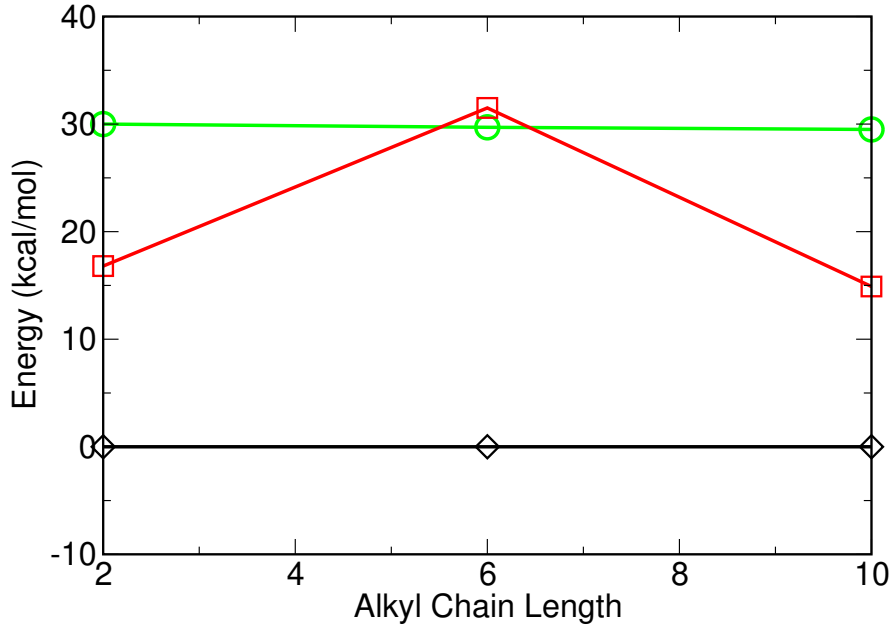
The interaction strength of the Fe-S bond in the complex also informs about the effect of the spin change as well as the effect of dielectric media on the active site species. This bond length was computed to be 2.22 Å for the resting state before binding which increases in response to the binding of substrate and spin change to 2.3 Å . This change in the interaction strength is shown by all of the members of the homologous series and conformations considered(Table 5.16).

5.5.3 First Reduction Step (S_2-S_3)

The binding of substrates to the active site FePCys molecule is followed by the first reduction step involving the uptake of an electron from the reductase domain in P-450(Figure 2.2. The addition of this electron alters the spin state of the complex and the stable state is shown to be quintet. For examining the spin states of the systems in this work, the relative energies were calculated for the possible spin states and given in Figure 5.47. It is clearly observed that for both sets of conformations, the quintet state is energetically most favorable.



(a)



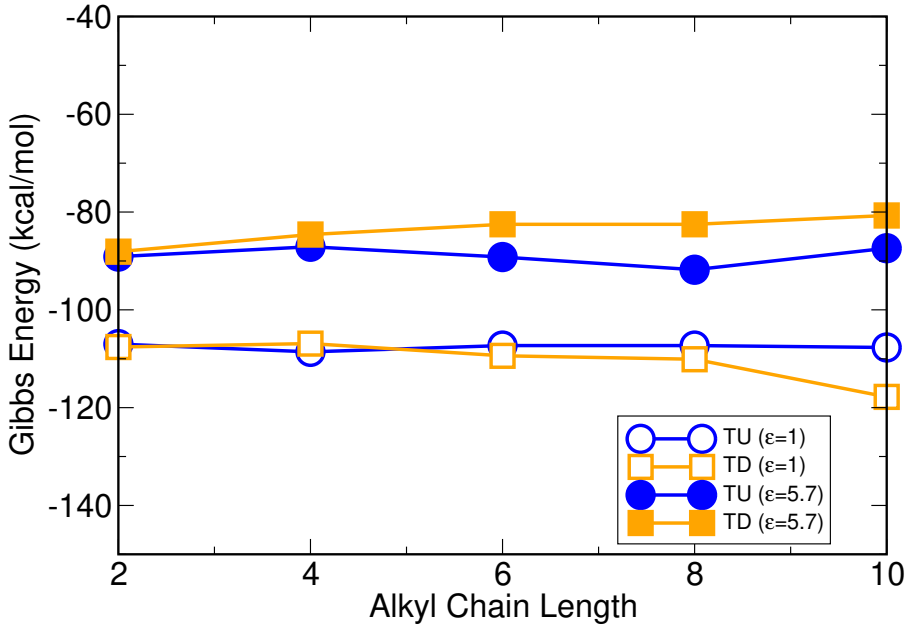
(b)

Figure 5.47: Relative energies of different spin states for S_3 (a) Tail Up and (b) Tail Down complexes optimized in ε = 1; Legend : Singlet○ , Triplet □, Quintet ◇

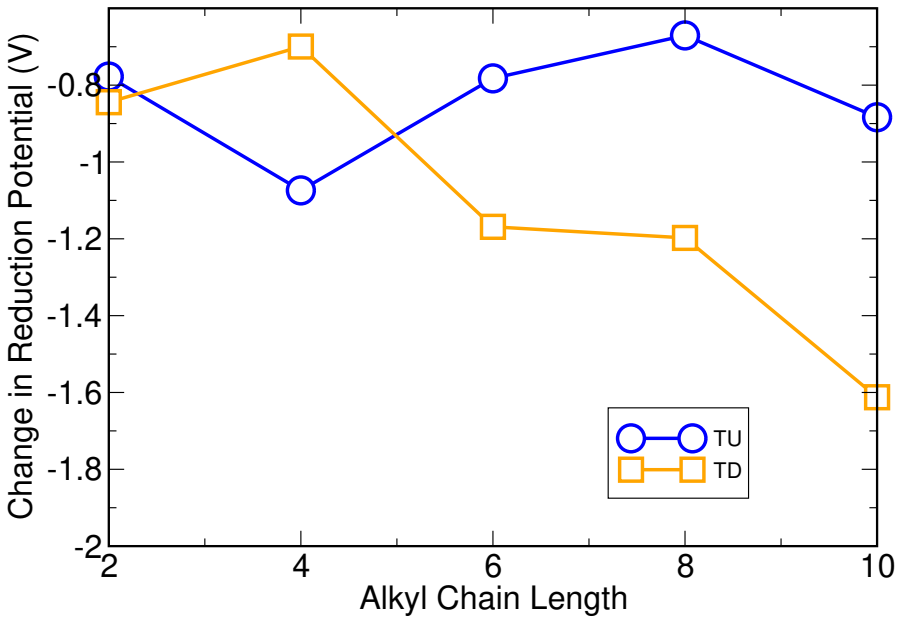
The gas phase optimized geometries are given in Figure A38 in the order of the homologous series of cations considered. The TU complexes retain their conformations from the previous binding step with the ring remaining parallel to porphyrin while the chain pointing away. In case of the TD complexes however, the ring

is drawn in towards the Fe center of the porphyrin while the chain remains in a slithered position. Figure 5.49 shows this character of S_3 complexes (in red) for C_6-C_{10} chain containing TD conformations. These conformations differ from TU ones due to the positioning of the alkyl chains. The free energy for this step is given in Figure 5.48a. For the TU complexes, the free energies remain within the energy of 1.5 kcal/mol throughout the homologous series owing to their similarity in geometry. In the case of TD complexes, however, the feasibility of the process is enhanced as the alkyl chain is increased on the cation with the energies ranging from -106.9 to -117.9 kcal/mol. Thus, favorability is enhanced for the TD conformations as compared to the TU ones.The relative free energies reflect this behavior as the chain length is enhanced from butyl to hexyl and greater with $[C_{10}mim]^+$ conformations being separated by 10 kcal/mol. In the interest of site selectivity, there seems to be equivalence in the orientation of the cation site exposed to Fe for nucleophilic attack. The electron addition is shown to remove the conformational bias previously exhibited by S_2 complexes.

The entropic contributions to the reduction process are given in Table 5.19. For the gas phase geometries, most of the entropic contributions are positive and energetically not significant. Only the $[C_6mim]^+$ and $[C_{10}mim]^+$ TD conformations deviate from this observation with 3 and 3.6 kcal/mol, respectively.



(a)



(b)

Figure 5.48: Gibbs energy associated with (a) Reduction of FePCys in complex (b) Relative Reduction Potential w.r.t ε = 5.7 computed for conformations in ε = 1

In order to compute the affinity for the reduction step, the reduction potentials of the systems in complex and substrate free states were computed. In the gas phase, due to higher free energies as compared to the embedded phase, the reduction potentials were shown to be greater in magnitude. This is expressed in the Figure 5.50b in which the relative potential were calculated to be consistently

above embedded geometries. The substrate-free potential was computed to be about 2.1 V that was far exceeded by the presence of cations irrespective of the size of alkyl chain length and conformation presented to FePCys.

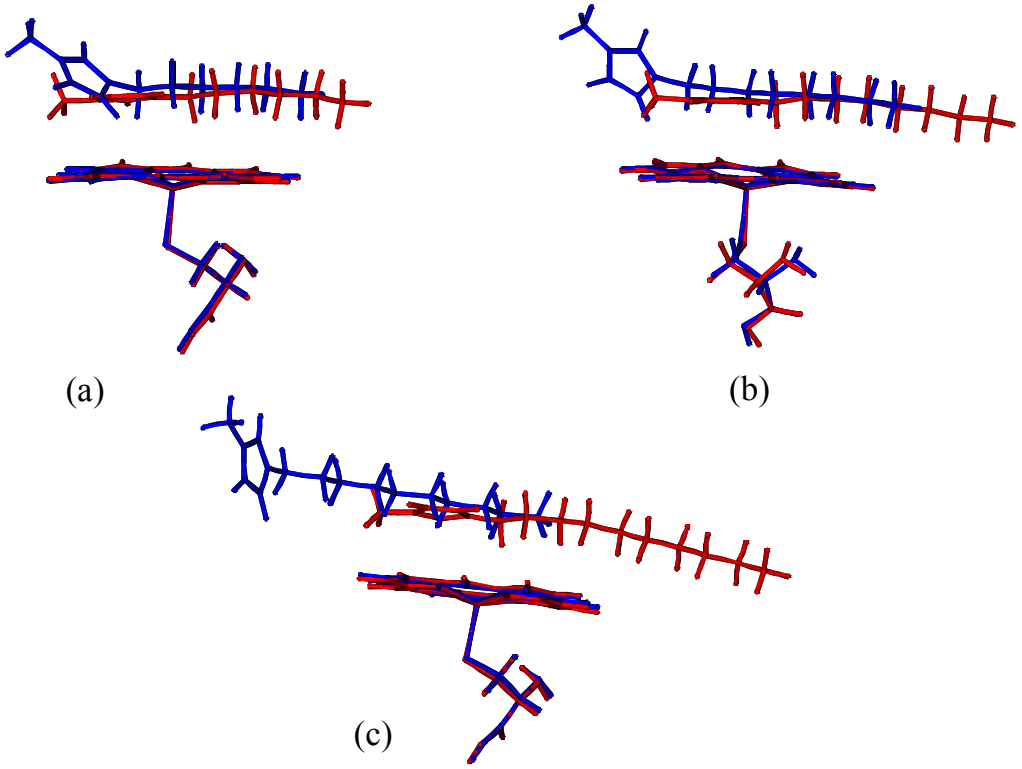


Figure 5.49: Alignment of (a) $[C_6mim]^+$ (b) $[C_8mim]^+$ and (c) $[C_{10}mim]^+$ containing complexes (in blue) with the geometries after reduction (S_3) (in red)

Geometries optimized in the embedded phase are given in Figure A39. For TU conformations(1-5), except for $[C_8mim]^+$, none of the other cations place their rings parallel to FePCys. For ethyl, hexyl and decyl cations, the alkyl chain points away from the porphyrin and the geometries are similar to the ones observed for S_2 complexes. The chain facing TD geometries (Figure A39, 6-10) for $[C_2mim]^+$ and $[C_4mim]^+$ show a slithering behavior of the alkyl chain with the carbon nearest to the ring exposed to the Fe center. For chain lengths greater than four, the chain is

Table 5.19: Entropic contribution to the formation of S_3 complex for gas and embedded phase geometries (kcal/mol)

Cation	ϵ = 1		ϵ = 5.7	
	TU	TD	TU	TD
$[C_2mim]^+$	1.6	−0.7	−0.6	0.8
$[C_4mim]^+$	1.3	0.5	−0.9	−6.4
$[C_6mim]^+$	0.4	3.0	−0.8	0.4
$[C_8mim]^+$	1.7	1.1	2.7	−0.3
$[C_{10}mim]^+$	0.9	3.6	1.1	0.6

tilted although the terminal carbons attached to the alkyl chain are still exposed to the Fe center for probable reactive attack. The free energies (Figure 5.48a) reflect this trend in the geometries with the $[C_8mim]^+$FePCys TU complex yielding a free energy of almost -92 kcal/mol which is the most favorable among the whole homologous series. For the TD conformations as well, the free energy undergoes a significant drop from $[C_4mim]^+$ to $[C_6mim]^+$ containing complexes with the difference between them being about 8 kcal/mol with the $[C_6mim]^+$ being relatively unstable. It is also noteworthy that with the exception of ethyl cation, all other complexes show a gradual decrease in their favorability as the alkyl chain length increases from 4-10.

Upon comparing these free energies with the ones computed for the gas phase geometries, it is clearly observed that embedding the geometries within the medium incurs an energetic penalty. This results in free energies in the gas phase to be consistently more favorable as compared to embedded ones, with the minimum and maximum relative free energies being 15 and 37 kcal/mol respectively. From the entropic contributions, it is clear that only $[C_8mim]^+$FePCys TU and $[C_4mim]^+$FePCys TD complexes are energetically affected by it. For them, the energies are reported to be 2.7 and -6.4 kcal/mol, respectively. All other geometries show contributions of 1.1 kcal/mol or less.

Reduction potentials computed for the step in embedded geometries(Figure 5.50c) could be directly compared to some of the known and biodegradable substrates of cytochrome P-450 reported in literature. In the present work, the substrate free

potential level was estimated to be at about 3.38 V as compared to 3.7 V for a wild type P-450 calculated with a similar model in Shaik et al. [8]. In the presence of cations, except for the $[C_4mim]^+$FePCys complex, the TU conformations report a more favorable potential as compared to TD complexes. However, the reduction of the imidazolium-based cation complexes are less favorable as compared to free energies for some of the natural substrates bound to the P-450 BM3 wild type enzyme. Among these, camphor has been shown to be the most favorable, which is followed by arachidonic and palmitic acid, with propylbenzene being the closest to the level of reduction promoted by ionic liquid cations.

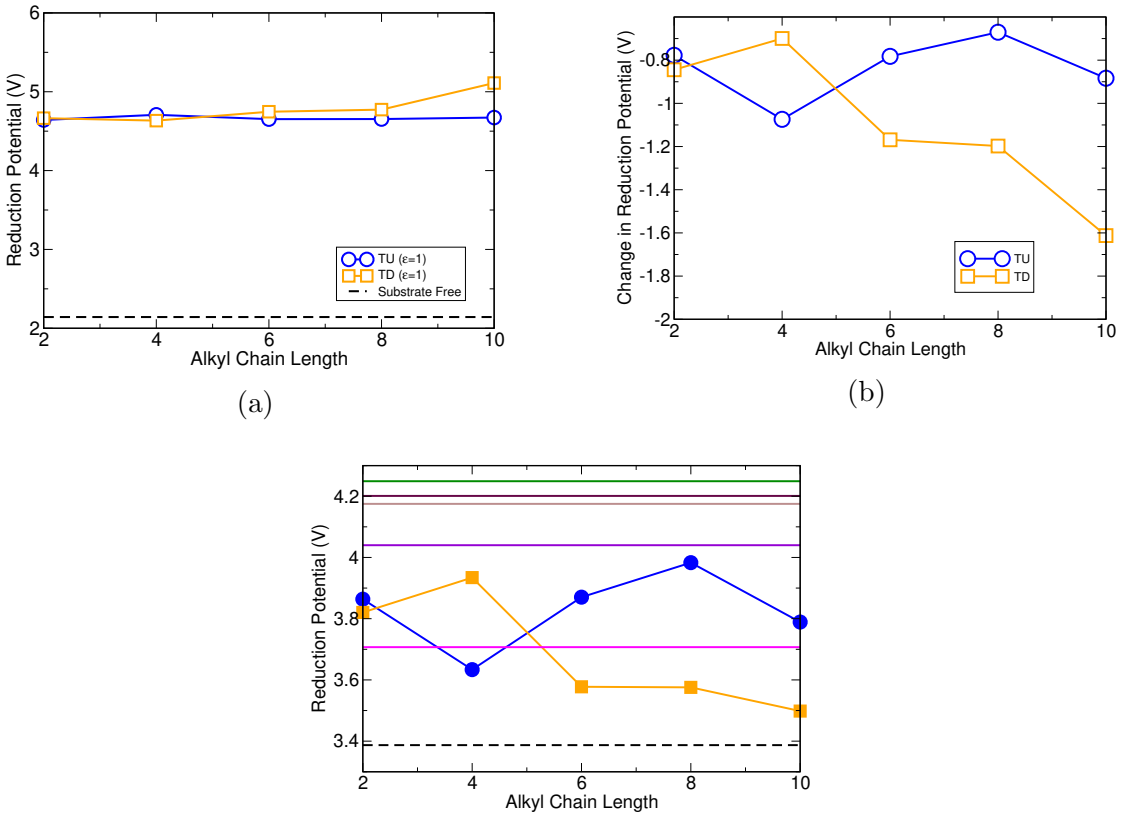
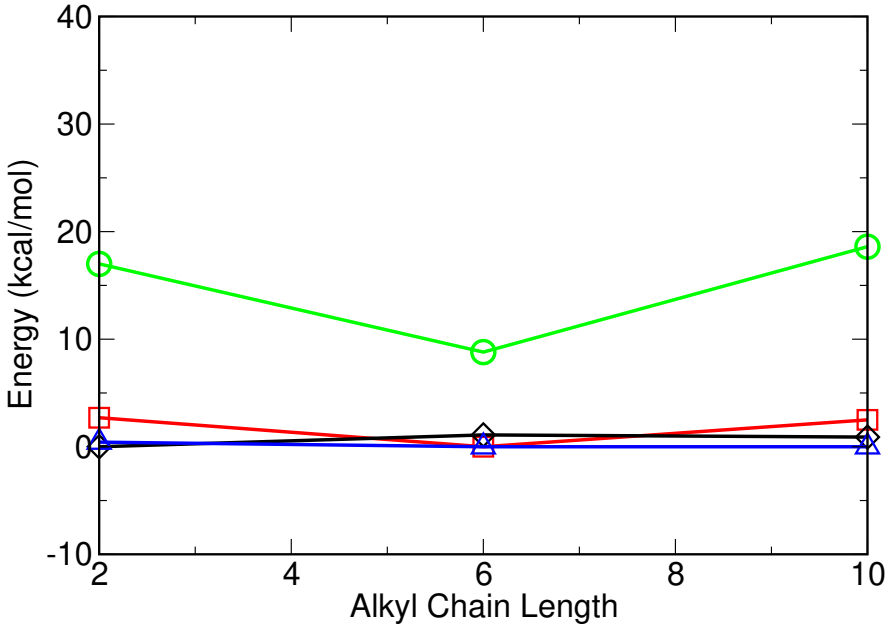


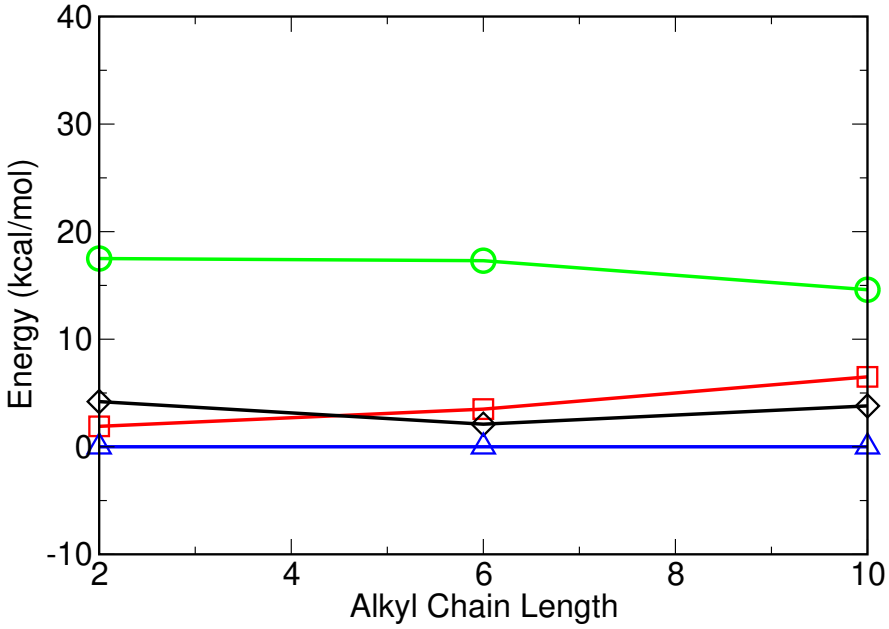
Figure 5.50: (a) One electron reduction potential computed for Gas Phase (b) Relative Reduction Potential w.r.t ε = 5.7 computed for conformations in ε = 1 (c) One electron reduction potential computed for Embedded system (ε = 5.7) ; Legend for (c): TU ● , TD ■ , Substrate Free – – – , Substrate Free [7] (exp) , Camphor [8] (pred) , Palmitic Acid [9] (exp) , Arachidonic Acid [9] (exp) , Propylbenzene [9](exp)

5.5.4 Oxidation Step (S_3-S_4)

Reduction of FePCys is followed by addition of triplet dioxygen to the Fe-center of heme in P-450 that competitively binds with the ligand bound to the porphyrin molecule. The resulting spin state becomes an open shell singlet bearing a radical nature. The relative energies in the gas phase are given in Figure 5.51.



(a)



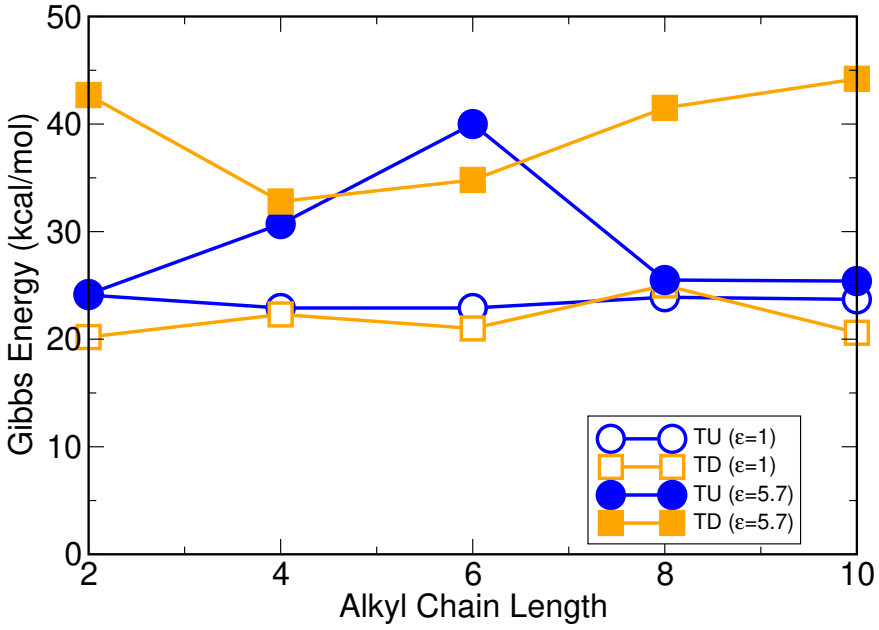
(b)

Figure 5.51: Relative energies of different spin states for S_4 (a) Tail Up and (b) Tail Down complexes optimized in ε = 1; Legend : Singlet○ , Triplet □, Quintet ◇ Open Shell Singlet △

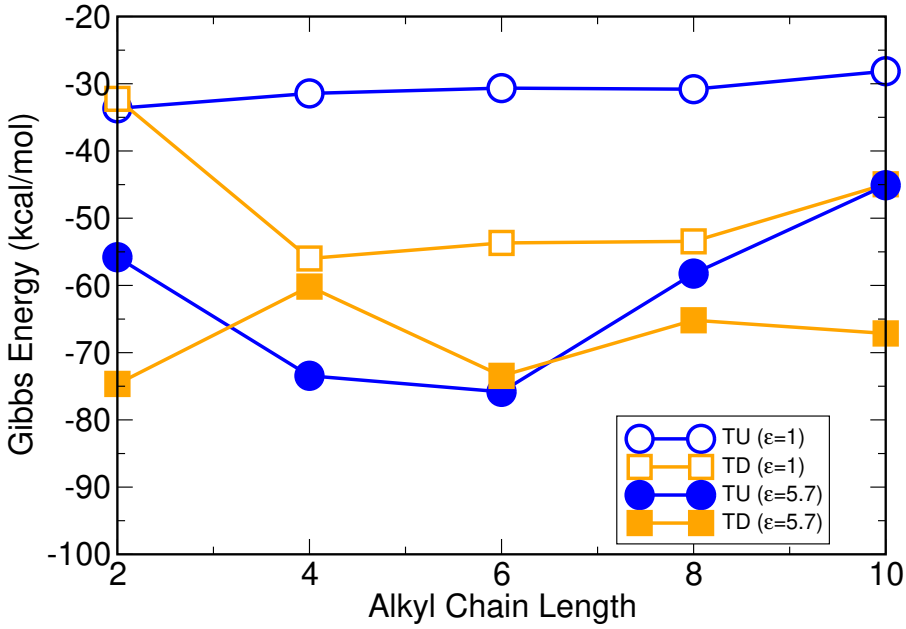
In the substrate free mode, the P450 binds a triplet dioxygen in an end on coordination scheme [10] and it was used as a template for the substrate binding.The end oxygen atom is attached to Fe (O_1) and the other oxygen atom is attached to the former (O_2). The gas phase optimized geometries are given in Figure A40.

For TU geometries(1-5), the ring interacts with O_2 and creates a potential contact with the acidic C2 hydrogen. This observation also holds true for the TD complexes in which the ring does not show a parallel arrangement w.r.t the plane of the FePCys molecule.

Figure 5.52a shows the free energies associated with the formation of oxy-ferrous complexes. Calculations indicate that this process is thermodynamically infeasible irrespective of the electronic environment provided to the ligand-FePCys complex. The free energies irrespective of the conformation and system size, are within the bounds of 5kcal/mol. It is worth noting that our calculations for camphor using the same treatment of level of theory and basis set also yield a thermodynamically unfavorable free energy of 43.2 kcal/mol. It has been established that camphor is biodegradable entity [8] and aids in the movement of the catalytic cycle of cytochrome P-450. The thermodynamic contributions to the computed free energies are given in Figure 5.53a. It can be easily observed that unlike the two former steps, entropic contributions are not only energetically significant but exceed the enthalpic levels. The energetic infeasibility can be attributed to two reasons, the first being competitive binding of the electron rich molecule. Since the ring of the cations are associated with FePCys in parallel arrangement, a certain energetic penalty is encountered when the favorable $\pi - \pi$ arrangement is disturbed. This is true in case of TU complexes where the O_2 moiety is inserted into the interplanar arrangement and the positively charged imidazolium core is displaced from its equilibrium position in the S_3 complex.



(a)



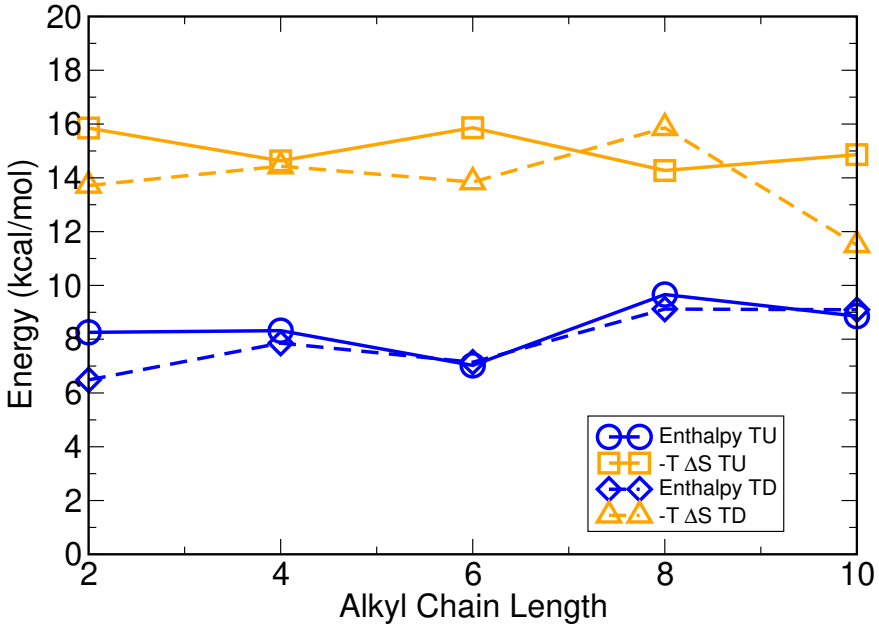
(b)

Figure 5.52: Free energies computed associated with the formation of S_4 and S_5 complexes in gas and embedded phase

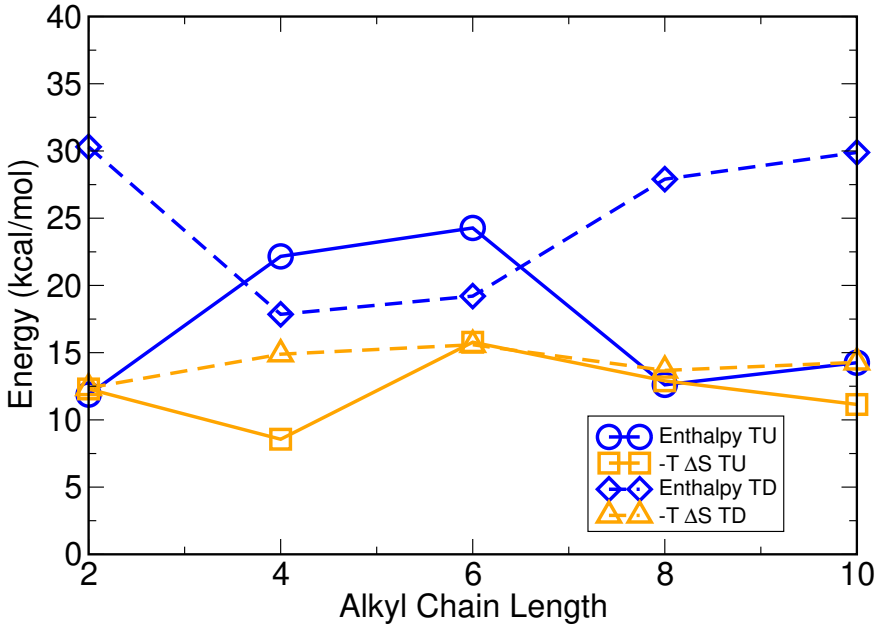
Figure A41 gives the geometries optimized in embedded phase. For TU structures(1-5), there is no clear trend or tendency reported for the cations in the complex. Except for the $[C_4mim]^+$FePCys complex, all of the rings remain upright with multiple atoms within the hydrogen bonding range (2-2.5 Å) that contribute to their stability. In the case of the butyl complex, only the hydrogen attached to the

3'-methyl carbon is probable of participating in nucleophilic attack by O_2. The TD geometries show these probable interactions for all of the cations in complex and the sites are vary from ring to alkyl chain hydrogens. For $[C_4mim]^+$ cation, the probable site of interaction belongs to the C2 hydrogen atom located on the ring. The ethyl and hexyl cations in complex expose the hydrogen attached to the immediate chain carbon atom to the N3 position. For the two greatest cations, the hydrogens closest to the proximal oxygen are at $\omega - 3$ and ω positions.

The free energies for the complexes in embedded phase also show a thermodynamic infeasibility similar to the gas phase geometries. In fact, these energies are even more positive as compared to the gas phase values which might be attributed to the presence of dielectric media as for the two previous steps (binding and first reduction). The thermodynamic contributions (Figure 5.53b) do not show a specific trend although, similar to gas phase, the entropic contributions are energetically significant and play an important role in dictating the final free energies for the process. For the TU geometries, the entropies are comparable in case of $[C_2mim]^+$ and $[C_8mim]^+$FePCys complexes and exceed the enthalpies by about 0.4 and 0.3 kcal/mol. On the other hand for the TD complexes, the enthalpic contributions significantly exceed the entropies irrespective of the size of cation in complexation with FePCys. The binding of the substrate also influences the strength of the dioxygen ligand bound to the FePCys molecule to reflect the competitive binding behavior.



(a)



(b)

Figure 5.53: Thermodynamic contributions to the free energies for S_4 complexes optimized in (a) gas phase (b) ε = 5.7 for TU and TD conformations; ΔG = $\Delta H - T\Delta S$, the quantities plotted are Enthalpy (ΔH) and Entropic Contributions ($-T\Delta S$)

Table 5.20: Gibbs free energies for the formation of ferric-peroxo complexes for structures optimized from gas phase (ϵ = 1) and solvated (ϵ = 5.7) environments (kcal/mol)

Cation	ϵ = 1		ϵ = 5.7	
	TU	TD	TU	TD
$[C_2mim]^+$	−33.7	−32.2	−55.8	−74.7
$[C_4mim]^+$	−31.4	−55.9	−73.4	−60.2
$[C_6mim]^+$	−30.7	−53.7	−75.8	−73.4
$[C_8mim]^+$	−30.8	−53.4	−58.2	−65.2
$[C_{10}mim]^+$	−28.1	−45.0	−45.1	−67.2

5.5.5 Second Reduction Step (S_4-S_5)

Formation of the ferric-peroxo complexes is associated with an electron addition to the oxyferrous complexes.

The addition of an electron to the oxyferrous complex does not lead to a significant change in their geometries in the gas phase(Figure A46). For all of the members of the homologous series, the TU complexes retain the contact with the C2 acidic hydrogen and hence, the site selectivity is unchanged. For TD complexes, the rings were shown to be non-parallel for all the geometries except for the $[C_{10}mim]^+FePO_2Cys$ (S_4). In the case of the S_5 complexes, the $[C_{10}mim]^+$ containing complex also shows this parallel behavior although the hydrogen site closest to O_2 is still shown to be the acidic one as mentioned above. In line with the S_4 complexes, the first three cations, i.e. ethyl, butyl and hexyl show an upright ring position with the same site exposed to O_2. The free energy profile is described in Figure 5.52b. For TU complexes, the free energies are shown to be between -33.6 to -28 kcal/mol with a linear relationship with respect to the alkyl chain length that can be related to geometries. In case of the TD complexes, the variability in free energies are shown in the profile with the range extending from -56 to -32.2 kcal/mol. While the free energy for the formation of the $[C_2mim]^+$ complex is the lowest, for all the other cations, this free energy is significantly

higher than $[C_2mim]^+$. From a thermodynamic standpoint, Table 5.21 gives the entropic contributions to the free energy associated with the formation of these complexes. As the chain length is enhanced for both the sets of conformations (TU and TD), the entropic contributions become steadily unfavorable. The most important observation is related to the reduction potential associated with this process. Figure 5.54a shows these potentials in both substrate free and substrate bound modes. It can be easily inferred that in the absence of the cation, the negative one electron reduction potential reflects a positive free energy. In the bound state, this potential reflects thermodynamic feasibility as they are elevated by 3 V or higher irrespective of the size of the cation and conformation considered.

Table 5.21: Entropic contribution to the formation of S_5 complex for gas and embedded phase geometries (kcal/mol)

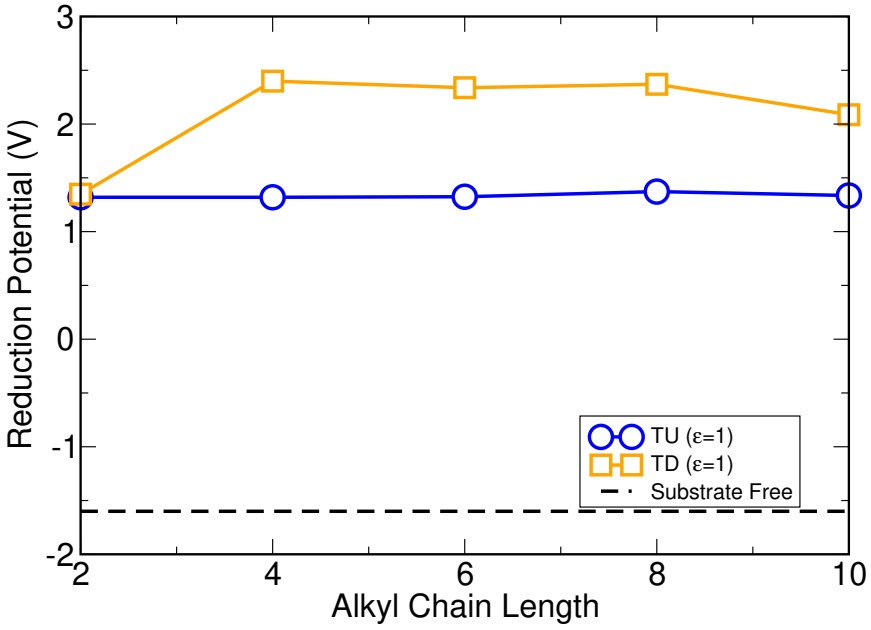
Cation	ϵ = 1		ϵ = 5.7	
	TU	TD	TU	TD
$[C_2mim]^+$	−3.2	−1.1	0.9	0.2
$[C_4mim]^+$	−1.0	−0.7	1.3	0.7
$[C_6mim]^+$	−0.1	0.2	−0.2	−0.1
$[C_8mim]^+$	0.8	1.2	−0.3	1.3
$[C_{10}mim]^+$	2.7	3.2	2.2	0.0

Figure A47 shows the geometries of S_5 complexes in the embedded phase. Similar to the gas phase, the geometries remain relatively unperturbed by the addition of the electron. In line with the oxy ferrous complexes (S_4), the TU geometries all show interaction of the O_2 atom with the C2 acidic hydrogen. The TD complexes however, show different sites of exposure to O_2 with the ethyl cation showing $\omega - 1$, the butyl showing the C2 acidic hydrogen and upon increasing chain length from 4-6 and beyond, the alkyl chain sites are being exposed. In the interest of site selectivity, the two greatest cation chain lengths show terminal (ω) site for potential nucleophilic attack. The free energies (Figure 5.52b) for the embedded

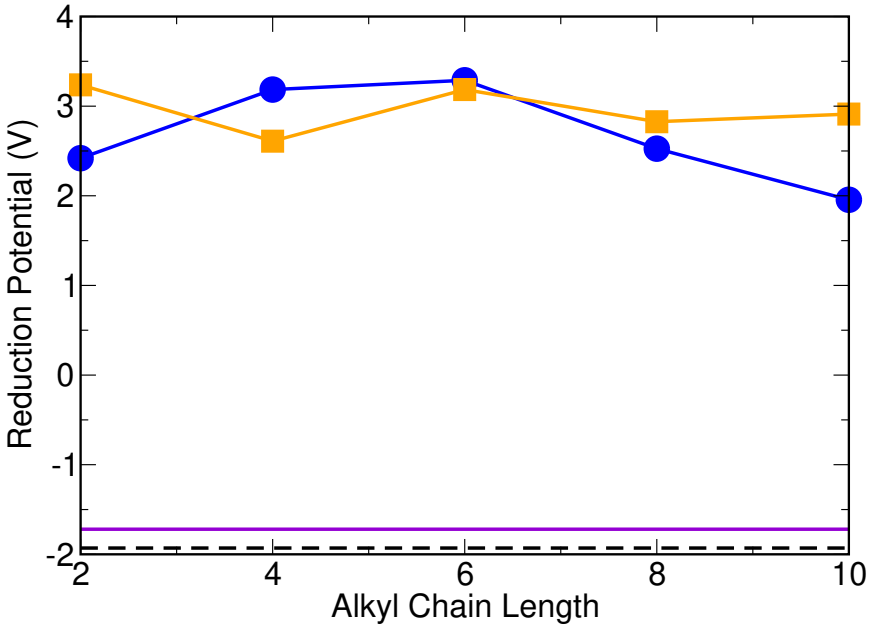
phase show more favorability as compared to the gas phase but conformationally a specific trend is not observed. Although, the effect of the medium is clearly inferred by comparing a certain conformation from both phases with the embedded phase always being significantly more favorable. The entropic contribution (Table 5.21) was shown to have no observable trend and the effect of alkyl chain length is not reflected. Only the $[C_{10}mim]^{+}FePO_2Cys^{-}$ TU complex shows a change in entropy of 2.2 kcal/mol which energetically significant. The reduction potentials (Figure 5.54b) show the effect of the binding in case of the active site ferric-peroxo complex. In the absence of substrate the one electron reduction potentials show a negative value that has also been observed for the oxy-heme complexes by Wang et al. [10] in a QMMM framework and Rydberg [11] in a QM only scheme. The presence of the $[C_n mim]^{+}$ elevates these reduction potentials significantly irrespective of the size and conformation presented to the active site molecule. These substrate induced potentials have been shown to be essential in driving the catalytic cycle of cytochrome P-450.

5.5.6 First and Second Protonation Steps (S_5-S_6 and S_6-S_7)

The first protonation step forms a ferric-hydroperoxide complex in which a proton is attached to the existing dioxygen moiety on the distal side of the FePCys. Figure 5.55a shows the Gibbs free energies associated with this process. On a cumulative scale, the first protonation is the most thermodynamically feasible step out of all in the catalytic cycle. The computed free energies are computed to be even more favorable than -300 kcal/mol for the gas phase. Chemically, the affinity of the electron rich peroxo complex towards the incoming proton suggests the same. A variety of studies [260, 261] have suggested this high affinity in line with our observations. In the gas phase, there is a clear conformational dependence for the addition of the initial proton with energies deviating in the order of $\simeq$15 kcal/mol with tail up conformations making it more favorable with the exception of the C_2 cation containing systems. This favorability is of the order of 14 kcal/mol and although entropically this process is shown to be not favorable. Table 5.22 shows the



(a)



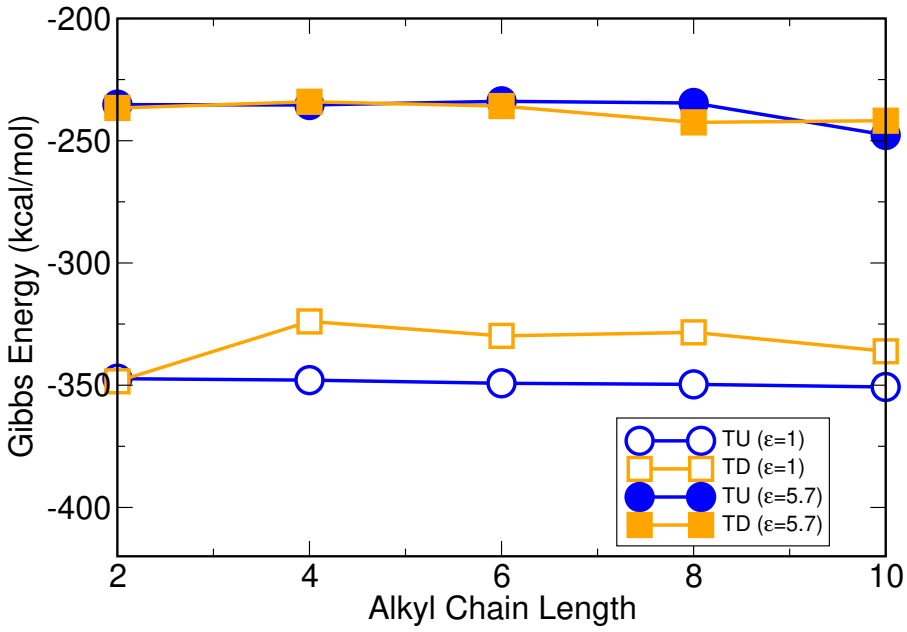
(b)

Figure 5.54: Reduction Potential for the second reduction step in (a) gas phase and (b) embedded phase; Legend for (b) TU ● , TD ■ , Substrate Free (QMMM, calc) [10] – – , Substrate Free (QM, ε = 80) [11]

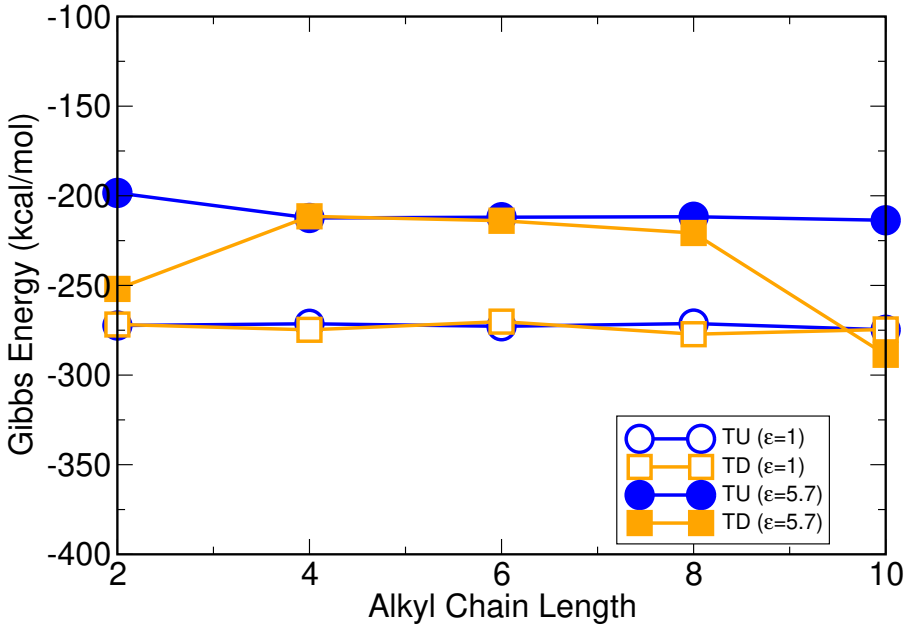
change in entropy for the gas phase proton addition to be positive for both sets of conformations. Upon closer inspection, the entropy changes become less positive as the chain length is enhanced for TU conformations. A subsequent protonation leads to the formation of the highly stable ultimate active species of the cycle (S_7).

The energies associated with this process are depicted in Figure 5.55b. Upon inspection, there seems to be little conformational dependence towards this process with the exception of C_2-containing complex in dielectric medium. All of the energies in the gas phase are within $\simeq$ 6 kcal/mol which is not highly significant. Thus, it can be concluded that formation of the doublet active site species is not dependent on the conformations presented to the FePCys molecule derived from a synchronous approach. Table 5.23 shows that the entropy change associated with this process is favorable as compared to the previous protonation. Together the two protonation steps are crucial in driving the overall catalytic cycle to thermodynamic favorability as they are 10 fold or higher than the S_4 formation step.

For the embedded phase, the free energies are also shown to be comparable to the gas phase and highly favorable with respect to the other steps in the catalytic cycle. The free energies for this step are computed to be between -233 to -248 kcal/mol and are less favorable as compared to the gas phase. Another key observation is the fact that the relative energies between the conformations is not as pronounced as those for the gas phase. Upon partitioning the free energy into enthalpic and entropic contributions, it is inferred that change in entropy plays a vital role in dictating the free energies. Table 5.22 shows the importance of this change in entropies ranging from 4.0 to 7.4 kca/mol. Thus, the favorability due to the presence of media is on the geometry itself is still energetically significant.
The second protonation leading to the formation of the ultimate active species has also shown to be highly favorable in literature for the embedded phase [260]. Our computations suggest that the free energies for this step is in the range of -209 to -288 kcal/mol. In the work by Harris et al. [260], the proton affinity leading to S_7 is computed to be -334 kcal/mol and comparing our results to this value, both the conformations reduce this affinity significantly. Only C_2 and C_{10} TD conformations yield a free energy higher than it. Entropically, the contributions are even more pronounced for this step as compared to the previous protonation step. They are energetically favorable leading to the stable ultimate active species and range from -43 to -4 kcal/mol. It can be concluded that the second hydroxylation



(a)



(b)

Figure 5.55: Gibbs energy associated with (a) First protonation (b) Second Protonation steps for complexes in gas phase and dielectric solvation

is entropically favorable while the first one is not.

Table 5.22: Entropic contribution to the formation of S_6 complex for gas and embedded phase geometries (kcal/mol)

Cation	ϵ = 1		ϵ = 5.7	
	TU	TD	TU	TD
$[C_2mim]^+$	8.6	10.4	6.9	4.0
$[C_4mim]^+$	8.1	9.2	6.7	7.4
$[C_6mim]^+$	6.9	6.1	7.5	6.5
$[C_8mim]^+$	6.5	4.6	5.8	7.7
$[C_{10}mim]^+$	4.3	6.5	7.4	6.2

Table 5.23: Entropic contribution to the formation of S_7 complex for gas and embedded phase geometries (kcal/mol)

Cation	ϵ = 1		ϵ = 5.7	
	TU	TD	TU	TD
$[C_2mim]^+$	−5.8	−6.1	−4.0	−43.1
$[C_4mim]^+$	−6.2	−8.3	−5.5	−5.5
$[C_6mim]^+$	−6.6	−5.4	−4.7	−5.9
$[C_8mim]^+$	−4.9	−7.5	−5.8	−5.8
$[C_{10}mim]^+$	−4.8	−9.8	−5.6	−4.6

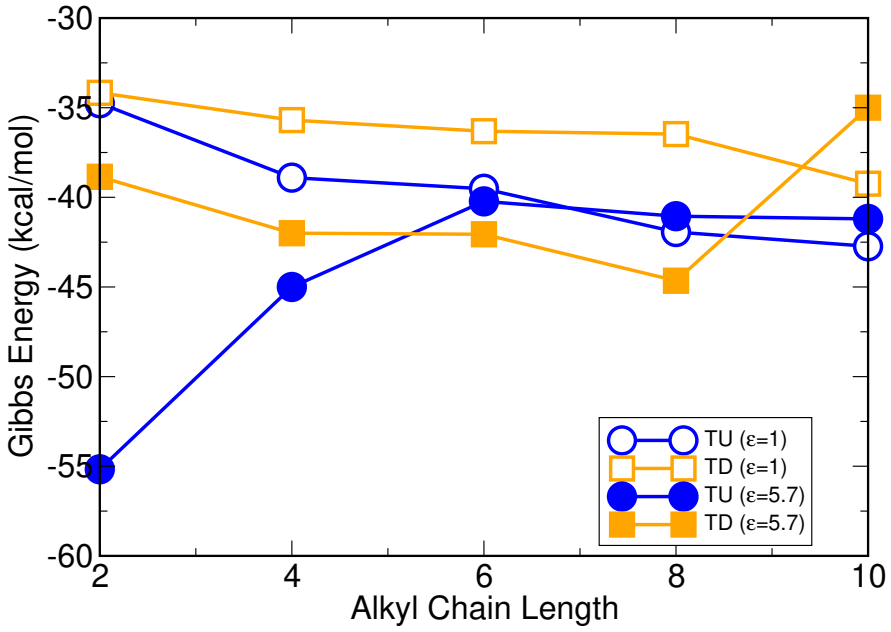
5.5.7 Hydroxylation Step (S_7-S_1)

Finally, the oxygen insertion into the cation takes place which culminates the catalytic cycle. The free energies computed for this step have been provided in Figure 5.56a. For computing the free energies associated with hydroxylation, the cationic proton that was closest to the distal oxygen attached to FePCys was considered and a list of these sites are provided in Table 5.25.

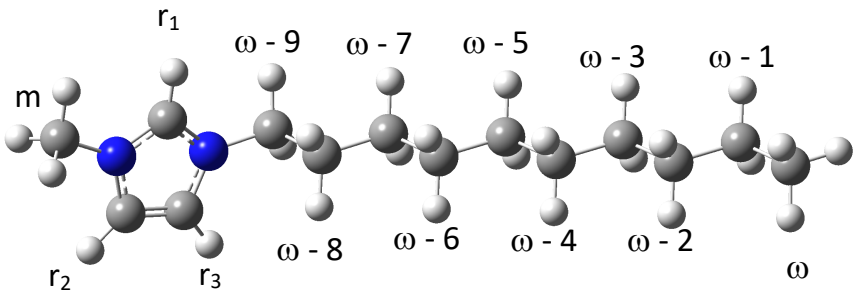
All of the values suggest that formation of alcohol from the complexes consisting of the active site species are thermodynamically feasible. The computed free energies lie in the range of -34 to -43 kcal/mol. While for the TU conformations, the free energies for the homologous series lie within 8 kcal/mol, for the TD conformations, it is 5 kcal/mol. For both sets of conformations, the free energies become more favorable as the alkyl chain is increased. On a relative scale, the free energies are more favorable towards the TU conformations as compared to the TD conformations. The entropic part of the free energy (Table 5.24) is shown to be favorable for all the complexes studied and energetically significant for all of the TU complexes. In case of the TD complexes, only C_2 and C_6 cations show energies beyond 2 kcal/mol. Table 5.26 gives the most probable sites for hydroxylation and for most of the geometries, the r_1 position i.e. the most acidic hydrogen is shown to be closest to the distal oxygen. The only exception being the C_2 TD complexes that shows exposure at the ω hydrogen for electrophilic attack.

In the embedded phase, the free energies are also shown to be favorable with consideration to the size and conformation presented to the ultimate active species. The free energies are predicted to be in the range of -35 to -55 kcal/mol. The TU conformations show that the free energies become less favorable as the alkyl chain length is increased which is in contrast to the observation for gas phase TU geometries. The TD geometries do not show any such observable trends w.r.t the chain length of the cation. Upon partition of these free energies and looking at the entropic contribution (Table 5.24), it is evident that the process is entropically favored for all of the systems similar to the gas phase. For this dielectric envi-

ronment, the entropic contributions are on an average basis, higher than those of the gas phase. Also, except for the C_4 cation containing system, all others show the entropic contribution to be energetically significant. The most probable sites for hydroxylation in case of the embedded geometries vary in position. For TU conformations, the ring sites r_1 and m are shown to be probable, which are the C2 acidic hydrogen and methyl sites respectively. In case of the TD conformations, the hydroxylation sites migrate from ring to alkyl chain levels with the highest cation expected to yield an alcohol at the terminal position.



(a)



(b)

Figure 5.56: (a) Free energies associated with the hydroxylation step leading to the formation of the resting state species for gas phase and embedded geometries and (b) Possible cationic sites for hydroxylation

Table 5.24: Entropic contribution to the formation of S_1 complex for gas and embedded phase geometries (kcal/mol)

Cation	ϵ = 1		ϵ = 5.7	
	TU	TD	TU	TD
$[C_2mim]^+$	−2.0	−3.2	−4.8	−2.9
$[C_4mim]^+$	−2.6	−1.8	−1.8	−1.3
$[C_6mim]^+$	−2.8	−3.5	−6.0	−4.1
$[C_8mim]^+$	−3.2	−1.8	−4.9	−5.1
$[C_{10}mim]^+$	−4.0	−1.3	−3.8	−2.3

Table 5.25: Gibbs free energies for hydroxylation step for structures optimized from gas phase (ϵ = 1) and solvated (ϵ = 5.7) environments (kcal/mol)

Cation	ϵ = 1		ϵ = 5.7	
	TU	TD	TU	TD
$[C_2mim]^+$	−34.8	−34.2	−55.2	−39.1
$[C_4mim]^+$	−38.9	−35.7	−45.3	−42.1
$[C_6mim]^+$	−39.5	−36.3	−40.2	−42.1
$[C_8mim]^+$	−41.7	−36.5	−41.1	−44.7
$[C_{10}mim]^+$	−42.7	−39.2	−41.2	−35.6

Upon analyzing the total free energies (Table 5.26) and supporting energy profiles (Figures 5.57 and 5.58), it is evident that the catalytic cycle as a whole is feasible and conformational bias towards these energies is quite minimal as compared to the initial binding step. The free energies are more favorable in the case of the gas phase and there seems to be a significant difference (≥ 15 kcal/mol) in this favorability when compared to the presence of dielectric media. It can be easily inferred from the results that in the gas phase, most of the conformations actually present an alkyl chain site to the ultimate active site species for de-protonation. In fact, only the ethyl cation optimized from the TD pose presents the terminal position for the hydroxylation process. Although, in case of the protein embedding, as the alkyl chain length is enhanced for the TD poses, chain sites seem to be exposed to the active site species for reactive attack and for $[C_{10}mim]^+$, the terminal site is the most favorable for this attack. This presents an encouraging view of the binding process in the pure DFT framework consisting of just the ac-

Table 5.26: Total Gibbs free energies for entire catalytic cycle for structures optimized from gas phase (ϵ = 1) and solvated (ϵ = 5.7) environments (kcal/mol)

Cation	ϵ = 1				ϵ = 5.7			
	TU		TD		TU		TD	
	Site	ΔG	Site	ΔG	Site	ΔG	Site	ΔG
$[C_2mim]^+$	r_1	-798.8	ω	-798.7	r_1	-641.3	r_2	-638.4
$[C_4mim]^+$	r_1	-799.8	r_1	-799.8	m	-652.7	r_1	-626.1
$[C_6mim]^+$	r_1	-799.9	r_1	-799.8	r_1	-642.4	$\omega - 4$	-640.3
$[C_8mim]^+$	r_1	-800.1	r_1	-800.1	r_1	-641.5	$\omega - 4$	-643.3
$[C_{10}mim]^+$	r_1	-800.8	r_1	-800.8	m	-641.7	ω	-629.5

tive site species as our protein model.

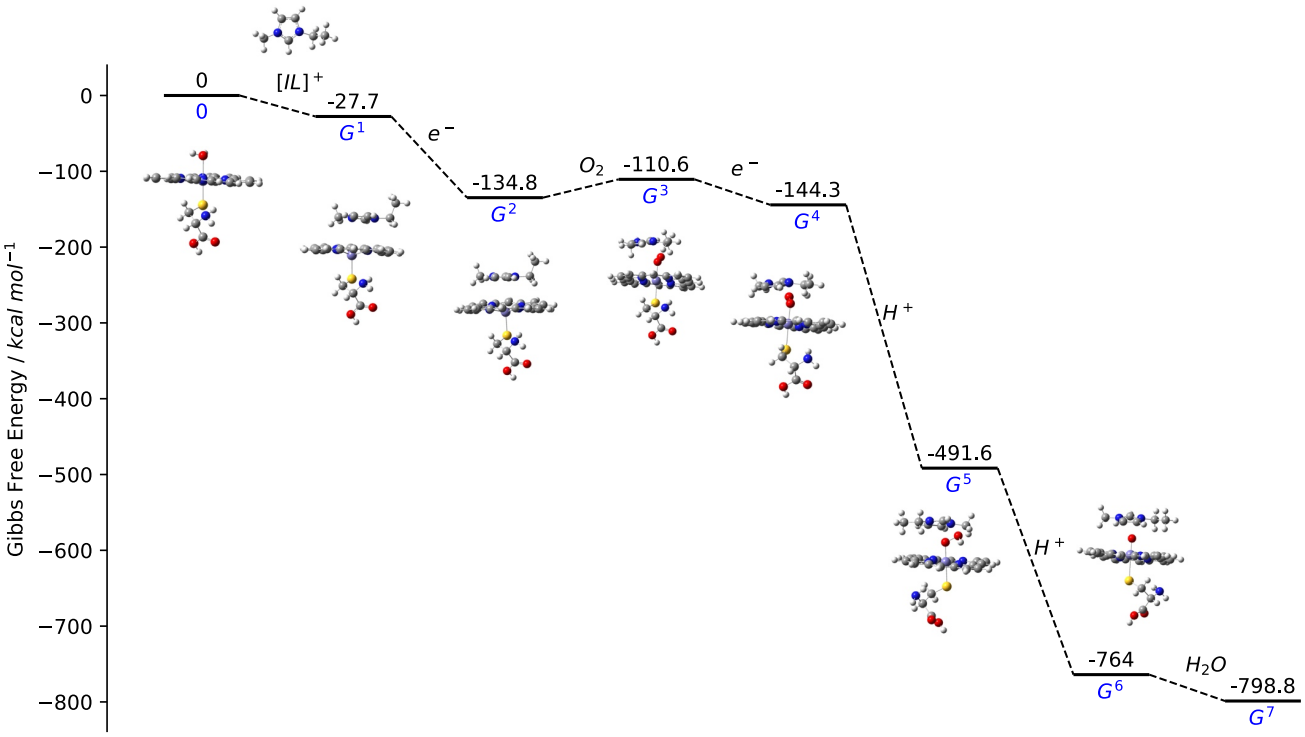


Figure 5.57: Energy profile for $[C_2mim]^+$ tail up complexes optimized in gas phase

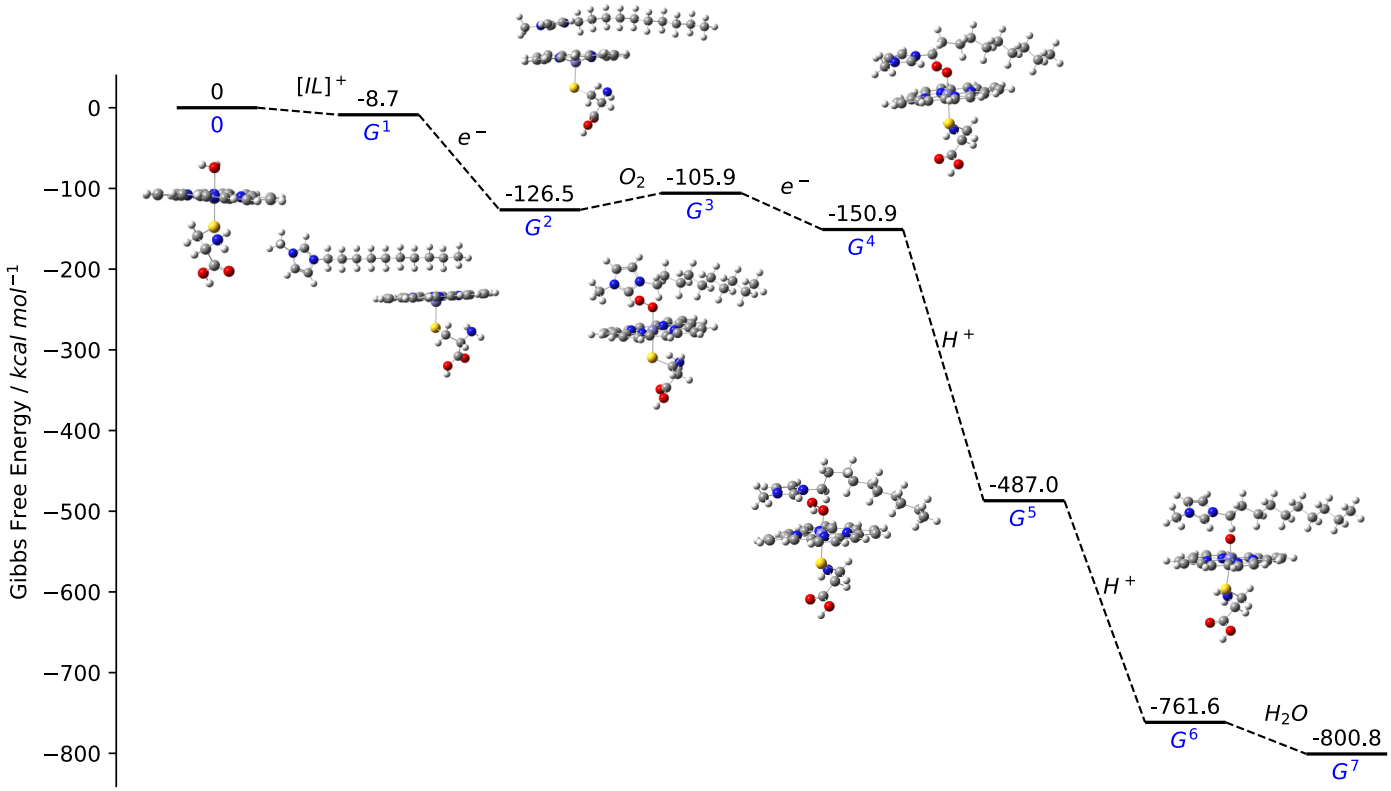


Figure 5.58: Energy profile for $[C_{10}mim]^+$ tail down complexes optimized in gas phase

5.6 Docking and QMMM Models

The docking procedure and methodology has been laid out in the previous sections and here we will discuss some of the key analysis performed for docking. This will be followed by QMMM modeling and structural analysis involving both docking and QMMM calculations. The section will be completed by a brief discussion of inferences obtained from the analysis and implications for the binding of imidazolium-based ionic liquid cations to cytochrome P-450.

The docking was performed on the native crystal structures of BM3 by using two different cubic search volumes bearing the same center. A collection of the utilized receptor biomolecules and the docking energy scores have been tabulated for the reader in the Appendix (Table 8.1). As the search space dictates the placement and fit of the substrate molecule, it was considered important to check for effect, if any, encountered by changing the search volume while docking. Two different search volumes bearing cubic dimensions of 10 and 12 Å were used for analyzing the docking scores and the results have been presented in Figure 5.59. It is easily observed that the change in the search volume does not affect the overall average energy scores for the substrate. Even changing the alkyl chain length on the substrate, the overall energy from all of the docking poses remains within the bounds of 0.3 kcal/mol which is energetically insignificant. From this observation, the further calculations and analyses were conducted on the geometries and docking poses obtained from the dimensions of 10 ÅİIt is important to be noted that Autodock Vina conducts rigid docking and hence the geometrical coordinates of the macromolecule/receptor remain intact during the docking process. These positions are allowed to relax during the QMMM calculations.

Firstly, the docking calculations were conducted on the substrate free 1bvy WT [262] crystal structure devoid of any mutations on key residues. The docking performed in all of the cases is a rigid type where the macromolecule is not perturbed and the substrate is allowed to accommodate itself in the search volume.

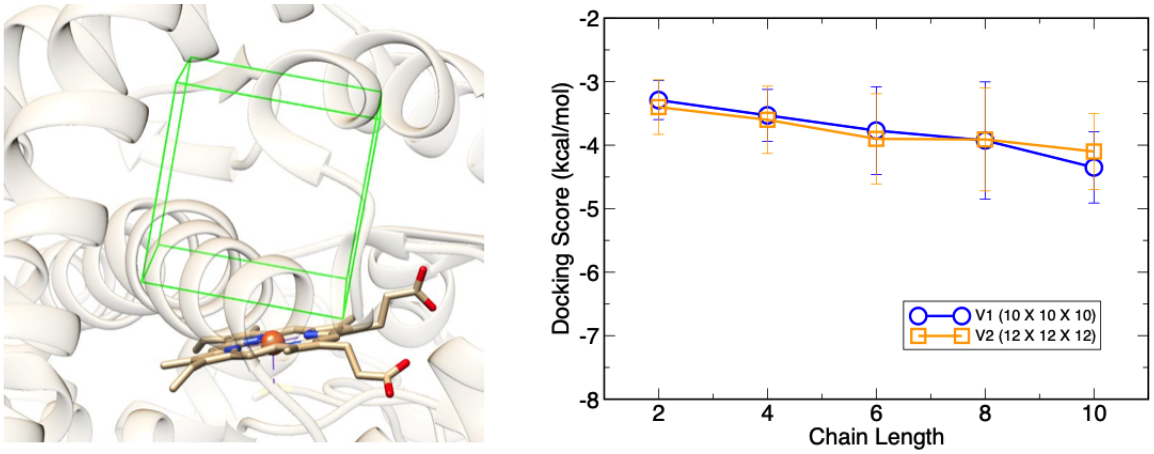


Figure 5.59: Search volume shape (left) and the comparison of docking scores from two different search spaces (right) with ranges in the scores

The average energy of the docking poses obtained from the calculations seem to increase as a function of the alkyl chain length. This can be expressed as the favorable interactions between the hydrophobic alkyl chain of the imidazolium cation and the similarly natures P-450 BM3 binding pocket. From the docking scores, it is also evident that as size of the substrate is increased, the degeneracy of the energy levels is affected. For example, the difference in the energies between the scores of the 1st and 5th most favorable mode in case of $[C_2mim]^+$ is 0.3 kcal/mol while the same quantity is about 2 kcal/mol for the $[C_{10}mim]^+$ molecule. Conformationally, the energy scores can also be connected to the geometric make up of the substrate in which the the smallest cation is shown to adopt very different geometries with docking scores. While for the longer cations, especially, $[C_8mim]^+$ and $[C_{10}mim]^+$, different conformations establish different energetic scores that energetically separated from each other significantly. In case of our cluster calculations, the binding profiles indicate that the difference between conformations in quite significant. The binding pocket is thus able to attenuate that effect due to the inclusion of neighboring residues that exercise physical effects on the substrate. The primary observation from the docking calculations is the ability of cytochrome P-450 to accommodate conformations of a wide variety (with different sites exposed to the active site) within its binding pocket without incurring a significant energetic penalty.

In the next step, the docking calculations were performed in the presence of the water molecules in the binding pocket of the P-450 BM3. The crystallographic water molecules in the structure of 1bvy were retained to reflect the environment in the pocket and all the other waters were removed. A comparison of the docking from both the wild type and the aqueous pocket containing wild type structures has been given in Figure 5.60. All of the possible ranks/poses of the substrate were retained for each of the members of the homologous series. It is clearly observed that the average docking score in the presence of water is less favorable as compared to the wild type. This lowering in the favorability is also reflected in the conformations that are obtained as a result of the docking exercise. The substrate free environment as well as the most favorable poses for all of the cations from $[C_2mim]^+$ to $[C_{10}mim]^+$ have given in the Figure 5.61 as a reference. Interestingly, all of the cations in the series show a tail facing conformation as the most favorable one. This can be further tied to the analysis performed for the gas phase where the tail down conformations were shown to be least stable as compared to others as the alkyl chain length is increased. Adding onto that, even the two shortest chains, ethyl and butyl, containing cations show this property. From our initial hypothesis, the alkyl monooxygenase is expected to oxidize the substrate at the alkyl chain for oxygen insertion which will further promote its breakdown. The tail facing conformations are supportive of this statement because of the fact that the initial docked geometries do not change their regioselectivity after the QMMM treatment. The first step in the catalytic cycle is the removal of the water molecule on the distal site of the heme molecule and the associated change in the spin state of heme.

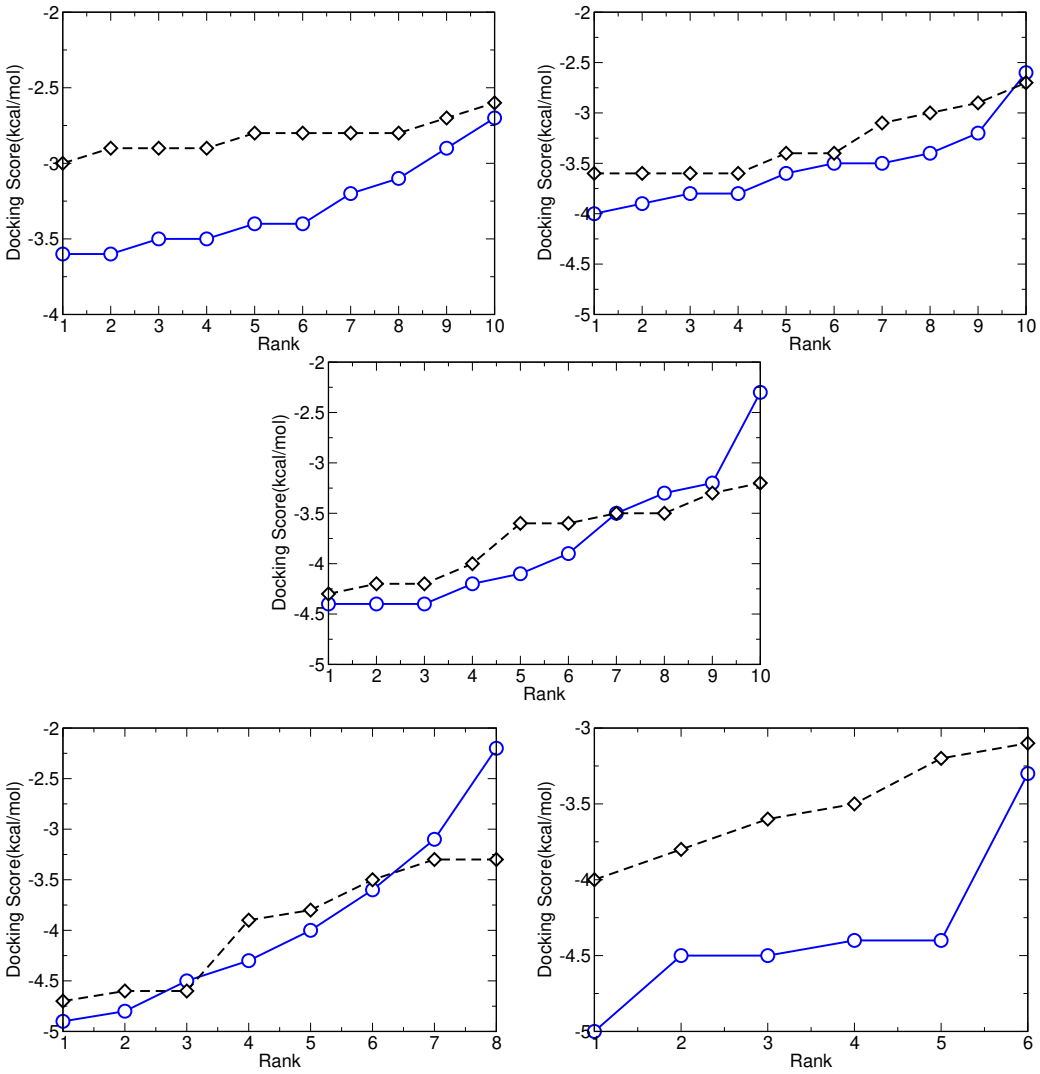


Figure 5.60: Comparison of docking scores from 1bvy WT and 1bvy WT + waters (in pocket) , Legend : blue circles (WT), black diamond (WT + water)

The first QMMM calculations using the ONIOM were carried out on a reduced model as mentioned in the methodology section. The alignment between the docked and QMMM optimized structures for this model have been given for all the cations in Figure 5.63. From the figure, it is evident that structural rearrangement occurs for all of the cations but the site selectivity remains intact for all of the cations except for $[C_2mim]^+$. This is a key observation as binding is the first step in the catalytic cycle. The exposure of key reactive sites along the alkyl chain of the cation is expected to allow for the attack from the P-450 monooxygenase and catalyze the eventual degradation of the imidazolium cation and in turn the ionic liquid. The different structural domains of the P-450 BM3 topology have been depicted in Figure 5.62. These are the collection of the various secondary structures within the crystal structure of the protein chain. While the ones highlighted in yellow are the β- sheets, the blue ones are the α helixes. The motion of these domains determine the structural rearrangement of the protein. The presence of the substrate induces a certain degree of shift in the volume as well as shape of the binding pocket that influences the helices in immediate contact with it. Hence, the inclusion of substrate changes the geometrical makeup of the binding pocket also influencing the secondary structural elements associated with it. The motion of these elements and domains become essential in determining key switches for electron transport and proton shuttling that are central to the activity of the P450 BM3. The next step in the catalytic cycle as described in the methods is the electron transport to the binding domain and addition to the central iron atom. The tilting of the binding domain [239] that is activated by the inclusion of the substrate is central to this action and aids in the electron transport. It reduces the distance of approach for the electron transport significantly.

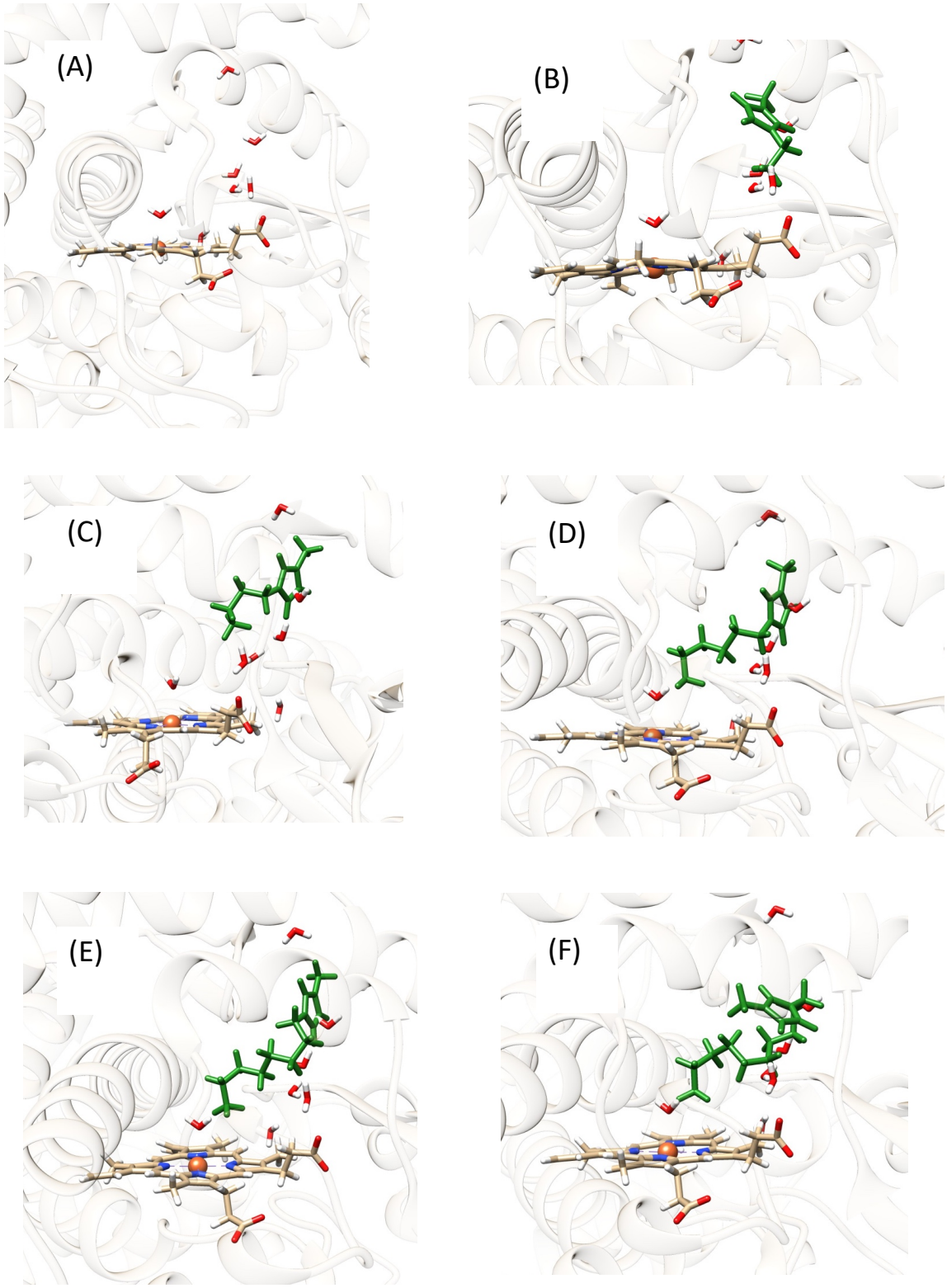


Figure 5.61: (A) The substrate free receptor (1bvy +water(only pocket)) and the most favorable docking poses for (B) $[C_2mim]^+$ (C) $[C_4mim]^+$ (D) $[C_6mim]^+$ (E) $[C_8mim]^+$ (F) $[C_{10}mim]^+$

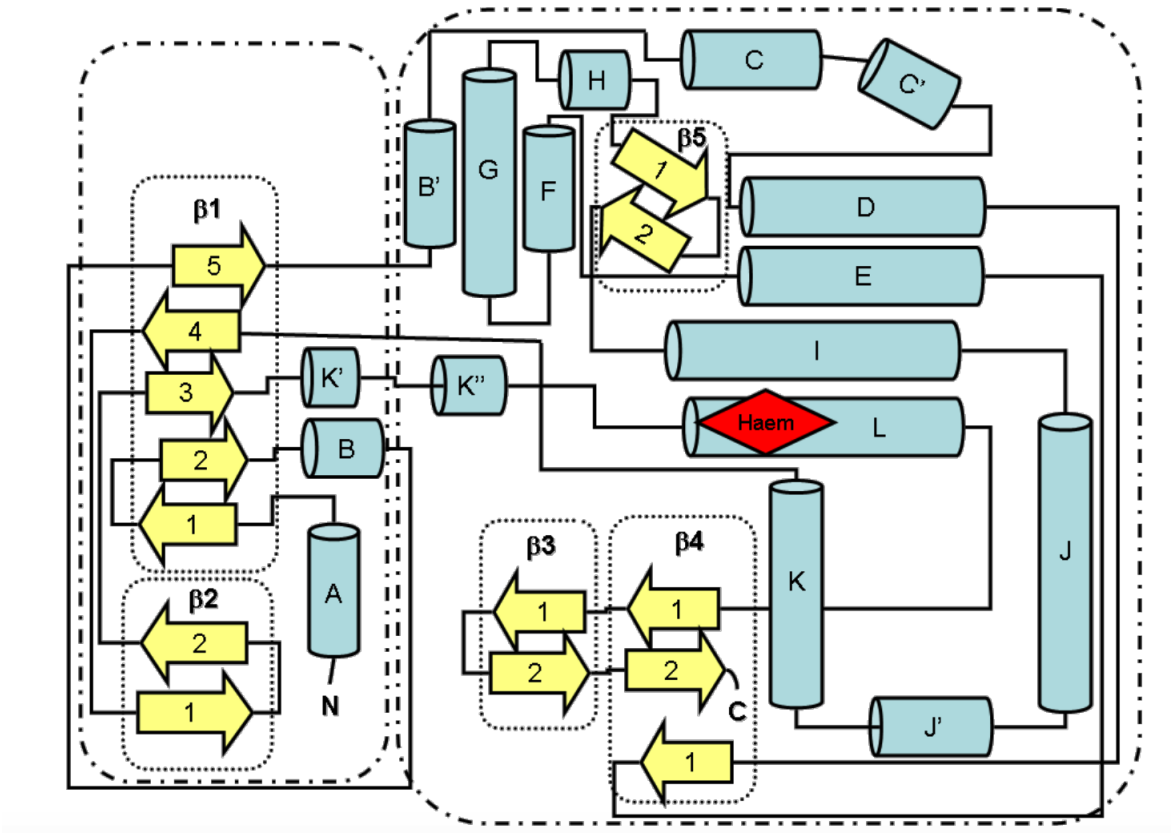


Figure 5.62: Schematic Showing domains of the P-450 BM3 protein (adapted from [12])

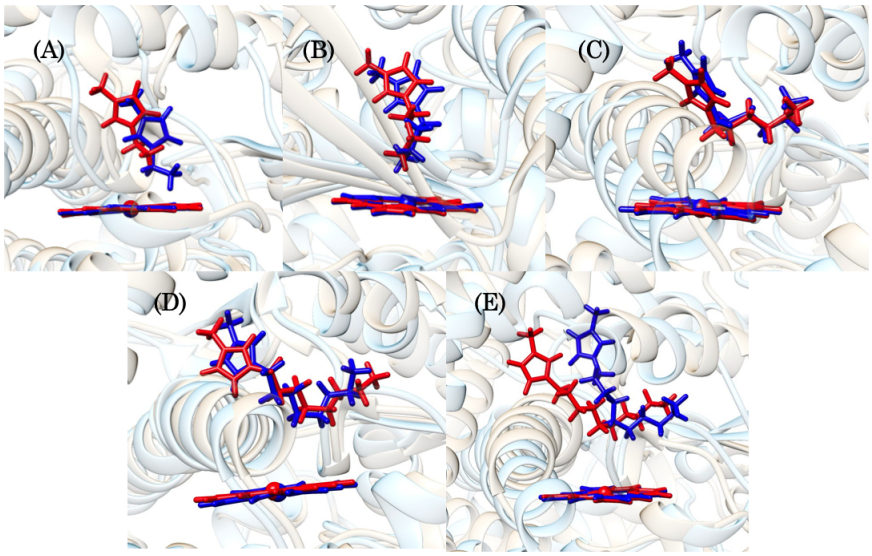


Figure 5.63: Alignment between Docked (blue) and QMMM optimized (red) reduced model for (A) $[C_2mim]^+$ (B) $[C_4mim]^+$ (C) $[C_6mim]^+$ (D) $[C_8mim]^+$ (E) $[C_{10}mim]^+$

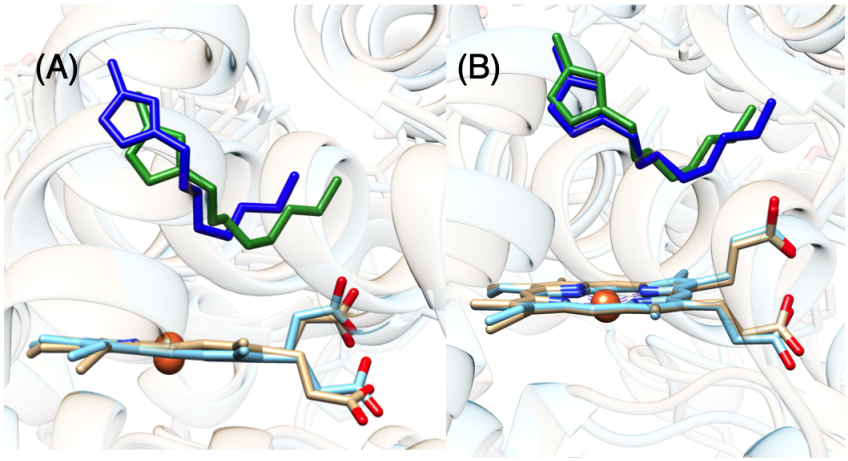


Figure 5.64: Schematic showing the different stages of optimization for the full model (A) Quadratic Macro-optimization and (B) Microiteration Enabled (Color key : Blue (Docked), Green (QMMM Optimized)

In the next step, the full model of the central heme molecule was included into the model region , i.e. including the subsidiary allyl, alkyl and propionate chains. Due to change in the model size, the convergence of the system becomes rather non-trivial due to the interactions between the propionate residues and the nearby atoms surrounding it. In order to overcome this challenge, the optimization of the full model (QMMM2) was performed in two stages. In the first stage, the

opt=quadmac option in Gaussian was chosen to quadratically couple the QM and MM regions and relax the region governed by SCF without performing microiterations for the MM region. In the next step, the microiterations are enabled that allow the relaxation of the MM region for every SCF cycle performed on the model region. The effect of both of these steps in given for a $[C_8mim]^+$ containing system are depicted in Figure 5.64. On the left (A), the quadratic macro-optimization causes a greater displacement in the substrate atoms as compared to the microiterations stage shown on the right (B). The optimization including the microiterations are the final ones that have been used for the evaluation of the molecular properties.

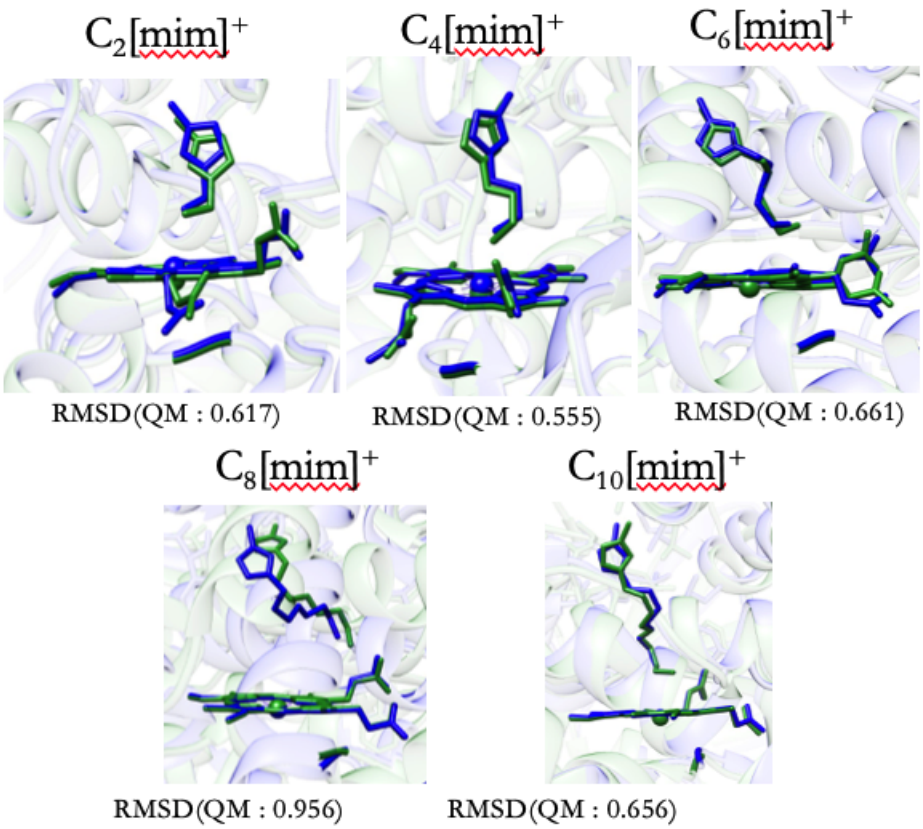


Figure 5.65: Alignment of the most favorable tail down conformations from different chain lengths (2-10) for the $[C_n\text{mim}]^+$ in P-450 BM3 , docked (blue) and optimized (green)

Figure 5.65 gives the alignments between the most favorable tail facing geometries from each member of the homologous series with their docked structures. It is

clearly observed that docked structures dictate the site selectivity and orientation of the cation within the binding pocket. The sites referred to here are with respect to the iron center of the heme molecule representing the ones most favorable for a nucleophilic attack. This is inferred due to the fact that all of the carbon atoms that are shown to be exposed to the Fe atom are maintained after the QMMM optimization. The RMSD between the QM models contained in the structures are given in brackets underneath each of the alignments. Out of the alignments depicted in the work, the $[C_8mim]^+$ has the greatest deviation from the docked structure. The treatment of the most favorable tail facing geometries were of particular interest due to the mechanistic hypothesis laid out for the hydroxylation of the cation and favorability of the conformation amongst the different possible poses obtained from the docking exercise.

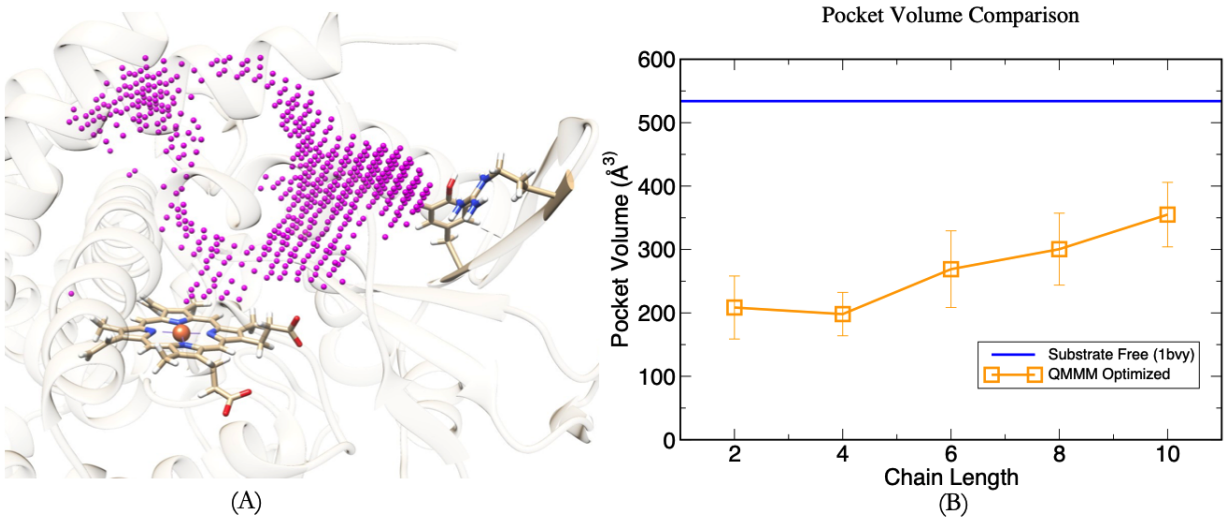


Figure 5.66: (A) Point field used for calculating the pocket volume of substrate free P-450 BM3 and (B) the comparison of bound volumes from different poses

To account for the change in the binding space due to the inclusion of the substrate, the volume of the binding pocket was calculated. In order to perform this calculation, the program POVME [263] was used that uses a distance criterion to construct a field of spheres, given a geometric center an an inclusion radius. Out of the total area under consideration, exclusion areas are also specified in order

to reduce the steric clashes between atoms and to approach the pocket area with sufficient accuracy. The pocket volume of the substrate free 1bvy crystal structure is 534 $Å^3$ that is significantly rearranged upon binding of the substrate. The effect of this inclusion is shown in Figure 5.66 where the substrate free and bound volume are compared. The cation seems to constrict the pocket volume of the protein due to its interaction with the neighboring residues and this is reflected by the volumes reported for different cation sizes. The free volume is the difference between the volume evaluated on the bound structure and the volume reported for the wild type substrate-free structure. Naturally, the inclusion of a bigger substrate molecule is supposed to constrict this free volume further than smaller cation variants. This key conclusion is verified from the calculations of the average volume of the pocket in the optimized crystal structures that range from 200 to about 350 $Å^3$ as the alkyl chain is enhanced from ethyl to decyl progressively by adding two carbon atoms to the alkyl chain at each step.

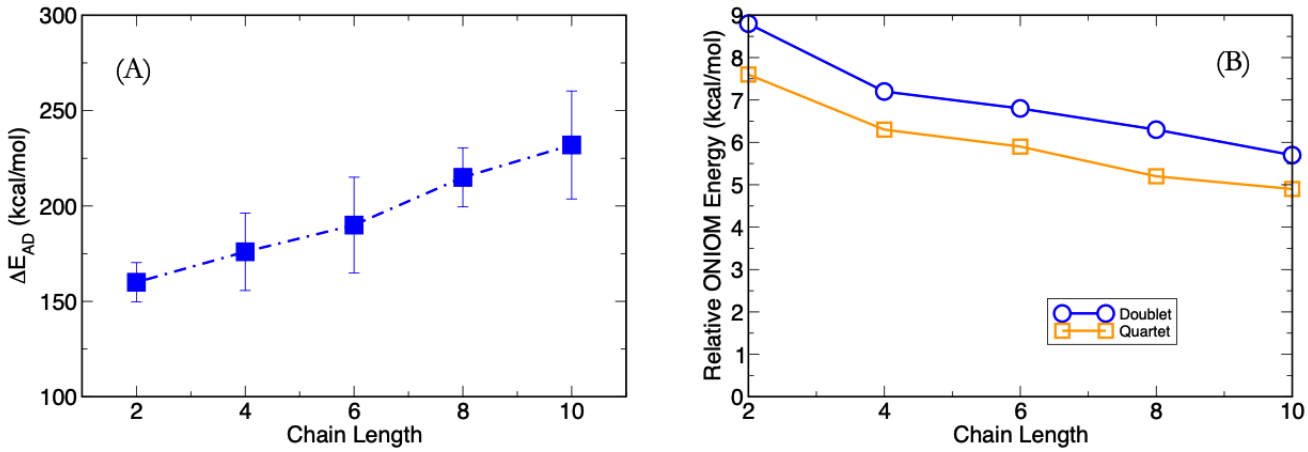


Figure 5.67: (A) Adiabatic stabilization of the QM region (IL cation + Heme +Cys400) in the presence of electrostatic embedding (B) Relative single point QMMM energies with respect to the sextet electronic state of the QM region

In order to the gauge the effect of the protein environment, the adiabatic stabilization of the model region was calculated. This was done by taking the difference of the single point energies between the QM region in the presence and absence of the surrounding protein. The adiabatic stabilization is an indicator of the strength

of the electrostatic embedding provided by the point MM charges to the model region encapsulated within the surrounding protein. From Figure 5.67 (A), it is clear the adiabatic stabilization is a function of the number of atoms in the QM region (heme + IL cation + Cys400) with the protein field. The stabilization seems to be a linearly related to the size of the substrate bound to the pocket. The difference in the energies are higher as compared to those observed for P-450 cam complexes that range from 90-120 kcal/mol [145, 146]. Thus, the adiabatic stabilization is also an indicator of the strength of interaction between the binding pocket and bound substrate. The difference in ΔE_{AD} is about 75 kcal/mol between the averages computed for $[C_{10}mim]^+$ and $[C_2mim]^+$, respectively.

To ascertain the correct spin for the calculations and aligning them with the cluster model explained in the previous section for cysteinated porphyrins, the single point energies of the different structures were computed. These single point energies reflected that the relative separation in energy between the sextet and the other two states (doublet and quartet) is significant. The average separation between the states reduces as the alkyl chain is enhanced by clearly shows that sextet state is the most stable among the ones possible. This is in line with our observations inferred from the cluster calculations in which the resting state molecule, $FePCysH_2O$ undergoes a spin change from doublet to sextet upon the binding of a substrate at the distal end.

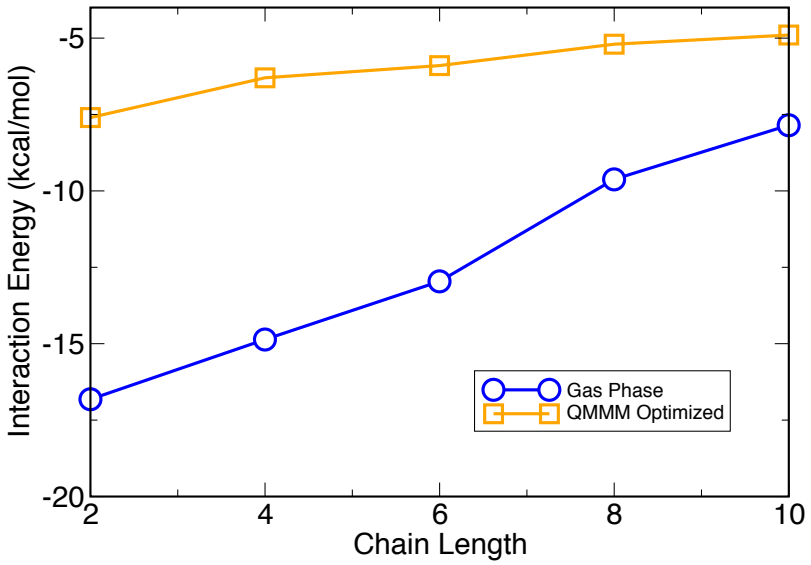


Figure 5.68: Interaction energies from gas phase and QMMM optimized (most favorable docked) tail down binding modes

Finally, the interaction energies for the most favorable tail down poses for the full model were calculated and compared with those obtained from the gas phase FeP containing model. The inclusion of the protein environment makes the interaction weaker between the respective monomers, i.e. the heme and the imidazolium cation. Also, it is noteworthy that the slope for the gas phase calculations is quite drastic as compared to the one seen from QMMM optimization. This is due to the fact that the presence of electrostatic embedding stabilizes the electronic environment around the binding pocket so that interactions between the active site and substrate are reduced.

5.7 Metal Porphyrins

The methodology discussed for the iron porphyrin containing complexes were applied to substituted porphyrins as well. All of the results presented here have been optimized at M06/6-31g(d,p) configuration similar to the FeP containing complexes. Their binding and and reductive abilities were analyzed in the presence of 1-alkyl-3-methyl imidazolium cations. Figure 5.69 gives the binding profile for all of the metals considered in the work. These profiles show the relationship between the binding energies of the two conformations of the cations studied in the presence of the different porphyrins. The geometries arising from the optimization have been given in the Appendix 3 (Figure A48 - A52). The snapshots clearly show the similarity in the binding geometries arising from the optimization and their similarity with the ones obtained from iron porphyrins(Figure A7). The TU geometries clearly have a tendency to align themselves in an interplanar position to maximize $\pi - \pi$ interactions. This is in line with the observations made for iron porphyrin containing complexes. Also the geometries do not show any significant offset in terms of the plane of the imidazolium ring with respect to the center of the underlying porphyrin thus reflecting that electrostatic repulsion is the most dominating physical force governing the geometries of the complex. As with the analysis for binding in the presence of FeP, the relative energetics become important in determining the inherent stability of the complexes and the trend of destabilization along the homologous series.

The relative energetics for the complexes formed due to the different metals follow the one shown by FeP earlier in the work. The TD conformation are mostly relatively destabilized as compared to the TU ones for all of the considered systems as the alkyl chain length is enhanced to hexyl and higher. In all of the analyzes, 2 kcal/mol has been considered to be the datum for significant instability between conformations. For NiP containing complexes, the instability switches from TU to TD as the chain length bound increases from $C_4 - C_6$ although only the C_{10} cation is shown to be significantly unstable as compared to its TU counterpart. The

reason behind this instability is discussed in detail for the FeP complexes, and the same inferences can be drawn for these porphyrins. The ring progressively moves away from the center of the porphyrin leaving the entire complex less stabilized as the chain slithers in front of the underlying macromolecule. In case of zinc, the instability is significant itself from C_6 and higher having relative energies 3.6, 2.2 and 2.3 kcal/mol respectively. Both of the NiP and ZnP are singlet species and the binding energetics are dictated by the geometry presented to the porphyrin. Moving on to the doublet and sextet species, the CoP shows significant relative instability for C_8 and C_{10} TD systems while CuP only yields C_{10} to have the same characteristic. Finally, the manganese containing systems that are sextet, show C_8 and C_{10} TD systems to be significantly unstable as compared to TU species at 2.8 and 2.9 kcal/mol respectively. The key inference drawn from the analysis is that the change in multiplicity of the macromolecule does not interfere with the relative energies and their characteristic across the spectrum of metals considered in the work. While NiP and ZnP are singlet, FeP is triplet and the three others, CoP, CuP and MnP are doublet, doublet and sextet respectively, the electronic configuration does not change the binding characteristics of the overall complex. However, the effect of the number of unpaired electrons present in the system does influence the reductive ability of the porphyrin in the system.

The electrophilicity index of the metals has been described in Figure 5.70. All of the porphyrins' reductive ability has been enhanced due to the presence of the positive charge of the imidazolium cation. The reactivity index reflects that the reductive ability is enhanced by about 80 kcal/mol or more as compared to the cation free state irrespective of the metal present in the system. The change in the electrophilicity index is variable across the different metals considered in the work. Also, conformational effects on the profiles are significant with the indexes for TD conformations being less favorable as the alkyl chain is enhanced beyond butyl systems. Numerically, the electrophilicity index for manganese porphyrins are the highest with the indexes ranging up to even 167 kcal/mol. This shows the effect of the spin state on the porphyrin molecule as the manganese porphyrins

are sextet while the other systems have either 1 or 2 unpaired electrons.

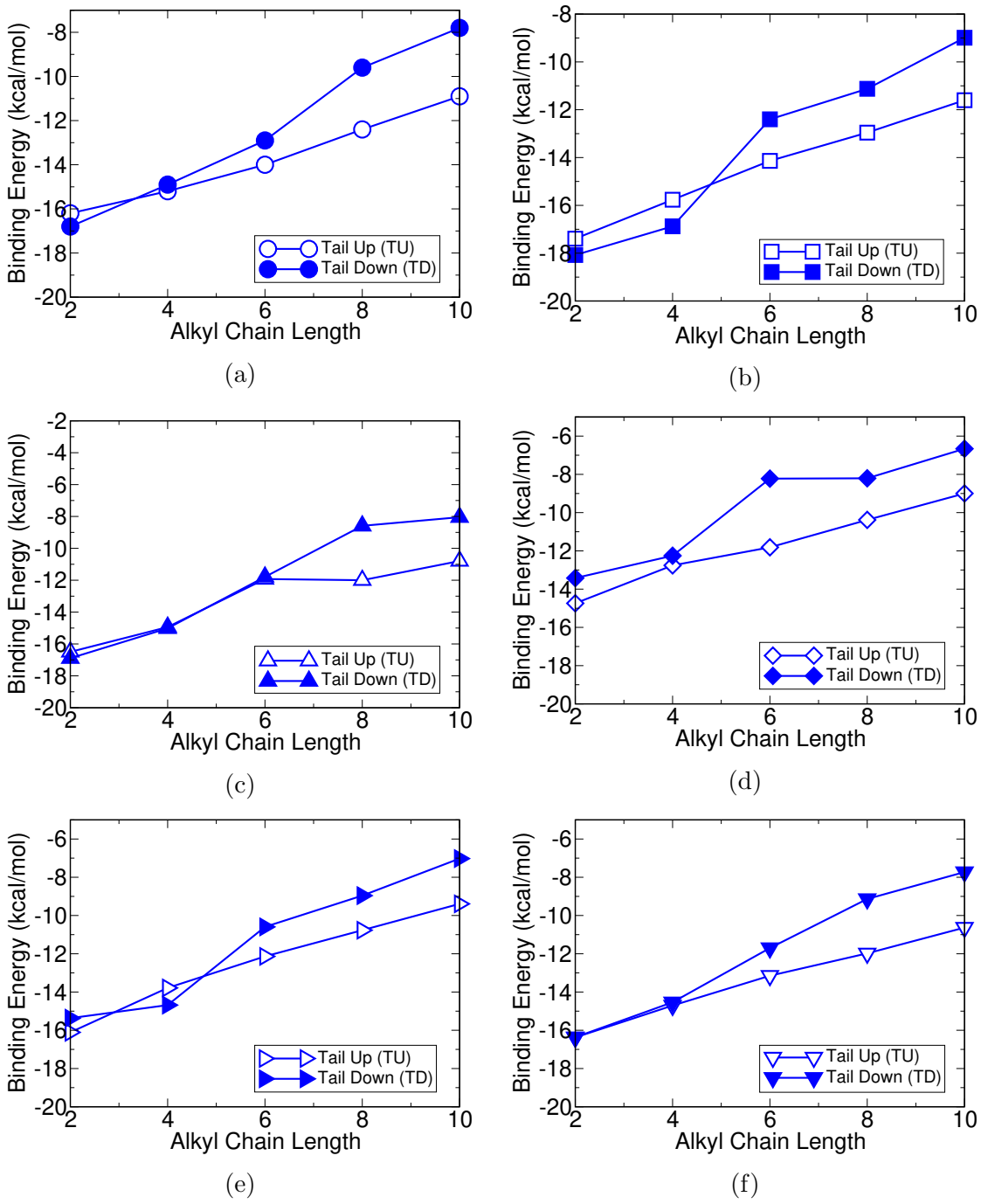


Figure 5.69: Binding Energy profiles for Metal Porphyrin Cation complexes (MP[$C_n mim$]$^+$) M is (a) Fe (b) Ni (c) Co (d) Zn (e) Cu (f) Mn

To account for the donor acceptor interactions between porphyrins and the cation, the intermolecular stabilization was evaluated using NBO analysis. The NBO stabilization was calculated by summing all the intermolecular contributions

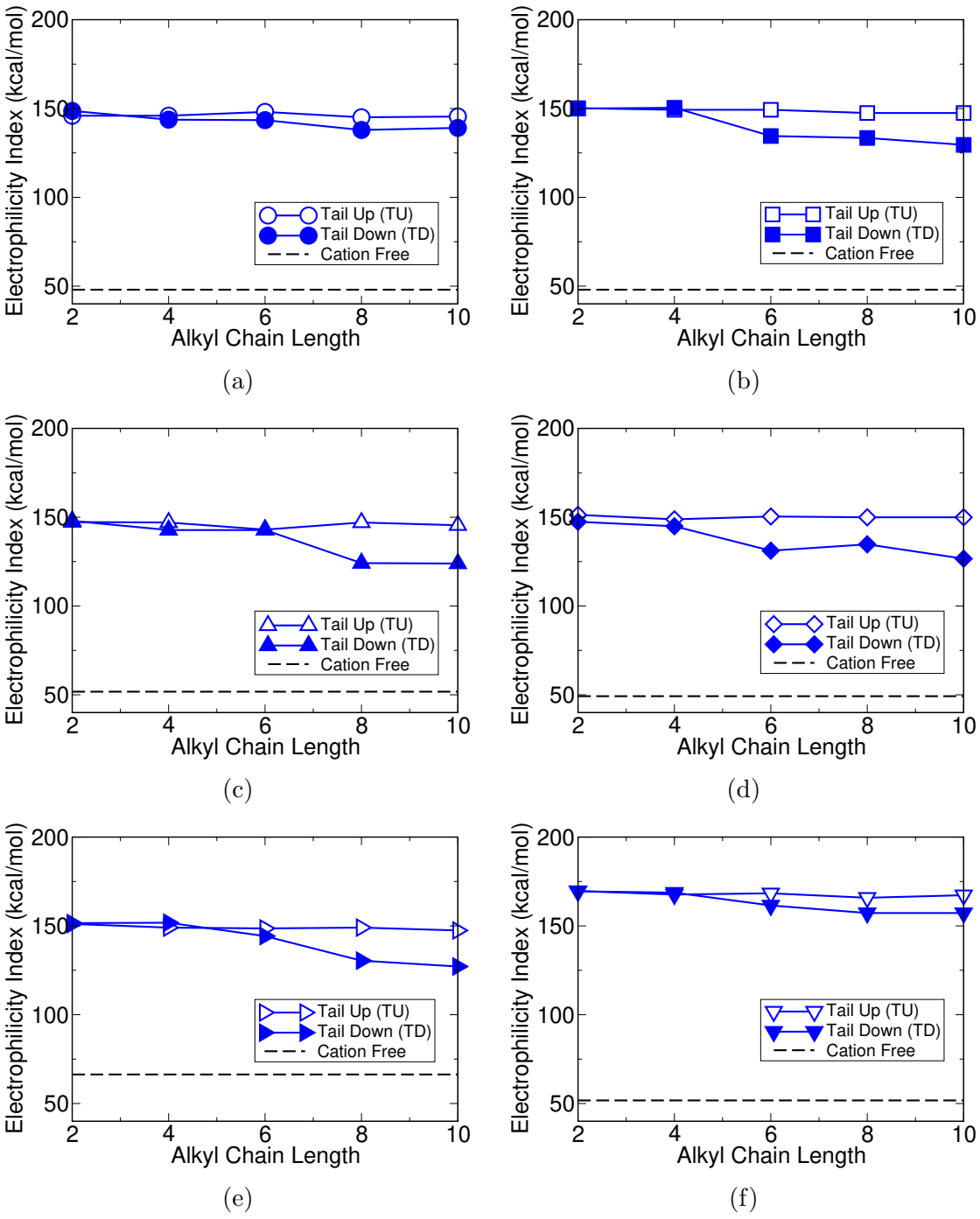


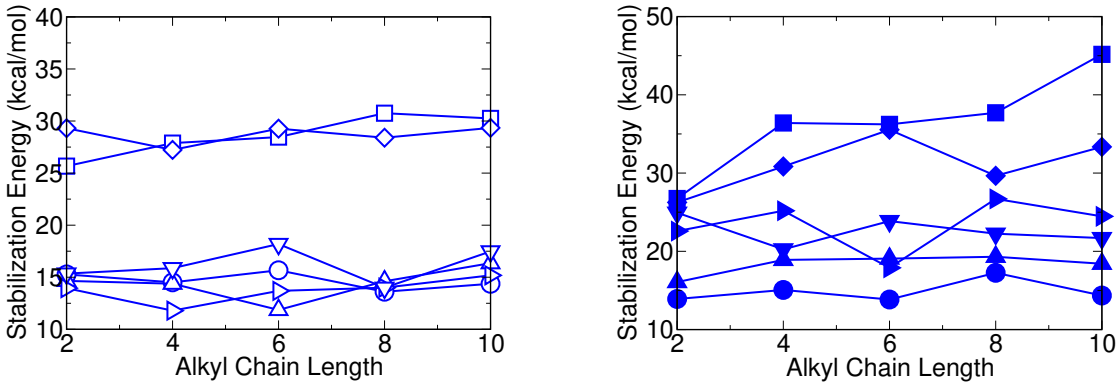
Figure 5.70: Electrophilicity Indexes for Metal Porphyrin Cation complexes (MP[$C_n mim$]$^+$) M is (a) Fe (b) Ni (c) Co (d) Zn (e) Cu (f) Mn

from the cation and porphyrins towards each other. The individual contributions summed from the three units as laid out in the NBO program, i.e. the cation, the porphyrin and the metal atom were added in the intermolecular framework.

The resulting stabilization gives an indication of the perturbation to the natural Lewis wavefunction . Upon comparing the stabilizations for individual conformations for each metal, some interesting trends emerge that reflect the orbital exchange. In case of FeP, there does not seem to be a clear preference for a certain conformation to show a higher degree of stabilization. Thus, the donor acceptor interactions are equivalent for most of the conformations. The only difference is seen for the C_8 series where the TD is shown to have higher stabilization as compared to the TU conformation by about 3 kcal/mol. Moving on to Nickel and other metals, the distinction between the two conformations for preference in stabilization is clear. For NiP containing complexes, except for the shortest cation, $[C_2mim]^+$, all the other members show that stabilization for TD conformations is energetically more significant.

The reason behind this preference can be explained by the fact that the alkyl chain participates in more donor-acceptor interactions with the underlying porphyrin atoms than the imdazolium ring. The interaction between NBO's is dictated by the distance between the atoms as well as prospective electronic configuration of the system. For TU conformations, the majority of the interactions is between the ring and the metal porphyrins while for TD conformations, the interaction regime is dominated by the alkyl chain. ZnP also shows a similar profile as NiP, with the only exception to TD conformations showing more stabilization is the ethyl cation in complex. For systems containing odd number of unpaired electrons or having an even multiplicity, the trend seems to be clearly favorable for the TD conformations without any exceptions. For all the three remaining metals, i.e. Co, Mn and Cu, the second order perturbation analysis shows that participation of donor acceptor interactions between the alkyl chain and porphyrin is dominant as compared to the imidazolium ring. A combined treatment of the total stabilization for TU and TD conformations is given in Figure 5.71(a-b). It compares the same overall stabilization for all the metal complexes to gauge the effect of the presence of different metals in the complex. Interestingly, the metals bearing a singlet state show greater stabilization as compared to the ones that con-

sist of unpaired d-electrons. Nickel and Zinc containing porphyrins clearly show this trend for the entire homologous series for TU conformations. Thus, conformationally the presence of unpaired electron advocates the NBO exchange between the chain and porphyrin to a greater extent while the overall stabilization is still greater for spin neutral systems. The TD conformations also show the same trend where the greatest stabilization among all of the metal porphyrins are shown by Nickel and Zinc while all others are below their level. The only exception to this is ethyl containing system in which the Manganese porphyrin is shown to have similar perturbation corrections.



(a) (b)

Figure 5.71: Total NBO stabilizations for all of the considered metals for (a) TU (b) TD conformations; Legend - circle (Fe), square (Ni), upper triangle (Co), diamond (Zn), right triangle (Cu), down triangle (Mn) ; open symbols (TU) , closed symbols (TD)

The presence of imidazolium cations in the porphyrin systems induces a marked non-planarity in the latter. This non-planarity can be understood on the basis of the underlying vibrational modes that occur for all of these porphyrins. These non-planar vibrational motions have been represented for porphyrin models in Figure A78 with their associated displacement vectors for individual atoms. Out of the possible motions, the ones that have been found to be most significant have been tabulated for all of the metal porphyrins in isolation (in absence of cation) in Tables 8.2- 8.7. The most important motions are namely, doming, ruffling, waving and saddling that occur at distinct frequencies for each of the different

metal porphyrins. Out of these, the doming frequencies are associated with the greatest IR intensity while others show zero intensity in the absence of the cations in complex. Other notable observations are the presence of two complementary waving motions that occur at the same frequency for all of the macrocycles. Upon analyzing the Fukui functions, it is also noticed that the local electrophilicity and nucleophilicity of the metal atoms is uniform across the two conformations.

CHAPTER VI

CONCLUSION

6.1 Conclusions

The present work is an attempt to understand the molecular level insights into the binding, and then later, into the associated chemical changes in cytochrome P-450 and ionic liquid cations. It was conceptualized based on the hypothesis related to the biodegradability of ionic liquid cations that "a suitable alkyl monooxygenase is expected to introduce oxygen-rich moiety into the alkyl chain of the ionic liquid cation". This oxidation would cause the eventual breakdown of the cationic part of the ionic liquid, which is directly related to the overall degradability of the fluid. Although it is understandable that biodegradation is a very sophisticated process to model using one computational technique, the work presented in this thesis is an attempt to develop a basic computational framework to be followed by research efforts at higher scales. Under the guidelines laid out for Green Chemistry, identification of structural effects on substrates in contact with biological macromolecule often forms the backbone of the detailed design process. It is deemed essential for two reasons, firstly because structure and reactivity are intertwined and determination of geometrical factors would undeniably aid in understanding chemical reactivity. The second reason being that modeling effect of conformation forms the basis of modeling at further levels of complexity such as molecular dynamics, computational fluid dynamics, kinetic reaction analysis, pilot level synthesis and

industrial realization, in that order.

Out of the different forms of oxidation possible, the monooxygenase P-450 is capable of a plethora of oxygen insertion options and mechanistically is becomes challenging to pin point a certain type of reaction. In this vein, the metabolic insight provided by experimentalists becomes important as the first step is shown to be hydroxylation in an aerobic environment. This hydroxylation has also been further studied based on mass spectrometry to give rise to fragments indicating terminal nucleophilic attack and subsequent alcohol formation. Thus, out working hypothesis being hydroxylation of the ionic liquid cation being possible in the presence of cytochrome P-450 and the geometrical factors dictating this process would enable site specificity. The site most conducive for the hypothesis to be verified is the terminal position.

Before explaining the inferences drawn from the different modeling stages and the electronic structure calculations performed, it is important to understand the motivation behind the attempt. Cytochrome P-450, although being widely regarded and studied as multifaceted enzyme capable of performing a wide range of reactions, has not been tested for ionic liquids. This poses a major logistical challenge as no known crystallographic structure exists for the ionic liquid within the P450 binding pocket. Thus, the determination of feasible binding mechanism and conformational effects have to be associated with *ab initio* studies. The absence of any known PDB structure as mentioned above also indicates that the model to be developed needs to be inclusive of all the representative effects from a realistic enzymatic environment. The necessary structural effects that are incurred from the secondary structural elements play a vital role in driving the function of the enzyme but the central part of the enzyme is the active site. The active site is the heart of any biomolecular assembly and no model of substrate association can be formulated without considering the active site. In the present work, as the title suggests, the efforts have been dedicated towards developing an understanding of the active site models at various levels.

The problem was formulated by beginning the work to include a very reductionist approach to the active site model of cytochrome P-450 BM3 with the ionic liquid

cation. The choice of the cation over the anion was made because of its greater impact on the biodegradability of the overall ionic liquid. In literature, it has been expressed that anionic effects are minimal and geometrical insight into the cation is directly reflected in their biodegradability levels in various environments. Also, the effect of the cation has been clearly shown to be indicative of the degradation potential in the presence of a wide variety of anions. The relative sensitivity to the anionic part of the ionic liquid is not as pronounced and the anionic part is expected to remain in solution while the cation plays the major role in chemical reactivity.

In the first part of the work, a model representative of the association between cytochrome P-450 active site and an imidazolium-based ionic liquid cation was developed. As inferred from literature, the active site of cytochrome is the heme consisting of a porphyrin macromolecule at its center. To accommodate the effects of the porphyrin as well as the metal center of the porphyrin, electronic structure calculations were conducted on both free base porphyrin (FBP) and iron porphyrin (FeP). To include the conformational effects of the cation, three distinct initial conformations were considered for the initial model, namely, tail up, tail down and interplanar. Out of these, the binding energies as well the relative energetics show that the tail down conformations become least stable as the size of the alkyl chain length is enhanced from 4-6 carbon atoms in the cation. Our initial treatment with the pure B3LYP DFT functional yielded geometries that gave highly unfavorable binding energies for both FBP and FeP containing models. The B3LYP cluster model provided insight into the placement of the energy states populated as the cation comes in contact with the porphyrin molecule. The frontier energy orbitals, HOMO and LUMO of the cations as well as the porphyrin populate distinct levels on the energy scale in the complex. The effect of the porphyrin on the cation frontier orbitals was also tested and the effect on the LUMO was far more pronounced as compared to the HOMO. This effect was also related to the geometries presented to the underlying porphyrin. The placement of the LUMO and the HOMO orbitals of the cation in isolation were shown to be on the

ring and shifting to the alkyl chain respectively. As the alkyl chain was enhanced, the HOMO was distributed along the chain. The exposure of these parts of the cation to the porphyrin thus, dictated the change in the HOMO and LUMO energies from their isolated values in the gas phase. A charge distribution analysis also uncovered that the exposure of the alkyl chain to the porphyrin is able to effect maximum retention of the electronic density on the protein as a result of least charge donation to the cation. The tail down conformation showed the least charge transfer while the ones from tail up and interplanar conformations were comparable.

Upon addition of the dispersion effect to our initial cluster model, the geometries and binding energies were significantly different. As compared to the pure B3LYP, all of the geometries obtained from the dispersion enabled B3LYP-D2 level of theory yielded favorable binding energies. Thus, dispersion was shown to play a very important role in estimating the interaction between the ionic liquid cation and the underlying porphyrins. Geometrically, the D2 correction has its most pronounced effect on the tail down conformations, in which chain slithering was observed. The progressive slithering was also noticed as the size of the cation was enhanced. From a computational point of view, the results from the empirically corrected B3LYP-D2 were also tested with the parametrized M06 functional. The key conclusion was that although absolute energetics from the two theories were quite different, the relative energetics between the conformations were captured by both of them. From the results, it was clear that tail facing geometries would yield the most unstable conformations as the size of the chain is increased. Upon decomposing the binding energies, it was also inferred that monomer deformation and counterpoise correction energies do not play an important role in dictating the overall binding energies. The energy of association between the monomers is dependent on the interaction energies determined at the equilibrium geometry. To gauge for the reductive ability of the porphyrin in the presence of the cation, an electrophilicity index from conceptual DFT was utilized and computed. Our calculations indicated that the presence of the positive charge of the cation raises the index significantly irrespective of the conformation presented to porphyrin.

The conformational effect of the tail down geometries is also reflected from the electrophilicity indexes computed for porphyrins in the complex.

Owing to the unstable nature of the tail down conformations, the next part of the work was dedicated towards understanding the nature of them for different sets of cations. For this part, three aromatic and one non-aromatic cation was considered and their binding energies were computed in the presence of FeP as in the previous model. Geometrically, the tail down conformations were shown to be more unstable for imidazolium and pyridinium based cations but the trend was reversed for pyrrolidinium and thiazolium containing systems. In fact, for pyrrolidinium, with the exception of the C_6 cation, all other relative energies were within quantum chemical accuracy to establish same energetic level for binding. Upon comparing the binding energies arising from the tail down conformations from all the four classes of cations, they were found to bind within an energetic window of 2 kcal/mol once the alkyl chain length was sufficiently large. This highlighted the fact that as the ring was progressively displaced from the center of the porphyrin and in the absence of any extra bulky groups, the binding of the cation would occur at the same level irrespective of the nature of the ring head. A natural bond orbital analysis to understand the donor-acceptor based interactions between the cation and the porphyrin showed that the atoms participating in these interactions were dependent on the conformations. While for the tail up conformations, most of the interaction signatures were shown for the ring atoms, the tail down conformations showed more dominant participation from chain alkyl atoms. The extent of the stabilization was also dictated by the nature of cation as well as the conformation exposed to FeP. Out of the four cation groups studied, thiazolium tail up complexes gave a unique T-shaped complex that highlighted the effect of the directional lone pair on sulphur and the propensity of $\sigma - \pi$ interactions rather than dominant electrostatic repulsion arising from a $\pi - pi$ stacking.

The model developed for the FeP containing system was further extended by including elements of the crystal structure of P-450 BM3. To this end, the prox-

imal end of the FeP was extended to include a deprotonated cysteine residue as a representative of the Cys400 residue in the PDB structure. This model was then utilized to perform calculations related to the catalytic cycle of cytochrome P-450 in a synchronous manner. The synchronous approach was undertaken in order to preserve the conformational effects that were developed across the different steps leading to the final hydroxylation step of the bound cation. From our DFT treatment of the catalytic cycle, it is evident that all of the steps leading to and including the hydroxylation are thermodynamically favorable, with the exception of the dioxygen addition. In literature, this dioxygen binding has been shown to be kinetically driven rather than being thermodynamically induced and similar calculations for camphor suggested a positive free energy. In the substrate free or cation free state, the individual steps in the cycle were also evaluated and it was observed that in addition to the dioxygen addition, the second reduction step also is thermodynamically favorable as has also been observed in literature. According to our calculations, the protonation steps leading the formation of the ultimate active species are the most energetically favorable steps followed by the first reduction, second reduction, hydroxylation, binding and then dioxygen addition all of which are evaluated in the presence of the imidazolium cation. Out of the steps evaluated, most of them apart from the exception of the protonations, do not show any significant contribution to the free energy from the entropy. Geometrically, all of the gas phase geometries yield hydroxylation sites on the ring rather than being on the chain. Although, our calculations on the embedded media suggest that terminal and sub terminal hydroxylation is feasible for the two greatest cations in complexation with the cysteinated porphyrin.

In the final part of the model containing the P450 BM3 protein with the cation, docking and QMMM calculations were conducted on the system. Our docking calculations suggested that the binding pocket of the P450 BM3 induces a significant conformational freedom to the ionic liquid cation. This was proved by observing the docking scores obtained from different conformations and their relative closeness in energy. A wide range of conformations were obtained within a chemical

accuracy of about 1kcal/mol suggesting that the P450 binding pocket is capable of binding the imidazolium based cation in a variety of poses without significant energetic penalty. Upon performing QMMM calculations, it is also interpreted that the docked structures serve as appreciable guesses to them. The site selectivity of the docked structures are retained after the optimization is completed. The spin states as observed from the cluster model were also echoed by the QMMM model including the protein environment. Adiabatic stabilization and electrostatic embedding also showed that the model region or the active site is significantly stabilized by the point MM charges surrounding it and the stabilization is linearly dependent on the size of the cation in the binding pocket. The interaction energies evaluated for the protein also show that they are thermodynamically feasible but significantly less favorable as compared to the gas phase counterparts.

An additional part of the work consisted of the substitution of the central metal of the porphyrin with other metals in the same row as iron. For this, five other metals other than Fe were considered namely, Nickel, Zinc, Cobalt, Copper and Manganese. The binding profile of 1-alkyl-3-methylimidazolium based cations were studied from all of these porphyrins and conclusion regarding to the stability of the complexes were drawn. Similar to the observations made for the FeP containing cluster model, the tail down conformations were shown to be more unstable as the alkyl chain length on the cation was enhanced. The reductive ability of these porphyrin models were also raised significantly as compared to the cation free states. A vibrational analysis also showed that the non-planar modes of the porphyrin leading to its deformation are accentuated by the cation and was observed in the complexes.

6.2 Future Directions

The present work has been dedicated towards an exhaustive exploration of a reduced active site model of the ionic liquid cations with respect to metallic and initially non-metallic porphyrins. As such, the modeling at high scales such as QMMM and molecular dynamics need to be performed for these systems in order to gain further insight into binding and other process. These calculations would allow us to understand not only the structural implications but also inherent dynamics governing the entry and exit of the substrate and other small molecules within the binding pocket.

Of particular interest is the steered molecular dynamics approach that presents a free energy path for the entry of the substrate into the protein entry channel. This would be crucial in determining the tentative approach of the substrate into the binding pocket and upon comparison with the cluster model, provide a unified approach for determining the accurate site selectivity of the ionic liquid cation within cytochrome P-450.

Another point of interest is the reaction mechanism of the ionic liquid cation in the presence of the ultimate active species leading to its hydroxylation. In literature, the evidence has been presented for a set of different mechanisms and intermediate species. These mechanisms need to be individually tested for the ionic liquid cations as they have not been studied in the same light. Similarly, experimental collaboration to further elucidate the molecular level details presented in the work would be useful in validation as well progressing towards a unified biodegradability framework. The computational framework laid out here is the first step in understanding the association of the protein molecule with the non-natural cationic substrate and further studies, particularly, crystallographic ones, would be very crucial in aiding conformational insight and hybrid modeling such as QMMM.

6.3 Experimental Techniques

The present work is a strictly computational effort involving electronic structure calculations. The validation of the calculations performed for each of the models would require experimental work related to characterization and evaluation of thermodynamic quantities. In the interest of the same, some of the possible experimental techniques have been outlined to aid in the understanding of the research work.

6.3.1 Equilibrium Titrations

The association of the ionic liquid cation is reflected in the ligand affinity of P-450 BM3 and is one of the most important quantities that needs to be captured for successful prediction of its binding. A dissociation constant reflecting this affinity (K_s) can be potentially determined by plotting the absorbance changes while titrating P450 with a substrate vs ligand concentration. The fit obtained from such a process would be hyperbolic or quadratically nonlinear depending on the strength of binding. The intensity of a light source may also change with an increase in the ligand concentration. The absorbance can be measured by considering a series of solutions having a constant molarity but varying protein to ligand ratios. Dilution often promotes the light path presented to the sample.

6.3.2 Ligand Binding Kinetics

The P450 catalytic cycle is a series of reactive steps leading to the final hydroxylation of the substrate and product release. These individual reaction steps can be resolved by stopped-flow spectrophotometry and recovering the single or multi-wavelength kinetic data obtained after rapid mixing of the respective protein and ligand mixtures. Also, the step encountered after the substrate binding is the reduction of heme via a suitable redox partner. A marker such as carbon monoxide may be added to the mixture and the reduction can be monitored by appropriately

following the adduct formation [264, 265]. In another similar technique, first the CO-Fe bond is disrupted by a laser flash and the rebinding of CO is monitored. The rebinding kinetics reflect the rate of diffusion of CO through the protein indicative of the protein conformation and binding mode of the substrate. The CO binding rate in this case is directly proportional to the ligand access channel that is influenced by the conformational modulation of protein by the substrate.

6.3.3 Electron Paramagnetic Resonance (EPR) Spectroscopy

All active site species associated with the P-450 consists of either unpaired electrons or are radical in nature. Hence, EPR spectroscopy is one of the most effective techniques that could be applied to the present research problem to predict the oxidation states of iron in various species. A set of unpaired electrons leading to a certain oxidation state can contribute to either loss or gain in momentum, which is associated with the response of a species to a magnetic field. Each of the unpaired electrons translates to a unique line in the overall EPR spectrum. The spin state equilibrium induced by the presence of substrates can be measured using this technique.

6.3.4 Electrochemistry

Ferric substrate bound P450 species receives two electrons from its redox partner during the catalytic cycle. The addition of these electrons in crucial in promoting the functioning of the cycle. Presence of substrates have shown to enhance the rate of electron addition, which is also reflected in our calculations. The act of substrate binding leads to a large positive shift in the mid point potential of the active site species. The free energy of the reaction in the medium can be obtained from measurements of the redox potential. The semi-classical Marcus equation is given by

$$k \propto exp(\frac{-(\lambda + \Delta G)^2}{4\lambda k_B T}) \quad \text{(VI..1)}$$

where k is the rate electron addition, λ is the reorganization energy, ΔG is the free energy, k_B is the Boltzmann's constant and T is the temperature [266]. In the equation, the reorganization energy is shown to be significantly greater than the associated free energy of the process. This energy is enhanced in the presence of substrate that reduces the coordination number of Fe from 6 to 5. Both the oxidation states, i.e. reduced and non-reduced in the presence of the substrate are pentacoordinated that enhance the turnover rate as compared to the substrate-free state of the active site species.

Electrochemical studies in the case of P450 would be potentially applied to determine the redox potential or electron transfer processes. The interaction for the analysis is shaped to occur between the P450 and various electrodes, directly or indirectly. To electrochemically drive the P450 cycle, a suitable electrode is employed as the redox partner. A voltammogram consists of two peaks attributed to the redox pair of the Fe^{3+}/Fe^{2+} states. The heme reduction is influenced when the substrate is present and an increase in the cathodic peak current is observed [267, 268]. A similar observation can be drawn in the presence of oxygen, when the substrate is oxidized. The surface concentration of the P450 enzyme can be determined by the calculation of the total charge transferred to the P450. This is simply the integration of the reduction peak obtained under anaerobic conditions.

6.3.5 X-ray Crystallography

In order to determine the appropriate details for the association of the cation and the enzyme, X-ray crystallography is expected to provide insight into the interactions between the substrate and the P450 structure. For understanding each step of the cycle, several intermediate crystal structures need to determined to account for the conformational effects upon substrate inclusion and other chemical changes.

The substrate or binding partner of the protein needs to be co-crystallized in order to obtain the final structures and the co-crystallization efforts are dependent on the knowledge of kinetic parameters [262, 269]. This is needed to determine the concentration of the ligand to be added to the protein solution. A protein binding one ligand often consists of a bilobal organization via a suitable linker region.

CHAPTER VII

CALCULATIONS AND PROTOCOLS

In this section, some of the key aspects of the calculations, their setup and pitfalls will be discussed. It has been organized into sections related to each type of calculations and the authors' suggestions on them.

7.1 Geometry Optimization

The geometry optimization is the most important calculations laid out in the present work. The initial geometries have been given in the methodology section for reference. The choice of these geometries were based on the fact that they are chemically distinct form each other and the prospective interactions arising from them would be different. While for the TU complexes, the ring is expected to interact with the underlying porphyrin, the TD conformations present the tail to the active site. Finally, the interplanar complexes were constructed to reflect the possible $\pi - \pi$ interactions between the plane of imidazolium and porphyrin. It is important to note two separate aspects related to these calculations at the first stage. The first being, each of these conformations yield a distinct path for the optimization. This path becomes essential to determine local minima and discern between optimized and un-optimized geometries.

In the initial stage with B3LYP, this primitive theory is not able to capture the dispersion interaction between the two monomers and as such yields bind-

ing energies that are significantly less favorable. Upon the inclusion of dispersion corrections, both B3LYP-D2 and M06 theories are able to capture the dispersive attraction between the the molecules and give a realistic picture of the binding. For the TU complexes, the hydrogen pointing towards the Fe belongs to the methyl groups attached to the ring. During the course of optimization, the ring aligns itself with the plane of the porphyrin and approaches an interplanar shape. On the other hand, the TD conformations show that the alkyl chain slithers and tries to attain a parallel makeup with respect to the plane of the underlying porphyrin. These optimization paths are essential in determining if a geometry is stuck at a local minima and restarting the job from the appropriate end point from the initial calculation. The optimization path is dictated by the favorable physical interactions between the molecules participating in the optimization.

The second aspect related to this calculation is the determination of the distance of approach. This distance of approach is kept at 3 Åfor all the three conformations. The closest atom approaching the porphyrin is kept at this distance to start with the initial geometry. We will briefly discuss the two points underlying this assumption related to the monomeric separation. Firstly, the pure DFT calculations were carried out in the gas phase and in the absence of any neighboring residues or solvent effects. In the gas phase, it is not possible to capture these effects by simply relying on arbitrary geometries as initial ones for the optimization to progress. As such, it is important to understand the geometrical makeup of the binding pocket of cytochrome P-450. The binding pocket is shown to have a cylindrical orientation with an iron facing diameter of about 10-12 Å, with the distance depending on the crystal structure considered. If the cation needs to be accommodated within this distance, then the exposure of the chain as well as ring cannot be perfectly studied. So, the initial conformations were chosen keeping in mind that the interactions with the active site will eventually bring the substrate, in this case the IL cation, to a position that is chemically favorable, maximizing the expected interactions between the molecules. The distance of approach was determined by performing an initial computational experiment by varying it and

checking for the single point energy at every step. The resulting distance of 3 Åwas chosen from the potential energy profile obtained. The single point energy was used as the deciding quantity because the initial geometry dictates the optimization path as the geometry optimization progresses thus, choosing based on the optimized geometry would be biased.

The next part of this section would be dedicated to the optimization algorithms available in Gaussian and their applicability to the present structures. The DFT calculations related to metal atoms become challenging due to the fact that orbital coefficients are quick to fluctuate between different guesses. The reason behind this can be explained by the presence of orbitals that are close to being degenerate at the valence levels. For example, all the unpaired electrons in Fe(II) porphyrins are within the energetic level of about 1.5 kcal/mol and frequently switch based on geometrical and electronic conditions encountered during potential energy relaxation. As such, the development of the first initial guess becomes rather non-trivial and dependent on the algorithm that is provided. Another pitfall related to these calculations in called the spin contamination error. This is simply the difference in the calculated and estimated spins from the iterative process arising out of the wavefunction and eigenvalue calculations. To avoid the spin contamination and provide a legible guess for the geometry, the option scf=xqc or scf=qc is helpful as it restricts the optimization to go through linear searches when far from the minima and standard Newton Raphson as the forces start to converge to the threshold values. While using this non-standard optimization algorithm, it must be made sure that the intermediate geometries are not physically unrealizable. Some of the examples of physically unrealizable geometries are unnecessary curling, bond breaking, wavefunction or geometry collapsing. All of these problems can be solved by starting from an different initial geometry or starting from the final realistic geometry and then imposing slower convergence conditions on the system.

An issue often faced for these calculations is an extremely slow convergence or the algorithm being stuck at a local minima with the force and displacement

conditions not showing improvement even after several steps. In these cases, two solutions are useful. The first one being perturbing the structure along the path of optimization by a small value, normally upto 1 Ådoes not cause the minima to move very far from the intended one. The second solution would be to stop the calculation at the point where the forces and displacements are stalled and then let the program initialize a new guess from the geometry itself without providing a checkpoint file. The checkpoint file is very useful for jobs that require more time than what is provided initially through the submission file. In the calculations described in the present work, the authors' experience suggests that the forces converge quicker than the displacements for the metal containing complexes. When these forces are negligible and close to 0, the overall optimization can be considered to be successful. In none of the calculations performed for this work, the optimization tolerance was changed initially and the default values outlined by the Gaussian program were used.

7.2 Counterpoise Corrections

The geometry optimization is followed by a counterpoise correction calculation to determine the correction to the overall wavefunction. This corrected wavefunction yields the energy used to evaluate the binding energy. This calculations entails the calculation of the individual monomers as well as the wavefunction of the overall complex. The first step is the calculation of the complex wavefunction at the basis set level of the complex. If this step fails itself, then the geometry optimization might have to be repeated with tighter convergence conditions. This can be performed by adding to the already established grid size and making some other substitutions to the scf algorithm. The complex wavefunction calculation is followed by the monomer calculation at their own basis and also at the complex basis. Often, the monomer calculations do not converge at their own basis in the scheme of the algorithm. In that situation, it is helpful to repeat the optimization with an increased grid or integral size in Gaussian. If all of the above steps fail,

then the final solution is to change the algorithm for the counterpoise calculation. This was often to resorted to as the final step and must be noted for the individual geometry. Normally, the counterpoise calculation must be performed without using any special conditions on the algorithm.

Another important note is on the determination of counterpoise corrections. These corrections are needed to be applied because of the overstabilization of complex wavefunction due to the presence of unequal basis sets on the monomers. Hence, the number of atoms and the interaction pattern of the monomers is important in determining the overall counterpoise correction. The interaction strength between the monomers is directly proportional to the counterpoise energy to the complex wavefunction. Our calculations suggest that these corrections are in the order of 3-6 kcal/mol for the systems studied in the present work. At the absolute scale, these corrections are very important in determining the final binding energy. Although, on a relative scale between conformations, the difference in corrections are not energetically significant. This provides a clue to determine if the counterpoise correction applied to the complex is correct or not. A simple test to achieve this is to perform the counterpoise correction calculations for two different algorithms and test the complex energies obtained from them. The scf=xqc algorithm eventually optimized to the level where the presence of the linear search does not affect the Newton Raphson treatment.

7.3 Population Analysis

The population analysis provides the distribution of the energy levels from the outcome of the geometry optimization and the spatial orientation of these energy levels. It is essential for the population analysis to be performed only on an optimized geometry. In the common method, not special requirements should be imposed on the algorithm for the population analysis. To save the time for the orbital guess for the analysis, the checkpoint file from the previously converged

calculation can be used. For difficult convergence cases, an exception to this can be made by utilizing the scf=xqc option but should be avoided. A good example of this is the presence of cysteine in the active site of the present work, where this option was used to perform population analysis.

Several other sister analyses can also be performed that basically arise from accounting for the distribution of the electronic energy levels. The charge distribution on atoms is important for explaining intermolecular charge transfer. The atomic charges calculated before and after optimization can be used to analyze the above charge transfer for different conformations and systems. It must be kept in mind that the same scheme should be used for all of the geometries for which the calculation is desired. Absolute charges for atoms are generally not reliable and depend on the model as well as method used for their calculation. Most of the schemes available in literature use a hybrid method to determine point specific electrostatic potential that may or may not use van der Waals radii of individual elements. If the van der Waals radii are needed, then they can be obtained from literature and an initial fit might be useful to compare with literature.

One of the popular ones arising from charge distributions is the Fukui function used to describe the local electron accepting ability of the system. For this, the charges are evaluated for atomic sites at neutral and charged electronic states. The key for determining accurate Fukui charges is to always consider the neutral optimized geometry and impose different electronic distributions on it. Often, this creates convergence issues that can be solved by using scf=xqc keyword for all the geometries involved in the calculation. For the Fukui function calculations, from the different schemes available for use, Hirshfeld scheme was employed in this work due to its ability to describe porphyrin containing systems. Other schemes may also be used and tested for their relevance.

For the natural bond orbital calculations, it is necessary for the optimized geometry to be used and if possible, different components of the system to be

specified. NBO6 and Gaussian 16 allow it while it is not available in Gaussian 09. In Gaussian 09, the metal atom is considered as separate entity or unit, and should not be confused for a different part of the system. In order for calculating the perturbations and stabilization, the contributions arising from the different units can be combined and separated as needed.

7.4 Docking

The docking calculations were carried out using the Autodock Vina program which has a scoring function that includes both knowledge based and empirical terms. This program was chosen because of it tested performance on P-450 containing systems and the fact that the training set for the empirical part of the scoring function includes P-450 cam complexes. While docking, it is necessary to keep the center of the search volume and its dimensions consistent for all the substrates considered in the work. It is clearly understood that in the absence of specific literature, the center needs to be determined visually to accommodate the best geometrical fit for the binding pocket. If this is the case, then the dimensions of the search volume need to be consistent throughout all the calculations. Even after the docking preparation, it must be made sure that the protonation states and charge assignments are retained within the receptor. For the non-standard residues, the charge fit is important and must be performed from the extracted geometry at a QM level. While fitting the charges of these residues, the protonation states and geometry should remain unchanged otherwise spurious effects may be encountered at latter stages of the calculations or even during the QMMM calculations.

7.5 QMMM Calculations

The QMMM calculations depend on the initial geometry provided by the docking calculations. From the present work, it is evident that the regioselectivity is maintained from the docking process into the QMMM calculation. In Gaussian, the QMMM setup requires some attention in terms of charge and multiplicity as-

signments and configuration. Some of the intricacies associated with the setup of QMMM calculations in general also system-specific, have been discussed in the section related to Methodology. Here, some of the more system-specific details and guidelines are explained as well as discussed in some detail. Gaussian requires the specification of spin and multiplicity for the whole system i.e. QM and MM, the model system(QM) and the MM specification for the QM system. The whole system consists of the model part, the link atoms and the part of the system excluding both of the former elements. For the whole geometry, the charge and multiplicity are taken to be the most basic to reduce any artificial effects. It is considered as electrically neutral with a singlet state, so devoid of any unpaired electrons.

The model part that is treated at a high level of theory, generally follows the charge and spin assignments from established cluster models of the same nature. In the present case, the presence of the thiolate ligand attached to the heme bears a sextet state that has been validated experimentally and computationally. Hence, the sextet state and a charge assignment based on the cluster model was used for the model region in the present work. In addition, the spin state assignment was also tested for doublet, quartet and sextet configurations to make sure that the relative energies were favorable for the highest spin. As already mentioned in the Methodology, the -1 charge assignment was placed for the bigger model consisting of the additional propionate chains and +1 for the one considering the smaller model. A detailed discussion on the results from both have been given in the appropriate section. Finally, the MM assignment for the QM part was the same as the pure QM part, i.e the combination of -1 and +1 with a sextet multiplicity. This was done to ensure uniformity in the MM calculations for the QM part with the high level of theory treatment.

Upon having an initial run in Gaussian, it is important to check that the charge assignment on the individual residues and the collective system retain what has been assigned in the input. The charges on the link atoms would be taken up by the QM region so it will not perfectly be equal to the prior assignment. It

must be made sure that any ambiguous protonation states or charge placements are corrected before the geometry enters the optimization process. An initial run could be a simple ONIOM single point calculation suggesting the energy of the three layers and the charge of the residues. It is also important to note that the polarization of the QM part is not sensitive to the optimization step in the process. As the MM charges are basically spheres embedding the QM part, their charges remain the same during the optimization. Only charges for the active site/ QM model changes according to the SCF calculation and treatment of the link atoms.

The final point of discussion would be regarding convergence of QMMM calculations. As a rule of thumb, electronic embedding better polarizes than just simple mechanical embedding. This is due to the fact that electronic embedding includes electrostatic as well as van der waals interactions while calculating the common energy between QM and MM parts. For difficult convergence cases, the opt=quadmacro option can be used to obtain a good initial guess that can be further treated with the default microiteration scheme. As a final remark, the nosymm keyword is essential for the QMMM calculations. The reason behind this is the fact that Gaussian searches for a point group to transform coordinates to help with optimization. The authors experience suggests that imposing symmetry conditions in QMMM calculations makes the MM region clash with the point group and gives rise to enormously high energy terms. These terms interfere in providing the correct solution and often leads to the failure of microiterations.

REFERENCES

[1] Bernd Jastorff, Reinhold Störmann, Johannes Ranke, Kerstin Mölter, Frauke Stock, Boris Oberheitmann, Wolfgang Hoffmann, Jens Hoffmann, Matthias Nüchter, Bernd Ondruschka, and Juliane Filser. How hazardous are ionic liquids? Structure-activity relationships and biological testing as important elements for sustainability evaluation. *Green Chemistry*, 5(2):136–142, 2003.

[2] Natalia V. Plechkova and Kenneth R. Seddon. Applications of ionic liquids in the chemical industry. *Chemical Society Reviews*, 37(1):123–150, 2008.

[3] Maxim V. Fedorov and Alexei A. Kornyshev. Ionic liquids at electrified interfaces. *Chemical Reviews*, 114(5):2978–3036, 2014.

[4] Jennifer Neumann, Olav Grundmann, Jorg Thöming, Michael Schulte, and Stefan Stolte. Anaerobic biodegradability of ionic liquid cations under denitrifying conditions. *Green Chemistry*, 12(4):620–627, 2010.

[5] K Skylaris. DFT for calculations on molecules and materials. *Self-published*, pages 1–26, 2014.

[6] André R. Groenhof, Marcel Swart, Andreas W. Ehlers, and Koop Lammertsma. Electronic ground states of iron porphyrin and of the first species in the catalytic reaction cycle of cytochrome P450s. *Journal of Physical Chemistry A*, 109(15):3411–3417, 2005.

[7] François Ogliaro, Samuël P. De Visser, and Sason Shaik. The 'push' effect of the thiolate ligand in cytochrome P450: A theoretical gauging. *Journal of Inorganic Biochemistry*, 91(4):554–567, 2002.

[8] Sason Shaik, Shimrit Cohen, Yong Wang, Hui Chen, Devesh Kumar, and Walter Thiel. P450 Enzymes : Their Structure , Reactivity , and Selectivitys Modeled by QM / MM Calculations. pages 949–1017, 2010.

[9] Christopher J. C. Whitehouse, Stephen G. Bell, and Luet-Lok Wong. P450 BM3 (CYP102A1): connecting the dots. *Chem. Soc. Rev.*, 41(3):1218–1260, 2012.

[10] Dongqi Wang and Walter Thiel. The oxyheme complexes of P450cam: A QM/MM study. *Journal of Molecular Structure: THEOCHEM*, 898(1-3):90–96, 2009.

[11] Patrik Rydberg, Emma Sigfridsson, and Ulf Ryde. On the role of the axial ligand in heme proteins: A theoretical study. *Journal of Biological Inorganic Chemistry*, 9(2):203–223, 2004.

[12] Manca Povsic. PhD Thesis: Rational redesign of cytochrome P450 BM3 (CYP102A1) towards industrially relevant drug metabolites. 2010.

[13] Johannes Ranke, Stefan Stolte, Reinhold Störmann, Jürgen Aming, and Bernd Jastorff. Design of sustainable chemical products - The example of ionic liquids. *Chemical Reviews*, 107(6):2183–2206, 2007.

[14] Ken Geiser. *Chemicals Without Harm : Policies for a Sustainable World.* Urban and Industrial Environments. The MIT Press, 2015.

[15] *Trends in sustainable development : chemicals, mining, transport and waste management.* United Nations, New York, 2010.

[16] Thomas J Schwartz, Brandon J O'neill, Brent H Shanks, and James A Dumesic. Bridging the chemical and biological catalysis gap: Challenges and outlooks for producing sustainable chemicals. *ACS Catalysis*, 4(6), 2014.

[17] Megan R. Schwarzman. A new chemical economy. *Science*, 349(6253):1175.1–1175, 2015.

[18] *Developments in Sustainable Chemical and Bioprocess Technology*. 1st ed. 2013.. edition, 2013.

[19] *Sustainability in the chemical industry : grand challenges and research needs : a workshop report*. National Academy Press, Washington, D.C., 2006.

[20] *Transforming sustainability strategy into action : the chemical industry*. Wiley-Interscience, Hoboken, N.J., 2005.

[21] R S Boethling, E Sommer, and D DiFiore. Designing small molecules for biodegradability. *Chem. Rev.*, 107(12):2207–2227, 2007.

[22] Bianca Aparecida de Marco, Bárbara Saú Rechelo, Eliane Gandolpho Tótoli, Ana Carolina Kogawa, and Hérida Regina Nunes Salgado. Evolution of green chemistry and its multidimensional impacts: A review. *Saudi Pharmaceutical Journal*, 27(1):1–8, 2019.

[23] Robert S. Boethling, Philip H. Howard, William Meylan, William Stiteler, Julie Beauman, and Nestor Tirado. Group Contribution Method for Predicting Probability and Rate of Aerobic Biodegradation. *Environmental Science and Technology*, 28(3):459–465, 1994.

[24] Evan S. Beach, Zheng Cui, and Paul T. Anastas. Green Chemistry: A design framework for sustainability. *Energy and Environmental Science*, 2(10):1038–1049, 2009.

[25] Dingjun Zhang, Jianmin Chen, Yongming Liang, and Huidi Zhou. Facile synthesis of novel ionic liquids containing dithiocarbamate. *Synth. Commun.*, 35(4):521–526, 2005.

[26] J F Brennecke and E JMaginn. Ionic Liquids:Innovative Fluids for Chemical Processing. *AIChE J.*, 47(11):2384–2389, 2001.

[27] Tamar L. Greaves and Calum J. Drummond. Protic ionic liquids: Properties and applications. *Chem. Rev.*, 108(1):206–237, 2008.

[28] P Wasserscheid, A Bösmann, and C Bolm. Synthesis and properties of ionic liquids derived from the 'chiral pool'. Chemical Communications. 2002(3). *Chem. Commun.*, 3:200–201, 2002.

[29] K N Houk C A Hunter, M J Krische J Lehn, S V Ley M Olivucci, J Thiem M Venturi, and C Wong H Wong H Yamamoto. *Topics in Current Chemistry Topics in Current Chemistry*. 2001.

[30] Rohit L. Vekariya. A review of ionic liquids: Applications towards catalytic organic transformations. *J. Mol. Liq.*, 227:40, 2017.

[31] Kun Dong, Xiaomin Liu, Haifeng Dong, Xiangping Zhang, and Suojiang Zhang. Multiscale Studies on Ionic Liquids. *Chem. Rev.*, 117(10):6636–6695, 2017.

[32] Ksenia S. Egorova, Evgeniy G. Gordeev, and Valentine P. Ananikov. Biological Activity of Ionic Liquids and Their Application in Pharmaceutics and Medicine. *Chem. Rev.*, 117(10):7132–7189, 2017.

[33] B. Gurkan, B. F. Goodrich, E. M. Mindrup, L. E. Ficke, M. Massel, S. Seo, T. P. Senftle, H. Wu, M. F. Glaser, J. K. Shah, E. J. Maginn, J. F. Brennecke, and W. F. Schneider. Molecular design of high capacity, low viscosity, chemically tunable ionic liquids for CO2capture. *J. Phys. Chem. Lett.*, 1(24):3494–3499, 2010.

[34] Christopher P. Fredlake, Jacob M. Crosthwaite, Daniel G. Hert, Sudhir N. V. K. Aki, and Joan F. Brennecke. Thermophysical Properties of Imidazolium-Based Ionic Liquids. *J. Chem. Eng. Data*, 49(4):954–964, 2004.

[35] Matthew D. Green and Timothy E. Long. Designing imidazole-based ionic liquids and ionic liquid monomers for emerging technologies. *Polym. Rev.*, 49(4):291–314, 2009.

[36] Deborah Coleman and Nicholas Gathergood. Biodegradation studies of ionic liquids. *Chem. Soc. Rev.*, 39(2):600–637, 2010.

[37] Andrew Jordan and Nicholas Gathergood. Biodegradation of ionic liquids-a critical review. *Chem. Soc. Rev.*, 44(22):8200–8237, 2015.

[38] Y. Deng, I. Beadham, M. Ghavre, M. F. Costa Gomes, N. Gathergood, P. Husson, B. Légeret, B. Quilty, M. Sancelme, and P. Besse-Hoggan. When can ionic liquids be considered readily biodegradable? Biodegradation pathways of pyridinium, pyrrolidinium and ammonium-based ionic liquids. *Green Chem.*, 17(3):1479–1491, 2015.

[39] Stefan Stolte, Stephanie Steudte, Amaya Igartua, and Piotr Stepnowski. The Biodegradation of Ionic Liquids - the View from a Chemical Structure Perspective. *Curr. Org. Chem.*, 15(12):1946–1973, 2011.

[40] Johannes Ranke, Stefan Stolte, Reinhold Störmann, Jürgen Aming, and Bernd Jastorff. Design of sustainable chemical products - The example of ionic liquids. *Chem. Rev.*, 107(6):2183–2206, 2007.

[41] Jindal K. Shah, Joan F. Brennecke, and Edward J. Maginn. Thermodynamic properties of the ionic liquid 1-n-butyl-3-methylimidazolium hexafluorophosphate from Monte Carlo simulations. *Green Chem.*, 4(2):112–118, 2002.

[42] Jindal K. Shah and Edward J. Maginn. A Monte Carlo simulation study of the ionic liquid 1-n-butyl-3- methylimidazolium hexafluorophosphate: Liquid structure, volumetric properties and infinite dilution solution thermodynamics of CO2. *Fluid Phase Equilib.*, 222-223:195–203, 2004.

[43] Jitendra R. Harjani, Jeff Farrell, M. Teresa Garcia, Robert D. Singer, and Peter J. Scammells. Further investigation of the biodegradability of imidazolium ionic liquids. *Green Chem.*, 11(6):821–829, 2009.

[44] Yun Deng, Saibh Morrissey, Nicholas Gathergood, Anne Marie Delort, Pascale Husson, and Margarida F.Costa Gomes. The presence of functional groups key for biodegradation in ionic liquids: Effect on gas solubility. *ChemSusChem*, 3(3):377–385, 2010.

[45] Jairton Dupont, Roberto F. De Souza, and Paulo A.Z. Suarez. Ionic liquid (molten salt) phase organometallic catalysis. *Chemical Reviews*, 102(10):3667–3692, 2002.

[46] J Earle Martyn and R Seddon Kenneth. Ionic liquids. Green solvents for the future. *Pure and Applied Chemistry*, 72(7):1391, 2000.

[47] Andrew Jordan and Nicholas Gathergood. Biodegradation of ionic liquids-a critical review. *Chemical Society Reviews*, 44(22):8200–8237, 2015.

[48] Rohitkumar G. Gore, Lauren Myles, Marcel Spulak, Ian Beadham, Teresa M. Garcia, Stephen J. Connon, and Nicholas Gathergood. A new generation of aprotic yet Brønsted acidic imidazolium salts: Effect of ester/amide groups in the C-2, C-4 and C-5 on antimicrobial toxicity and biodegradation. *Green Chemistry*, 15(10):2747–2760, 2013.

[49] Stefan Stolte, Stephanie Steudte, Amaya Igartua, and Piotr Stepnowski. The Biodegradation of Ionic Liquids - the View from a Chemical Structure Perspective. *Current Organic Chemistry*, 15(12):1946–1973, 2011.

[50] Stefan Stolte, Salha Abdulkarim, Jürgen Arning, Anne-Katrin Blomeyer-Nienstedt, Ulrike Bottin-Weber, Marianne Matzke, Johannes Ranke, Bernd Jastorff, and Jorg Thöming. Primary biodegradation of ionic liquid cations, identification of degradation products of 1-methyl-3-octylimidazolium chloride and electrochemical wastewater treatment of poorly biodegradable compounds. *Green Chemistry*, 10(2):214–224, 2008.

[51] M. Markiewicz, J. Henke, A. Brillowska-Dąbrowska, S. Stolte, J. Łuczak, and C. Jungnickel. Bacterial consortium and axenic cultures isolated from activated sewage sludge for biodegradation of imidazolium-based ionic liquid. *International Journal of Environmental Science and Technology*, 11(7):1919–1926, 2014.

[52] Ewa Liwarska-Bizukojc, Cedric Maton, and Christian V. Stevens. Biodegradation of imidazolium ionic liquids by activated sludge microorganisms. *Biodegradation*, 26(6):453–463, 2015.

[53] Kathryn M Docherty, Michelle V Joyce, Konrad J Kulacki, and Charles F Kulpa. Microbial biodegradation and metabolite toxicity of three pyridinium-based cation ionic liquids. *Green Chemistry*, 12(4):701–712, 2010.

[54] Bodo Philipp, Malte Hoff, Florence Germa, Bernhard Schink, Dieter Beimborn, and Volker Mersch-Sundermann. Biochemical interpretation of quantitative structure-activity relationships (QSAR) for biodegradation of N-heterocycles: A complementary approach to predict biodegradability. *Environmental Science and Technology*, 41(4):1390–1398, 2007.

[55] Leigh Ford, Jitendra R. Harjani, Farzad Atefi, M. Teresa Garcia, Robert D. Singer, and Peter J. Scammells. Further studies on the biodegradation of ionic liquids. *Green Chemistry*, 12(10):1783–1789, 2010.

[56] Jennifer Neumann, Stephanie Steudte, Chul Woong Cho, Jorg Thöming, and Stefan Stolte. Biodegradability of 27 pyrrolidinium, morpholinium, piperidinium, imidazolium and pyridinium ionic liquid cations under aerobic conditions. *Green Chemistry*, 16(4):2174–2184, 2014.

[57] Ilia G. Denisov, Thomas M. Makris, Stephen G. Sligar, and Ilme Schlichting. Structure and Chemistry of Cytochrome P450. *Chemical Reviews*, 105(6):2253–2278, 2005.

[58] Jonathan P. Clark, Caroline S. Miles, Christopher G. Mowat, Malcolm D. Walkinshaw, Graeme A. Reid, Simon N. Daff, and Stephen K. Chapman. The role of Thr268 and Phe393 in cytochrome P450 BM3. *Journal of Inorganic Biochemistry*, 100(5-6):1075–1090, 2006.

[59] Peter Hlavica and Michael Lehnerer. Oxidative Biotransformation of Fatty Acids by Cytochromes P450: Predicted Key Structural Elements Orchestrating Substrate Specificity, Regioselectivity and Catalytic Efficiency. *Current Drug Metabolism*, 11(1):85–104, 2010.

[60] Abhik Ghosh. First-Principles Quantum Chemical Studies of Porphyrins. *Accounts of Chemical Research*, 31(4):189–198, 1998.

[61] Abhik Ghosh and Tebikie Wondimagegn. A theoretical study of axial tilting and equatorial asymmetry in metalloporphyrin - Nitrosyl complexes [16]. *Journal of the American Chemical Society*, 122(33):8101–8102, 2000.

[62] Mike M.Y. Chen, Christopher D. Snow, Christina L. Vizcarra, Stephen L. Mayo, and Frances H. Arnold. Comparison of random mutagenesis and semi-rational designed libraries for improved cytochrome P450 BM3-catalyzed hydroxylation of small alkanes. *Protein Engineering, Design and Selection*, 25(4):171–178, 2012.

[63] Masayuki Hata, Yoshinori Hirano, Tyuji Hoshino, Rie Nishida, and Minoru Tsuda. Theoretical study on compound I formation in monooxygenation mechanism by cytochrome P450. *Journal of Physical Chemistry B*, 108(30):11189–11195, 2004.

[64] Fredi Brühlmann, Laurent Fourage, Christophe Ullmann, Olivier P. Haefliger, Nicolas Jeckelmann, Cédric Dubois, and Denis Wahler. Engineering cytochrome P450 BM3 of Bacillus megaterium for terminal oxidation of palmitic acid. *Journal of Biotechnology*, 184:17–26, 2014.

[65] Jonathan P. Clark, Caroline S. Miles, Christopher G. Mowat, Malcolm D. Walkinshaw, Graeme A. Reid, Simon N. Daff, and Stephen K. Chapman. The role of Thr268 and Phe393 in cytochrome P450 BM3. *Journal of Inorganic Biochemistry*, 100(5-6):1075–1090, 2006.

[66] W.a. Pryor. Cytochrome P450: Structure, mechanism, and biochemistry. *Free Radical Biology and Medicine*, 21(2):251, 1996.

[67] Chul Ho Yun, Keon Hee Kim, Dong Hyun Kim, Heung Chae Jung, and Jae Gu Pan. The bacterial P450 BM3: a prototype for a biocatalyst with human P450 activities. *Trends in Biotechnology*, 25(7):289–298, 2007.

[68] Eugene G Hrycay. *Peroxidase and Peroxygenase Properties and Mechanisms of Cytochrome P450*. Number January. 2016.

[69] L. Ashley Cowart, John R. Falck, and Jorge H. Capdevila. Structural determinants of active site binding affinity and metabolism by cytochrome P450 BM-3. *Archives of Biochemistry and Biophysics*, 387(1):117–124, 2001.

[70] V. Helms and R. C. Wade. Thermodynamics of water mediating protein-ligand interactions in cytochrome P450cam: a molecular dynamics study. *Biophysical Journal*, 69(3):810–824, 1995.

[71] Devesh Kumar, Sam P. De Visser, and Sason Shaik. Multistate reactivity in styrene epoxidation by compound I of cytochrome P450: Mechanisms of products and side products formation. *Chemistry - A European Journal*, 11(9):2825–2835, 2005.

[72] Minor J. Coon. CYTOCHROME P450: Nature's Most Versatile Biological Catalyst. *Annual Review of Pharmacology and Toxicology*, 45(1):1–25, 2005.

[73] Martin Lochner, Markus Meuwly, and Wolf D. Woggon. The origin of the low-spin character of the resting state of cytochrome P450cam investigated by means of active site analogues. *Chemical Communications*, 12:1330–1332, 2003.

[74] J. R. Collins, P. Du, and G. H. Loew. Molecular Dynamics Simulations of the Resting and Hydrogen Peroxide-Bound States of Cytochrome c Peroxidase. *Biochemistry*, 31(45):11166–11174, 1992.

[75] H. Thomann, M. Bernardo, D. Goldfarb, P. M.H. Kroneck, and V. Ullrich. Evidence for Water Binding to the Fe Center in Cytochrome P450cam Obtained by 17O Electron Spin Echo Envelope Modulation Spectroscopy. *Journal of the American Chemical Society*, 117(31):8243–8251, 1995.

[76] Volkhard Helms and Rebecca C. Wade. Hydration energy landscape of the active site cavity in cytochrome P450cam. *Proteins: Structure, Function and Genetics*, 32(3):381–396, 1998.

[77] M. T. Green. Role of the axial ligand in determining the spin state of resting cytochrome P450 [7]. *Journal of the American Chemical Society*, 120(41):10772–10773, 1998.

[78] Rebecca C. Wade, Peter J. Winn, Ilme Schlichting, and Sudarko. A survey of active site access channels in cytochromes P450. *Journal of Inorganic Biochemistry*, 98(7):1175–1182, 2004.

[79] Nandun Thellamurege and Hajime Hirao. Water complexes of cytochrome P450: Insights from energy decomposition analysis. *Molecules*, 18(6):6782–6791, 2013.

[80] Keisuke Sakurai, Hideo Shimada, Takashi Hayashi, and Tomitake Tsukihara. Substrate binding induces structural changes in cytochrome P450cam. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, 65(2):80–83, 2009.

[81] Jan Hermans and Lu Wang. Inclusion of loss of translational and rotational freedom in theoretical estimates of free energies of binding. Application to a complex of benzene and mutant T4 lysozyme. *Journal of the American Chemical Society*, 119(11):2707–2714, 1997.

[82] Christopher F. Butler, Caroline Peet, Amy E. Mason, Michael W. Voice, David Leys, and Andrew W. Munro. Key mutations alter the cytochrome P450 BM3 conformational landscape and remove inherent substrate bias. *Journal of Biological Chemistry*, 288(35):25387–25399, 2013.

[83] Shabana Vohra, Maria Musgaard, Stephen G. Bell, Luet Lok Wong, Weihong Zhou, and Philip C. Biggin. The dynamics of camphor in the cytochrome P450 CYP101D2. *Protein Science*, 22(9):1218–1229, 2013.

[84] Andrew W. Munro, David G. Leys, Kirsty J. McLean, Ker R. Marshall, Tobias W.B. Ost, Simon Daff, Caroline S. Miles, Stephen K. Chapman, Dominikus A. Lysek, Christopher C. Moser, Christopher C. Page, and P. Leslie Dutton. P450 BM3: The very model of a modern flavocytochrome. *Trends in Biochemical Sciences*, 27(5):250–257, 2002.

[85] Yong Wang, Chuanlu Yang, Hongming Wang, Keli Han, and Sason Shaik. A new mechanism for ethanol oxidation mediated by cytochrome P450 2E1: Bulk polarity of the active site makes a difference. *ChemBioChem*, 8(3):277–281, 2007.

[86] Dan Fishelovitch, Sason Shaik, Haim J. Wolfson, and Ruth Nussinov. Theoretical characterization of substrate access/exit channels in the human cytochrome P450 3A4 enzyme: involvement of phenylalanine residues in the gating mechanism. *Journal of Physical Chemistry B*, 113(39):13018–13025, 2009.

[87] F. Peter Guengerich and William W. Johnson. Kinetics of ferric cytochrome P450 reduction by NADPH-cytochrome P450 reductase: Rapid reduction in the absence of substrate and variations among cytochrome P450 systems. *Biochemistry*, 36(48):14741–14750, 1997.

[88] Shiro Yoshioka, Satoshi Takahashi, Hiroshi Hori, Koichiro Ishimori, and Isao Morishima. Proximal cysteine residue is essential for the enzymatic activities of cytochrome p450cam. *European Journal of Biochemistry*, 268(2):252–259, 2001.

[89] P S Coelho, Z J Wang, M E Ener, S A Baril, A Kannan, F H Arnold, and E M Brustad Nat. Axial-Ligand Influence on P450 Reduction Potentials: Implications for Catalysis. *Chem. Biol*, pages 485–487, 2013.

[90] G Cant. Biodegradation of Cosmetics Products: A Computational Study of Cytochrome P450 Metabolism of Phthalates. *Inorganics*, 5(4):77, 2017.

[91] M.F. Perutz. Nature of the iron-oxygen bond and control of oxygen affinity of the haem by the structure of the globin in haemoglobin. *Advances in Experimental Medicine and Biology*, 148:31–48, 1982.

[92] François Ogliaro, Samuël P. de Visser, Shimrit Cohen, Jose Kaneti, and Sason Shaik. The Experimentally Elusive Oxidant of Cytochrome P450: A Theoretical "Trapping" Defining More Closely the "Real" Species. *ChemBioChem*, 2(11):848, 2001.

[93] M. T. Green. The structure and spin coupling of catalase compound I: A study of noncovalent effects [3]. *Journal of the American Chemical Society*, 123(37):9218–9219, 2001.

[94] James J. De Voss, Ole Sibbesen, Zhoupeng Zhang, and Paul R. Ortiz De Montellano. Substrate docking algorithms and prediction of the substrate specificity of cytochrome P450(cam) and its L244A mutant. *Journal of the American Chemical Society*, 119(24):5489–5498, 1997.

[95] Hiroshi Kuramochi, Louis Noodleman, and David A. Case. Density functional study on the electronic structures of model peroxidase compounds I and II. *Journal of the American Chemical Society*, 119(47):11442–11451, 1997.

[96] Christian Audergon, Krishna R. Iyer, Jeffrey P. Jones, John F. Darbyshire, and William F. Trager. Experimental and theoretical study of the effect of active-site constrained substrate motion on the magnitude of the observed intramolecular isotope effect for the P450 101 catalyzed benzylic hydroxylation of isomeric xylenes and 4,4'-dimethylbiphenyl. *Journal of the American Chemical Society*, 121(1):41–47, 1999.

[97] Alfin D.N. Vaz, Dermot F. McGinnity, and Minor J. Coon. Epoxidation of olefins by cytochrome P450: Evidence from site-specific mutagenesis for hydroperoxo-iron as an electrophilic oxidant. *Proceedings of the National Academy of Sciences of the United States of America*, 95(7):3555–3560, 1998.

[98] Robert P. Hanzlik and Kah Hiing John Ling. Active Site Dynamics of Xylene Hydroxylation by Cytochrome P-450 As Revealed by Kinetic Deuterium Isotope Effects. *Journal of the American Chemical Society*, 115(21):9363–9370, 1993.

[99] R. Davydov, T. M. Makris, V. Kofman, D. E. Werst, S. G. Sligar, and B. M. Hoffman. Hydroxylation of camphor by reduced oxy-cytochrome p450cam: Mechanistic implications of EPR and ENDOR studies of catalytic intermediates in native and mutant enzymes. *Journal of the American Chemical Society*, 123(7):1403–1415, 2001.

[100] Ilme Schlichting, Joel Berendzen, Kelvin Chu, Ann M Stock, Shelley A Maves, David E Benson, Robert M Sweet, Dagmar Ringe, Gregory A Petsko, G Stephen, Source Science, New Series, No Mar, Shelley A Maves, David E Benson, Robert M Sweet, Dagmar Ringe, Gregory A Petsko, and Stephen G Sligar. The Catalytic Pathway of Cytochrome P450cam at Atomic Resolution: The Catalytic Pathway of Cytochrome P450cam at Atomic Resolution. 287(5458):1615–1622, 2017.

[101] Karine Auclair, Zhengbo Hu, Dorothy M. Little, Paul R. Ortiz de Montellano, and John T. Groves. Revisiting the mechanism of P450 enzymes with the radical clocks norcarane and spiro[2,5]octane. *Journal of the American Chemical Society*, 124(21):6020–6027, 2002.

[102] S. P. De Visser, F. Ogliaro, N. Harris, and S. Shaik. Multi-state epoxidation of ethene by cytochrome P450: A quantum chemical study. *Journal of the American Chemical Society*, 123(13):3037–3047, 2001.

[103] Detlef Schröder, Sason Shaik, and Helmut Schwarz. Two-state reactivity as a new concept in organometallic chemistry. *Accounts of Chemical Research*, 33(3):139–145, 2000.

[104] F. Ogliaro, S. Cohen, S. P. De Visser, and S. Shaik. Medium polarization and hydrogen bonding effects on compound I of cytochrome P450: What kind of a radical is it really? [15]. *Journal of the American Chemical Society*, 122(51):12892–12893, 2000.

[105] K. Yoshizawa, T. Kamachi, and Y. Shiota. A theoretical study of the dynamic behavior of alkane hydroxylation by a compound I model of cytochrome P450. *Journal of the American Chemical Society*, 123(40):9806–9816, 2001.

[106] Shengxi Jin, Thomas M. Makris, Thomas A. Bryson, Stephen G. Sligar, and John H. Dawson. Epoxidation of olefins by hydroperoxo-ferric cytochrome P450. *Journal of the American Chemical Society*, 125(12):3406–3407, 2003.

[107] Sam P. De Visser and Sason Shaik. A proton-shuttle mechanism mediated by the porphyrin in benzene hydroxylation by cytochrome P450 enzymes. *Journal of the American Chemical Society*, 125(24):7413–7424, 2003.

[108] Martin Newcomb, Runnan Shen, Seung Yong Choi, Patrick H. Toy, Paul F. Hollenberg, Alfin D.N. Vaz, and Minor J. Coon. Cytochrome P450-catalyzed hydroxylation of mechanistic probes that distinguish between radicals and cations. Evidence for cationic but not for radical intermediates. *Journal of the American Chemical Society*, 122(12):2677–2686, 2000.

[109] Visvaldas Kairys and Michael K. Gilson. Enhanced docking with the mining minima optimizer: Acceleration and side-chain flexibility. *Journal of Computational Chemistry*, 23(16):1656–1670, 2002.

[110] Jin Young Park and Dan Harris. Construction and assessment of models of CYP2E1: Predictions of metabolism from docking, molecular dynamics, and density functional theoretical calculations. *Journal of Medicinal Chemistry*, 46(9):1645–1660, 2003.

[111] Damián A. Scherlis, Marcelo A. Martí, Pablo Ordejón, and Darío A. Estrin. Environment effects on chemical reactivity of heme proteins. *International Journal of Quantum Chemistry*, 90(4-5):1505–1514, 2002.

[112] Attila Szabo and Neil S. Ostlund. *Modern quantum chemistry : introduction to advanced electronic structure theory*. 1982.

[113] R. McWeeny. *Methods of Molecular Quantum Mechanics*. 1992.

[114] Peter Atkins and Ronald Friedman. *Molecular Quantum Mechanics*. 1997.

[115] Frank Jensen. *Introduction to Computational Chemistry*. 1999.

[116] Axel D. Becke. Perspective: Fifty years of density-functional theory in chemical physics. *Journal of Chemical Physics*, 140(18), 2014.

[117] Axel D Becke. Density-functional thermochemistry. I. The effect of the exchange-only gradient correction. *The Journal of Chemical Physics*, 96(3):2155–2160, feb 1992.

[118] T Hamilton and P Pulay. Direct inversion in the iterative subspace (diis) optimization of open- shell, excited-state, and small multiconfiguration scf wave functions. *Journal of Chemical Physics*, 84:5728–5734, 1986.

[119] Stefan Grimme. Semiempirical hybrid density functional with perturbative second-order correlation. *Journal of Chemical Physics*, 124(3), 2006.

[120] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *Journal of Chemical Physics*, 132(15), 2010.

[121] Stefan Grimme. Treatment of Dispersion Interactions with DFT-D3. *Interactions*, pages 1–6, 2011.

[122] Christopher J. Cramer and Donald G. Truhlar. Density functional theory for transition metals and transition metal chemistry. *Physical Chemistry Chemical Physics*, 11(46):10757–10816, 2009.

[123] C. David Sherrill, Tait Takatani, and Edward G. Hohenstein. An assessment of theoretical methods for nonbonded interactions: Comparison to complete basis set limit coupled-cluster potential energy curves for the benzene dimer, the methane dimer, benzene-methane, and benzene-H2S. *Journal of Physical Chemistry A*, 113(38):10146–10159, 2009.

[124] Narbe Mardirossian and Martin Head-Gordon. Thirty years of density functional theory in computational chemistry: An overview and extensive assessment of 200 density functionals. *Molecular Physics*, 115(19):2315–2372, 2017.

[125] David W. Schwenke and Donald G Truhlar. Systematic study of basis set superposition errors in the calculated interaction energy of two hf molecules. *Journal of Chemical Physics*, 82(5):2418–2426, 1 1985.

[126] Yan Zhao and Donald G. Truhlar. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four M06-class functionals and 12 other function. *Theoretical Chemistry Accounts*, 120(1-3):215–241, 2008.

[127] W J Hehre, R Ditchfield, and J A Pople. Self—Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian—Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.*, 56(5):2257–2261, March 1972.

[128] P C Hariharan and J A Pople. The Influence of Polarization Functions on Molecular Orbital Hydrogenation Energies. *Theoret. chim. Acta*, 28:213–222, April 2003.

[129] Zhen Cao, Shu Li, and Tianying Yan. Cation-π Interactions between a Free-Base Porphyrin and an Ionic Liquid: A Computational Study. *Chem. Phys. Chem*, 13(7):1743–1747, January 2012.

[130] Knut Angenendt and Patrik Johansson. Ionic Liquid Structures from Large Density Functional Theory Calculations Using Mindless Configurations †. *J. Phys. Chem. C.*, 114(48):20577–20582, December 2010.

[131] Gregorio Garc, Mert Atilhan, and Santiago Aparicio. Assessment of DFT methods for studying acid gas capture by ionic liquids. *Phys. Chem. Chem. Phys.*, 17:26875–26891, October 2015.

[132] Zhixiang Song, Haijun Wang, and Lijuan Xing. Density Functional Theory Study of the Ionic Liquid [emim]OH and Complexes [emim]OH(H2O) n (n=1,2). *J. Sol. Chem.*, 38(9):1139–1154, July 2009.

[133] Ekaterina I Izgorodina. Towards large-scale, fully ab initio calculations of ionic liquids. *Phys. Chem. Chem. Phys.*, 13(10):4189–19, 2011.

[134] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, N. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox. Gaussian - 09 Revision E.01. 2009. Gaussian Inc. Wallingford CT 2009.

[135] C David Sherrill. Counterpoise Correction and Basis Set Superposition Error. *Self-published*, pages 1–5, 2010.

[136] C. David Sherrill, Tait Takatani, and Edward G. Hohenstein. An assessment of theoretical methods for nonbonded interactions: Comparison to complete basis set limit coupled-cluster potential energy curves for the benzene dimer, the methane dimer, benzene-methane, and benzene-H2S. *Journal of Physical Chemistry A*, 113(38):10146–10159, 2009.

[137] Asger Halkier, Wim Klopper, Trygve Helgaker, Poul Jørgensen, and Peter R. Taylor. Basis set convergence of the interaction energy of hydrogen-bonded complexes. *Journal of Chemical Physics*, 111(20):9157–9167, 1999.

[138] Klaus R. Liedl. Dangers of counterpoise corrected hypersurfaces. Advantages of basis set superposition improvement. *Journal of Chemical Physics*, 108(8):3199–3204, 1998.

[139] J J Dannenberg. Using Perturbation and Frontier Molecular Orbital Theory To Predict Diastereofacial Selectivity. *Chem. Rev.*, 99(5):1225–1242, May 1999.

[140] Lian-Gang Zhuo, Wei Liao, and Zhi-Xiang Yu. A Frontier Molecular Orbital Theory Approach to Understanding the Mayr Equation and to Quantifying Nucleophilicity and Electrophilicity by Using HOMO and LUMO Energies. *Asian J. Org. Chem.*, 1(4):336–345, November 2012.

[141] P Geerlings, F De Proft, and W Langenaseker. Conceptual Density Functional Theory. *Chem. Rev.*, 103(5):1793–1874, May 2003.

[142] Timothy I Morrow and Edward J Maginn. Molecular Dynamics Study of the Ionic Liquid 1- n-Butyl-3-methylimidazolium Hexafluorophosphate. *J. Phys. Chem. B*, 106(49):12807–12813, December 2002.

[143] Pankaz K Sharma, Rouslan Kevorkiants, Sam P de Visser, Devesh Kumar, and Sason Shaik. Porphyrin Traps Its Terminator! Concerted and Stepwise Porphyrin Degradation Mechanisms Induced by Heme-Oxygenase and Cytochrome P450. *Angew. Chem. Int. Ed.*, 43(9):1129–1132, February 2004.

[144] Sason Shaik, Samuël P de Visser, and Devesh Kumar. One oxidant, many pathways: a theoretical perspective of monooxygenation mechanisms by cytochrome P450 enzymes. *J. Biol. Inorg. Chem.*, 9(6):661–668, July 2004.

[145] Jan C Schöneboom, Hai Lin, Nathalie Reuter, Walter Thiel, Shimrit Cohen, François Ogliaro, and Sason Shaik. The Elusive Oxidant Species of Cytochrome P450 Enzymes: Characterization by Combined Quantum Mechanical/Molecular Mechanical (QM/MM) Calculations. *J. Am. Chem. Soc.*, 124(27):8142–8151, July 2002.

[146] Jan C Schöneboom and Walter Thiel. The Resting State of P450 cam: A QM/MM Study. *J. Phys. Chem. B*, 108(22):7468–7478, June 2004.

[147] Ahmet Altun, Sason Shaik, and Walter Thiel. What is the Active Species of Cytochrome P450 during Camphor Hydroxylation? QM/MM Studies of Different Electronic States of Compound I and of Reduced and Oxidized Iron–Oxo Intermediates. *J. Am. Chem. Soc.*, 129(29):8978–8987, July 2007.

[148] Robert G Parr, László v Szentpály, and Shubin Liu. Electrophilicity Index. *J. Am. Chem. Soc.*, 121(9):1922–1924, March 1999.

[149] P. Hunt. Molecular Orbitals and Population Analysis. *Presentation File*, pages 1–10, 2009.

[150] Gabriela Bitencourt-Ferreira and Walter Filgueira de Azevedo. *Docking with SwissDock*, volume 2053. 2019.

[151] Dik Lung Ma, Daniel Shiu Hin Chan, and Chung Hang Leung. Molecular docking for virtual screening of natural product databases. *Chemical Science*, 2(9):1656–1665, 2011.

[152] Juan C. Alvarez. High-throughput docking as a source of novel drug leads. *Current Opinion in Chemical Biology*, 8(4):365–370, 2004.

[153] Yuan Ping Pang, Emanuele Perola, Run Xu, and Franklyn G. Prendergast. EUDOC: A computer program for identification of drug interaction sites in macromolecules and drug leads from chemical databases. *Journal of Computational Chemistry*, 22(15):1750–1771, 2001.

[154] Ajay N. Jain. Surflex: Fully automatic flexible molecular docking using a molecular similarity-based search engine. *Journal of Medicinal Chemistry*, 46(4):499–511, 2003.

[155] Richard A. Friesner, Jay L. Banks, Robert B. Murphy, Thomas A. Halgren, Jasna J. Klicic, Daniel T. Mainz, Matthew P. Repasky, Eric H. Knoll, Mee Shelley, Jason K. Perry, David E. Shaw, Perry Francis, and Peter S. Shenkin. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *Journal of Medicinal Chemistry*, 47(7):1739–1749, 2004.

[156] M. Stahl and M. Rarey. Detailed analysis of scoring functions for virtual screening. *Journal of Medicinal Chemistry*, 44(7):1035–1042, 2001.

[157] Miles Congreve, Christopher J. Langmead, Jonathan S. Mason, and Fiona H. Marshall. Progress in structure based drug design for G protein-coupled receptors. *Journal of Medicinal Chemistry*, 54(13):4283–4311, 2011.

[158] Max W. Chang, Christian Ayeni, Sebastian Breuer, and Bruce E. Torbett. Virtual screening for HIV protease inhibitors: A comparison of AutoDock 4 and Vina. *PLoS ONE*, 5(8):1–9, 2010.

[159] Sheng You Huang and Xiaoqin Zou. Advances and challenges in Protein-ligand docking. *International Journal of Molecular Sciences*, 11(8):3016–3034, 2010.

[160] Guillermo Sanchez. Las instituciones de ciencia y tecnología en los procesos de aprendizaje de la producción agroalimentaria en Argentina. *El sistema argentino de innovación: instituciones, empresas y redes. El desafío de la creación y apropiación de conocimiento.*, 26(February):15–26, 2013.

[161] Boris Gorelik and Amiram Goldblum. High quality binding modes in docking ligands to proteins. *Proteins: Structure, Function and Genetics*, 71(3):1373–1386, 2008.

[162] G. E. Terp, B. N. Johansen, I. T. Christensen, and F. S. Jørgensen. A new concept for multidimensional selection of ligand conformations (multiselect) and multidimensional scoring (multiscore) of protein-ligand binding affinities. *Journal of Medicinal Chemistry*, 44(14):2333–2343, 2001.

[163] Paul S. Charifson, Joseph J. Corkery, Mark A. Murcko, and W. Patrick Walters. Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. *Journal of Medicinal Chemistry*, 42(25):5100–5109, 1999.

[164] Didier Rognan, Sanne Lise Lauemøller, Arne Holm, Søren Buus, and Vincenzo Tschinke. Predicting binding affinities of protein ligands from three-dimensional models: Application to peptide binding to class I major histocompatibility proteins. *Journal of Medicinal Chemistry*, 42(22):4650–4658, 1999.

[165] William L. Jorgensen. The Many Roles of Computation in Drug Discovery. *Science*, 303(5665):1813–1818, 2004.

[166] Christopher A. Lipinski, Franco Lombardo, Beryl W. Dominy, and Paul J. Feeney. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, 64(SUPPL.):4–17, 2012.

[167] I. M. Kapetanovic. Computer-aided drug discovery and development (CADDD): In silico-chemico-biological approach. *Chemico-Biological Interactions*, 171(2):165–176, 2008.

[168] Brian K. Shoichet and Brian K. Kobilka. Structure-based drug screening for G-protein-coupled receptors. *Trends in Pharmacological Sciences*, 33(5):268–272, 2012.

[169] Chandrika B-Rao, Jyothi Subramanian, and Somesh D. Sharma. Managing protein flexibility in docking and its applications. *Drug Discovery Today*, 14(7-8):394–400, 2009.

[170] Jacob D. Durrant and J. Andrew McCammon. Computer-aided drug-discovery techniques that account for receptor flexibility. *Current Opinion in Pharmacology*, 10(6):770–774, 2010.

[171] Malgorzata N. Drwal and Renate Griffith. Combination of ligand- and structure-based methods in virtual screening. *Drug Discovery Today: Technologies*, 10(3):e395–e401, 2013.

[172] Sheng You Huang, Min Li, Jianxin Wang, and Yi Pan. HybridDock: A Hybrid Protein-Ligand Docking Protocol Integrating Protein- and Ligand-Based Approaches. *Journal of Chemical Information and Modeling*, 56(6):1078–1087, 2016.

[173] J. P. Hughes, S. S. Rees, S. B. Kalindjian, and K. L. Philpott. Principles of early drug discovery. *British Journal of Pharmacology*, 162(6):1239–1249, 2011.

[174] Dimitris K. Agrafiotis, Alan C. Gibbs, Fangqiang Zhu, Sergei Izrailev, and Eric Martin. Conformational sampling of bioactive molecules: A comparative study. *Journal of Chemical Information and Modeling*, 47(3):1067–1086, 2007.

[175] Lívia B. Salum, Igor Polikarpov, and Adriano D. Andricopulo. Structure-based approach for the study of estrogen receptor binding affinity and subtype selectivity. *Journal of Chemical Information and Modeling*, 48(11):2243–2253, 2008.

[176] Renxiao Wang, Yipin Lu, and Shaomeng Wang. Comparative evaluation of 11 scoring functions for molecular docking. *Journal of Medicinal Chemistry*, 46(12):2287–2303, 2003.

[177] Lyn H. Jones and Mark E. Bunnage. Applications of chemogenomic library screening in drug discovery. *Nature Reviews Drug Discovery*, 16(4):285–296, 2017.

[178] Sheng You Huang and Xiaoqin Zou. Advances and challenges in Protein-ligand docking. *International Journal of Molecular Sciences*, 11(8):3016–3034, 2010.

[179] Abdul-rahman Allouche. Software News and Updates Gabedit — A Graphical User Interface for Computational Chemistry Softwares. *Journal of computational chemistry*, 32:174–182, 2012.

[180] Thomas Simonson, Georgios Archontis, and Martin Karplus. Free energy simulations come of age: Protein-ligand recognition. *Accounts of Chemical Research*, 35(6):430–437, 2002.

[181] A. C. Pierce and W. L. Jorgensen. Estimation of binding affinities for selective thrombin inhibitors via Monte Carlo simulations. *Journal of Medicinal Chemistry*, 44(7):1043–1050, 2001.

[182] Lukas P. Pason and Christoph A. Sotriffer. Empirical Scoring Functions for Affinity Prediction of Protein-ligand Complexes. *Molecular Informatics*, 35(11-12):541–548, 2016.

[183] Maria Batool, Bilal Ahmad, and Sangdun Choi. A structure-based drug discovery paradigm. *International Journal of Molecular Sciences*, 20(11), 2019.

[184] Elizabeth Yuriev, Mark Agostino, and Paul A. Ramsland. Challenges and advances in computational docking: 2009 in review. *Journal of Molecular Recognition*, 24(2):149–164, 2011.

[185] Leonardo G. Ferreira, Ricardo N. Dos Santos, Glaucius Oliva, and Adriano D. Andricopulo. *Molecular docking and structure-based drug design strategies*, volume 20. 2015.

[186] Jan Jakubík, Esam E. El-Fakahany, and Vladimír Doležal. Towards predictive docking at aminergic G-protein coupled receptors. *Journal of Molecular Modeling*, 21(11):1–18, 2015.

[187] Sheng You Huang, Sam Z. Grinter, and Xiaoqin Zou. Scoring functions and their evaluation methods for protein-ligand docking: Recent advances and future directions. *Physical Chemistry Chemical Physics*, 12(40):12899–12908, 2010.

[188] Yunierkis Perez-Castillo, Stellamaris Sotomayor-Burneo, Karina Jimenes-Vargas, Mario Gonzalez-Rodriguez, Maykel Cruz-Monteagudo, Vinicio Armijos-Jaramillo, M. Natália D.S. Cordeiro, Fernanda Borges, Aminael Sánchez-Rodríguez, and Eduardo Tejera. CompScore: Boosting Structure-Based Virtual Screening Performance by Incorporating Docking Scoring Function Components into Consensus Scoring. *Journal of Chemical Information and Modeling*, 59(9):3655–3666, 2019.

[189] Julien Michel, Julian Tirado-Rives, and William L. Jorgensen. Energetics of displacing water molecules from protein binding sites: Consequences for ligand optimization. *Journal of the American Chemical Society*, 131(42):15403–15411, 2009.

[190] Alasdair T.R. Laurie and Richard M. Jackson. Q-SiteFinder: An energy-based method for the prediction of protein-ligand binding sites. *Bioinformatics*, 21(9):1908–1916, 2005.

[191] Hezekiel Mathambo Kumalo, Soumendranath Bhakat, and Mahmoud E.S. Soliman. Theory and applications of covalent docking in drug discovery: Merits and pitfalls. *Molecules*, 20(2):1984–2000, 2015.

[192] Michelle M.R. Arkin and James A. Wells. Small-molecule inhibitors of protein-protein interactions: Progressing towards the dream. *Nature Reviews Drug Discovery*, 3(4):301–317, 2004.

[193] Christopher R. Corbeil, Christopher I. Williams, and Paul Labute. Variability in docking success rates due to dataset preparation. *Journal of Computer-Aided Molecular Design*, 26(6):775–786, 2012.

[194] Simon Tietze and Joannis Apostolakis. GlamDock: Development and validation of a new docking tool on several thousand protein-ligand complexes. *Journal of Chemical Information and Modeling*, 47(4):1657–1672, 2007.

[195] Reiji Teramoto and Hiroaki Fukunishi. Supervised consensus scoring for docking and virtual screening. *Journal of Chemical Information and Modeling*, 47(2):526–534, 2007.

[196] Pablo Englebienne and Nicolas Moitessier. Docking ligands into flexible and solvated macromolecules. 5. Force-field-based prediction of binding affinities of ligands to proteins. *Journal of Chemical Information and Modeling*, 49(11):2564–2571, 2009.

[197] Oliver Korb, Thomas Stützle, and Thomas E. Exner. Empirical scoring functions for advanced Protein-Ligand docking with PLANTS. *Journal of Chemical Information and Modeling*, 49(1):84–96, 2009.

[198] Yat T. Tang and Garland R. Marshall. Phoenix: A scoring function for affinity prediction derived using high-resolution crystal structures and calorimetry measurements. *Journal of Chemical Information and Modeling*, 51(2):214–228, 2011.

[199] Freddie R. Salsbury. Molecular dynamics simulations of protein dynamics and their relevance to drug discovery. *Current Opinion in Pharmacology*, 10(6):738–744, 2010.

[200] Mengang Xu and Markus A. Lill. Induced fit docking, and the use of QM/MM methods in docking. *Drug Discovery Today: Technologies*, 10(3):e411–e418, 2013.

[201] Hao Fan, Dina Schneidman-Duhovny, John J. Irwin, Guangqiang Dong, Brian K. Shoichet, and Andrej Sali. Statistical potential for modeling and ranking of protein-ligand interactions. *Journal of Chemical Information and Modeling*, 51(12):3078–3092, 2011.

[202] Philip Prathipati and Anil K. Saxena. Evaluation of binary QSAR models derived from LUDI and MOE scoring functions for structure based virtual screening. *Journal of Chemical Information and Modeling*, 46(1):39–51, 2006.

[203] A Guide To and Biomolecular Simulations. *A Guide to Biomolecular Simulations*. 2006.

[204] Hernán Alonso, Andrey A. Bliznyuk, and Jill E. Gready. Combining docking and molecular dynamic simulations in drug design. *Medicinal Research Reviews*, 26(5):531–568, 2006.

[205] Matthew J. Harvey and Gianni De Fabritiis. High-throughput molecular dynamics: The powerful new tool for drug discovery. *Drug Discovery Today*, 17(19-20):1059–1062, 2012.

[206] Sara E. Nichols, Riccardo Baron, Anthony Ivetac, and J. Andrew McCammon. Predictive power of molecular dynamics receptor structures in virtual screening. *Journal of Chemical Information and Modeling*, 51(6):1439–1446, 2011.

[207] Suliman Adam, Michaela Knapp-Mohammady, Jun Yi, and Ana Nicoleta Bondar. Revised CHARMM force field parameters for iron-containing cofactors of photosystem II. *Journal of Computational Chemistry*, 39(1):7–20, 2018.

[208] Glen E. Kellogg and Deliang L. Chen. The importance of being exhaustive. Optimization of bridging structural water molecules and water networks in models of biological systems. *Chemistry and Biodiversity*, 1(1):98–105, 2004.

[209] Peter Monecke, Thorsten Borosch, Jürgen Brickmann, and Stefan M. Kast. Determination of the interfacial water content in protein-protein complexes from free energy simulations. *Biophysical Journal*, 90(3):841–850, 2006.

[210] Zheng Li and Themis Lazaridis. Water at biomolecular binding interfaces. *Physical Chemistry Chemical Physics*, 9(5):573–581, 2007.

[211] Yu Chian Chen. Beware of docking! *Trends in Pharmacological Sciences*, 36(2):78–95, 2015.

[212] Gerhard Klebe. Virtual ligand screening: strategies, perspectives and limitations. *Drug Discovery Today*, 11(13-14):580–594, 2006.

[213] Cristiano R.W. Guimarães and Mario Cardozo. MM-GB/SA rescoring of docking poses in structure-based lead optimization. *Journal of Chemical Information and Modeling*, 48(5):958–970, 2008.

[214] Alessio Amadasi, J. Andrew Surface, Francesca Spyrakis, Pietro Cozzini, Andrea Mozzarelli, and Glen E. Kellogg. Robust classification of "relevant" water molecules in putative protein binding sites. *Journal of Medicinal Chemistry*, 51(4):1063–1067, 2008.

[215] B. K. Muralidhara, Surendra Negi, Christopher C. Chin, Werner Braun, and James R. Halpert. Conformational flexibility of mammalian cytochrome P450 2B4 in binding imidazole inhibitors with different ring chemistry and side chains: Solution thermodynamics and molecular modeling. *Journal of Biological Chemistry*, 281(12):8051–8061, 2006.

[216] Cristina Shino Porro. Quantum mechanical molecular mechanics studies of Cytochrome P450BM3. *The University of Manchester*, PhD, 2010.

[217] A. Warshel and M. Levitt. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *Journal of Molecular Biology*, 103(2):227–249, 1976.

[218] MF Perutz. Electrostatic effects in proteins. *Science*, 201(4362):1187–1191, 1978.

[219] Arieh Warshel. Energetics of enzyme catalysis. *Proceedings of the National Academy of Sciences*, 75(11):5250–5254, 1978.

[220] WF Van Gunsteren and M Karplus. A method for constrained energy minimization of macromolecules. *Journal of Computational Chemistry*, 1(3):266–274, 1980.

[221] Lung Wa Chung, WMC Sameera, Romain Ramozzi, Alister J Page, Miho Hatanaka, Galina P Petrova, Travis V Harris, Xin Li, Zhuofeng Ke, Fengyi Liu, et al. The oniom method and its applications. *Chemical reviews*, 115(12):5678–5796, 2015.

[222] Marc W van der Kamp and Adrian J Mulholland. Combined quantum mechanics/molecular mechanics (qm/mm) methods in computational enzymology. *Biochemistry*, 52(16):2708–2728, 2013.

[223] Thom Vreven and Keiji Morokuma. Hybrid methods: Oniom (qm: mm) and qm/mm. *Annual reports in computational chemistry*, 2:35–51, 2006.

[224] Patrick Bultinck, Ramon Carbó-Dorca, and Wilfried Langenaeker. Negative Fukui functions: New insights based on electronegativity equalization. *Journal of Chemical Physics*, 118(10):4349–4356, 2003.

[225] P. Fuentealba, P. Pérez, and R. Contreras. On the condensed Fukui function. *Journal of Chemical Physics*, 113(7):2544–2551, 2000.

[226] Takehiro Ohta, Koji Matsuura, Kazunari Yoshizawa, and Isao Morishima. The electronic and vibrational structures of iron-oxo porphyrin with a methoxide or cysteinate axial ligand. *Journal of Inorganic Biochemistry*, 82(1-4):141–152, 2000.

[227] F. Ogliaro, N. Harris, S. Cohen, M. Filatov, S. P. De Visser, and S. Shaik. A model 'rebound' mechanism of hydroxylation by cytochrome P450: Stepwise and effectively concerted pathways, and their reactivity patterns. *Journal of the American Chemical Society*, 122(37):8977–8989, 2000.

[228] F. Peter Guengerich. Human Cytochrome P450 Enzymes. *Cytochrome P450: Structure, Mechanism, and Biochemistry*, pages 377–530, 2005.

[229] Dongmei Li, Yong Wang, and Keli Han. Recent density functional theory model calculations of drug metabolism by cytochrome P450. *Coordination Chemistry Reviews*, 256(11-12):1137–1150, 2012.

[230] Richard Lonsdale, Kerensa T. Houghton, Jolanta Zurek, Christine M. Bathelt, Nicolas Foloppe, Marcel J. De Groot, Jeremy N. Harvey, and Adrian J. Mulholland. Quantum mechanics/molecular mechanics modeling of regioselectivity of drug metabolism in cytochrome P450 2C9. *Journal of the American Chemical Society*, 135(21):8001–8015, 2013.

[231] Dongqi Wang, Jingjing Zheng, Sason Shaik, and Walter Thiel. Quantum and Molecular Mechanical Study of the First Proton Transfer in the Catalytic Cycle of Cytochrome P450cam and Its Mutant D251N. *The Journal of Physical Chemistry B*, 112(16):5126–5138, 2008.

[232] Shaik Sason, Kumar Devesh, P. de Visser Samuël, Altun Ahmet, and Thiel Walter. *Theoretical Perspective on the Structure and Mechanism of Cytochrome P450 Enzymes†*, volume 105. 2005.

[233] Junming Ho and Michelle L. Coote. A universal approach for continuum solvent pKa calculations: Are we there yet? *Theoretical Chemistry Accounts*, 125(1):3–21, 2009.

[234] Junming Ho. Are thermodynamic cycles necessary for continuum solvent calculation of pKas and reduction potentials? *Physical Chemistry Chemical Physics*, 17(4):2859–2868, 2015.

[235] Junming Ho and Mehmed Z. Ertem. Calculating Free Energy Changes in Continuum Solvation Models. *Journal of Physical Chemistry B*, 120(7):1319–1329, 2016.

[236] Bishnu Thapa and H. Bernhard Schlegel. Density Functional Theory Calculation of pKa's of Thiols in Aqueous Solution Using Explicit Water Molecules and the Polarizable Continuum Model. *Journal of Physical Chemistry A*, 120(28):5726–5735, 2016.

[237] Raphael F. Ribeiro, Aleksandr V. Marenich, Christopher J. Cramer, and Donald G. Truhlar. Use of solution-phase vibrational frequencies in continuum models for the free energy of solvation. *Journal of Physical Chemistry B*, 115(49):14556–14562, 2011.

[238] Tobias W B Ost, Jonathan Clark, Christopher G Mowat, Caroline S Miles, Malcolm D Walkinshaw, Graeme A Reid, Stephen K Chapman, and Simon Daff. Oxygen activation and electron transfer in flavocytochrome P 450 BM3. *J Am Chem Soc*, 125(49):15010–15020, 2003.

[239] Kshatresh Dutta Dubey and Sason Shaik. Choreography of the Reductase and P450BM3Domains Toward Electron Transfer Is Instigated by the Substrate. *Journal of the American Chemical Society*, 140(2):683–690, 2018.

[240] Abhinav Luthra, Ilia G. Denisov, and Stephen G. Sligar. Spectroscopic features of cytochrome P450 reaction intermediates. *Archives of Biochemistry and Biophysics*, 507(1):26–35, 2011.

[241] Thomas F. Hughes and Richard A. Friesner. Development of accurate DFT methods for computing redox potentials of transition metal complexes: Results for model complexes and application to cytochrome P450. *Journal of Chemical Theory and Computation*, 8(2):442–459, 2012.

[242] Lindsay E. Roy, Elena Jakubikova, M. Graham Guthrie, and Enrique R. Batista. Calculation of one-electron redox potentials revisited. Is it possible to calculate accurate potentials with density functional methods? *Journal of Physical Chemistry A*, 113(24):6745–6750, 2009.

[243] Oleg Trott and Aj Olson. NIH Public Access. *Journal of Computational Chemistry*, 31(2):455–461, 2010.

[244] Ramu Anandakrishnan, Boris Aguilar, and Alexey V. Onufriev. H++ 3.0: Automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Research*, 40(W1):537–541, 2012.

[245] Inacrist Geronimo, Catherine A. Denning, David K. Heidary, Edith C. Glazer, and Christina M. Payne. Molecular Determinants of Substrate Affinity and Enzyme Activity of a Cytochrome P450BM3 Variant. *Biophysical Journal*, 115(7):1251–1263, 2018.

[246] Rebecca C. Wade, Peter J. Winn, Ilme Schlichting, and Sudarko. A survey of active site access channels in cytochromes P450. *Journal of Inorganic Biochemistry*, 98(7):1175–1182, 2004.

[247] José N. Canongia Lopes and Agílio A.H. Pádua. Molecular force field for ionic liquids III: Imidazolium, pyridinium, and phosphonium cations; chloride, bromide, and dicyanamide anions. *Journal of Physical Chemistry B*, 110(39):19586–19592, 2006.

[248] Zhiping Liu, Shiping Huang, and Wenchuan Wang. A refined force field for molecular simulation of imidazolium-based ionic liquids. *Journal of Physical Chemistry B*, 108(34):12978–12989, 2004.

[249] Jesse G. McDaniel, Eunsong Choi, Chang Yun Son, J. R. Schmidt, and Arun Yethiraj. Ab Initio Force Fields for Imidazolium-Based Ionic Liquids. *Journal of Physical Chemistry B*, 120(28):7024–7036, 2016.

[250] Romelia Salomon-Ferrer, David A. Case, and Ross C. Walker. An overview of the Amber biomolecular simulation package. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 3(2):198–210, 2013.

[251] Xin Tian Feng, Jian Guo Yu, Ming Lei, Wei Hai Fang, and Shubin Liu. Toward understanding metal-binding specificity of porphyrin: A conceptual density functional theory study. *Journal of Physical Chemistry B*, 113(40):13381–13389, 2009.

[252] H. Y. Ammar and H. M. Badran. Effect of CO adsorption on properties of transition metal doped porphyrin: A DFT and TD-DFT study. *Heliyon*, 5(10):e02545, 2019.

[253] Chunying Rong, Shixun Lian, Dulin Yin, Aiguo Zhong, Ruiqin Zhang, and Shubin Liu. Effective simulation of biological systems: Choice of density functional and basis set for heme-containing complexes. *Chemical Physics Letters*, 434(1-3):149–154, 2007.

[254] Meng Sheng Liao, Ming Ju Huang, and John D. Watts. Iron porphyrins with different imidazole ligands. A theoretical comparative study. *Journal of Physical Chemistry A*, 114(35):9554–9569, 2010.

[255] Saeed Kazemiabnavi, Zhengcheng Zhang, Katsuyo Thornton, and Soumik Banerjee. Electrochemical Stability Window of Imidazolium-Based Ionic Liquids as Electrolytes for Lithium Batteries. *J. Phys. Chem. B*, 120(25):5691–5702, June 2016.

[256] Wim Buijs, Geert-Jan Witkamp, and Maaike C Kroon. Correlation between Quantumchemically Calculated LUMO Energies and the Electrochemical Window of Ionic Liquids with Reduction-Resistant Anions. *Int. J. Electrochem.*, 2012(7):1–6, 2012.

[257] Noel M O'boyle, Adam L Tenderholt, and Karol M Langner. cclib: A library for package-independent computational chemistry algorithms. *J. Comput. Chem.*, 29(5):839–845, 2008.

[258] Atiya Banerjee and Jindal K Shah. Elucidating the effect of the ionic liquid type and alkyl chain length on the stability of ionic liquid–iron porphyrin complexes. *The Journal of Chemical Physics*, 153(3):034306, 2020.

[259] Atiya Banerjee and Jindal K. Shah. Insight into conformationally-dependent binding of 1-n-alkyl-3-methylimidazolium cations to porphyrin molecules using quantum mechanical calculations. *Phys. Chem. Chem. Phys.*, 21(19):10095–10104, 2019.

[260] Danni L. Harris and Gilda H. Loew. Theoretical investigation of the proton assisted pathway to formation of cytochrome P450 compound I. *Journal of the American Chemical Society*, 120(35):8941–8948, 1998.

[261] Victor Guallar, Danni L. Harris, Victor S. Batista, and William H. Miller. Proton-transfer dynamics in the activation of cytochrome P450eryF. *Journal of the American Chemical Society*, 124(7):1430–1437, 2002.

[262] Irina F. Sevrioukova, Huiying Li, Hong Zhang, Julian A. Peterson, and Thomas L. Poulos. Structure of a cytochrome P450-redox partner electron-transfer complex. *Proceedings of the National Academy of Sciences of the United States of America*, 96(5):1863–1868, 1999.

[263] Jacob D. Durrant, César Augusto F. De Oliveira, and J. Andrew McCammon. POVME: An algorithm for measuring binding-pocket volumes. *Journal of Molecular Graphics and Modelling*, 29(5):773–776, 2011.

[264] Jason K. Yano, Michael R. Wester, Guillaume A. Schoch, Keith J. Griffin, C. David Stout, and Eric F. Johnson. The structure of human microsomal cytochrome P450 3A4 determined by X-ray crystallography to 2.05-Å resolution. *Journal of Biological Chemistry*, 279(37):38091–38094, 2004.

[265] Yoshitaka Hiruma, Mathias A.S. Hass, Yuki Kikui, Wei Min Liu, Betül Ölmez, Simon P. Skinner, Anneloes Blok, Alexander Kloosterman, Hiroyasu Koteishi, Frank Löhr, Harald Schwalbe, Masaki Nojiri, and Marcellus Ubbink. The structure of the cytochrome P450cam-putidaredoxin complex determined by paramagnetic NMR spectroscopy and crystallography. *Journal of Molecular Biology*, 425(22):4353–4365, 2013.

[266] Michael J. Honeychurch. The electrochemistry of cytochrome P450. What are we actually measuring? *Journal of Inorganic Biochemistry*, 96(1):151, 2003.

[267] Anuja Modi and John H. Dawson. *Peroxidase and Peroxygenase Properties and Mechanisms of Cytochrome P450*. 2015.

[268] Barry D. Fleming, Yanni Tian, Stephen G. Bell, Luet Lok Wong, Vlada Urlacher, and H. Allen O Hill. Redox properties of cytochrome P450BM3 measured by direct methods. *European Journal of Biochemistry*, 270(20):4082–4088, 2003.

[269] A. Hoeppner, L. Schmitt, and S.H.J Smits. Proteins and Their Ligands: Their Importance and How to Crystallize Them. *Advanced Topics on Crystal Growth, S.O.Ferreira (Eds.)*, pages 1–41, 2013.

APPENDICES

APPENDIX A

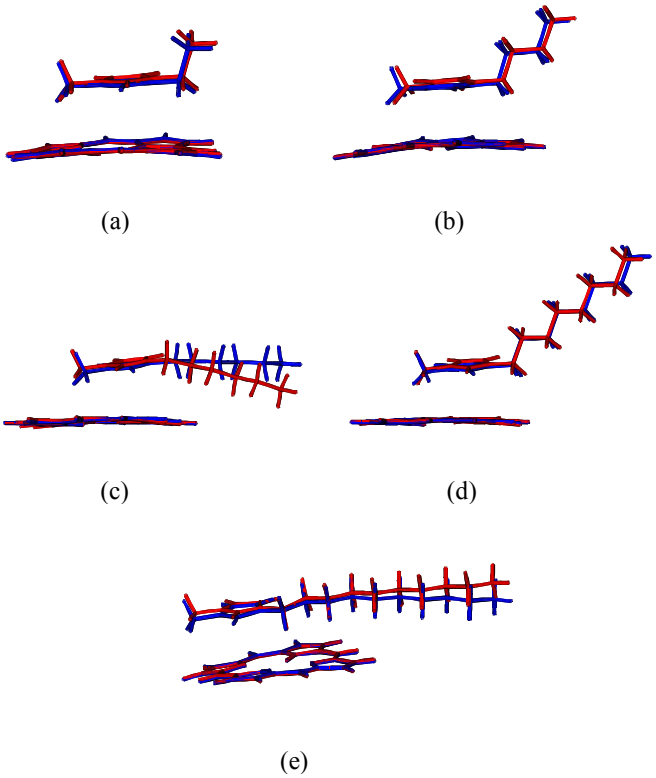


Figure A1: Alignment of $[C_n mim]^+$ FBP tail up (TU) optimized complexes at B3LYP-D2 (Blue) and M06 (Red) theories consisting of cations (a) $[C_2 mim]^+$ (b) $[C_4 mim]^+$ (c) $[C_6 mim]^+$ (d) $[C_8 mim]^+$ (e) $[C_{10} mim]^+$

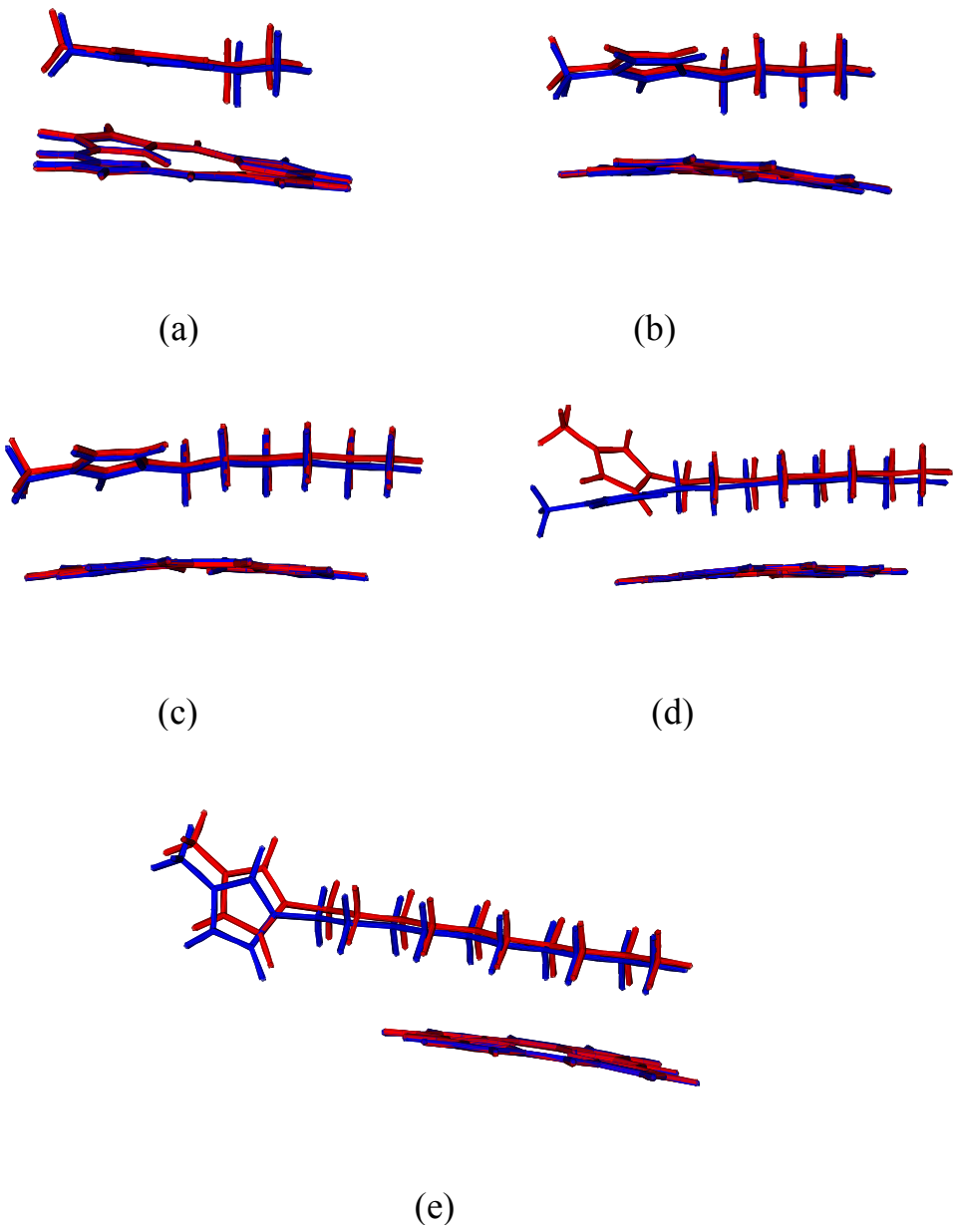


Figure A2: Alignment of $[C_n mim]^+$ FBP tail down (TD) optimized complexes at B3LYP-D2 (Blue) and M06 (Red) theories consisting of cations (a) $[C_2 mim]^+$ (b) $[C_4 mim]^+$ (c) $[C_6 mim]^+$ (d) $[C_8 mim]^+$ (e) $[C_{10} mim]^+$

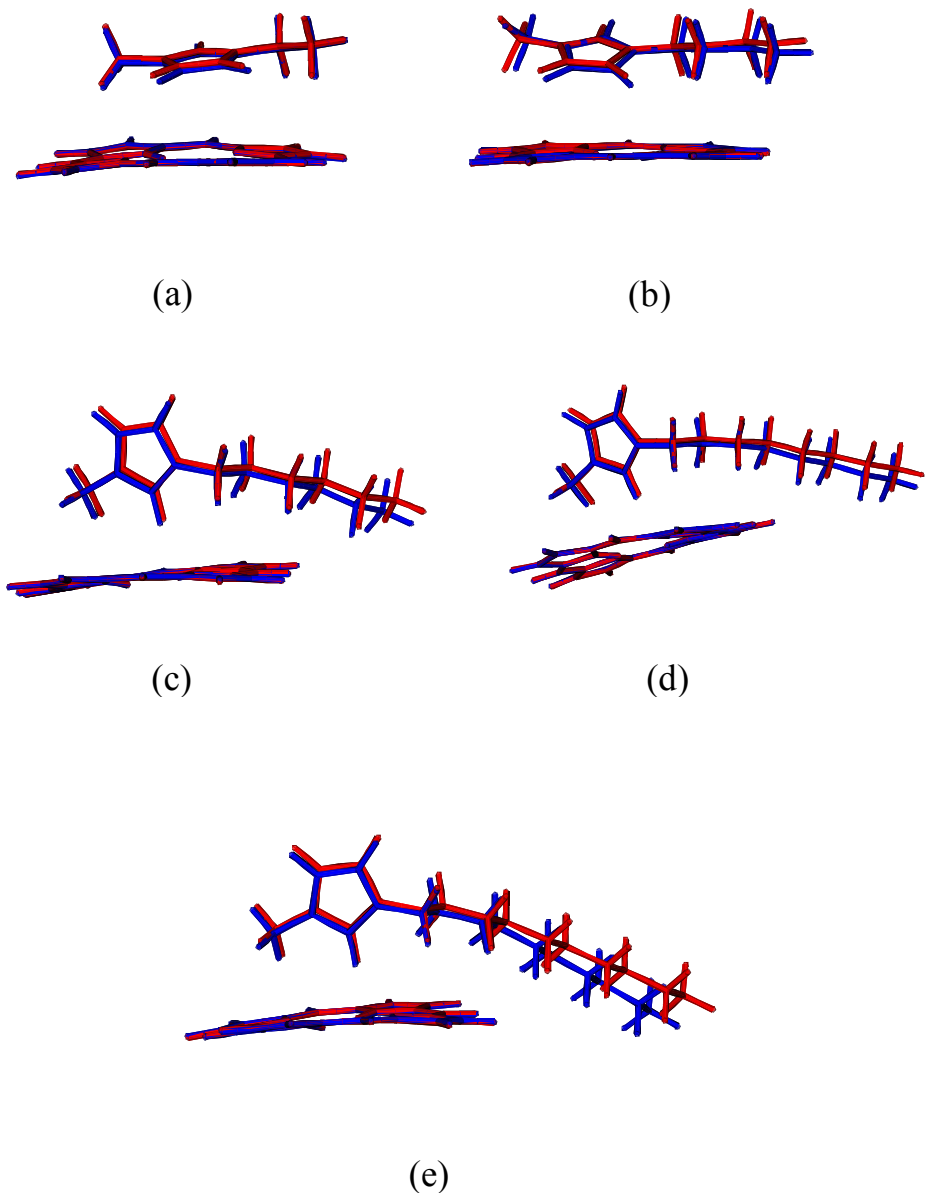


Figure A3: Alignment of $[C_n mim]^+$ FBP interplanar (IP) optimized complexes at B3LYP-D2 (Blue) and M06 (Red) theories consisting of cations (a) $[C_2 mim]^+$ (b) $[C_4 mim]^+$ (c) $[C_6 mim]^+$ (d) $[C_8 mim]^+$ (e) $[C_{10} mim]^+$

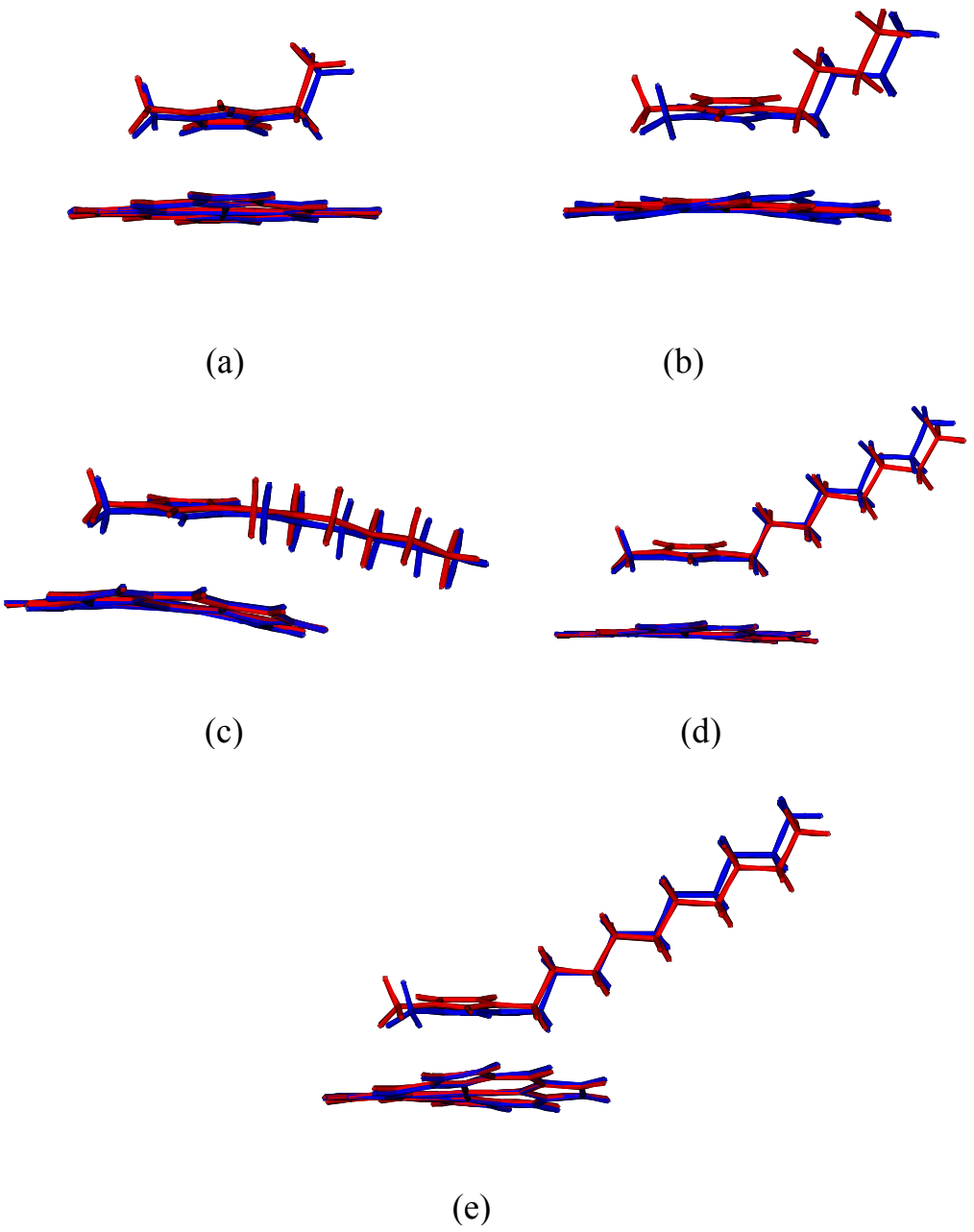


Figure A4: Alignment of $[C_n mim]^+$ FeP tail up (TU) optimized complexes at B3LYP-D2 (Blue) and M06 (Red) theories consisting of cations (a) $[C_2 mim]^+$ (b) $[C_4 mim]^+$ (c) $[C_6 mim]^+$ (d) $[C_8 mim]^+$ (e) $[C_{10} mim]^+$

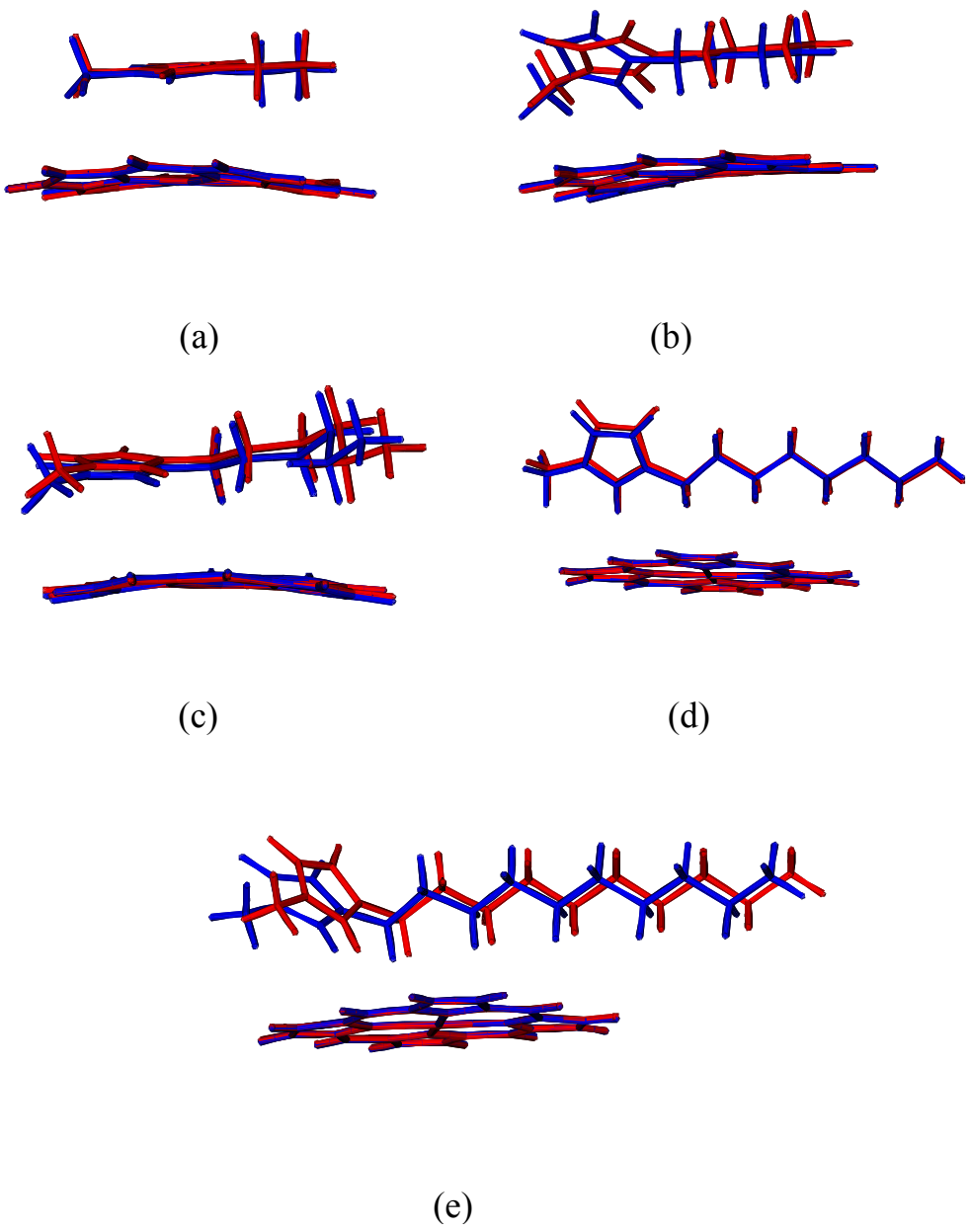
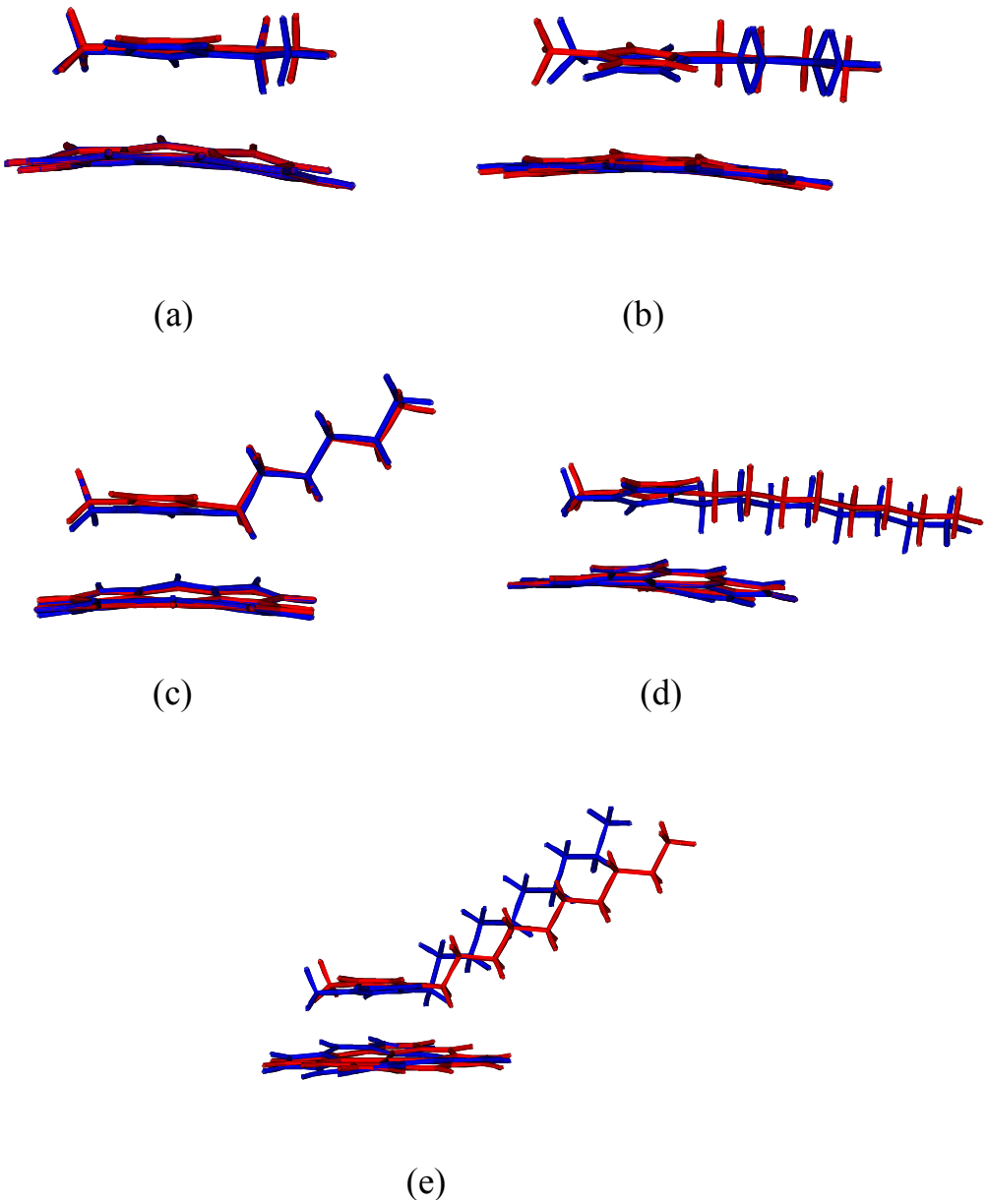


Figure A5: Alignment of $[C_n mim]^+$ FeP tail down (TD) optimized complexes at B3LYP-D2 (Blue) and M06 (Red) theories consisting of cations (a) $[C_2 mim]^+$ (b) $[C_4 mim]^+$ (c) $[C_6 mim]^+$ (d) $[C_8 mim]^+$ (e) $[C_{10} mim]^+$



center

Figure A6: Alignment of $[C_n mim]^+$ FeP interplanar (IP) optimized complexes at B3LYP-D2 (Blue) and M06 (Red) theories consisting of cations (a) $[C_2 mim]^+$ (b) $[C_4 mim]^+$ (c) $[C_6 mim]^+$ (d) $[C_8 mim]^+$ (e) $[C_{10} mim]^+$

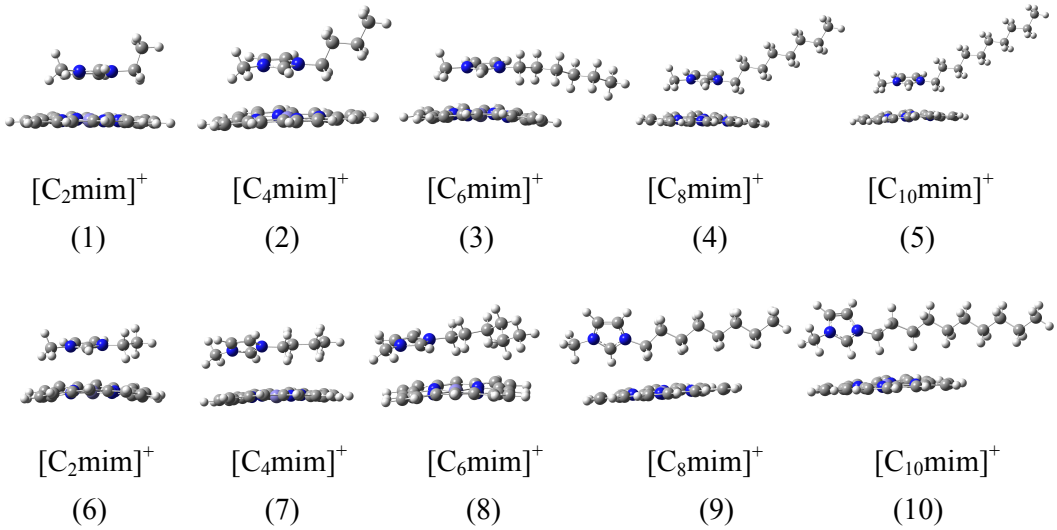


Figure A7: Tail Up (1-5) and Tail Down (6-10) optimized $[C_nmim]^+$FeP complexes optimized at M06/6-31g(d,p)

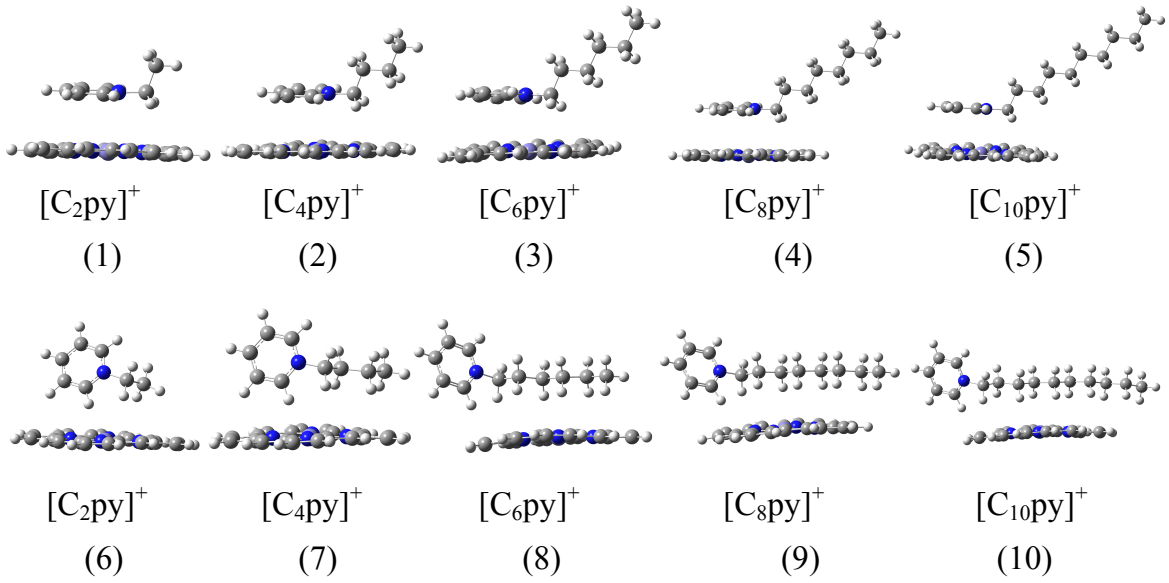


Figure A8: Tail Up (1-5) and Tail Down (6-10) optimized $[C_npy]^+$FeP complexes optimized at M06/6-31g(d,p)

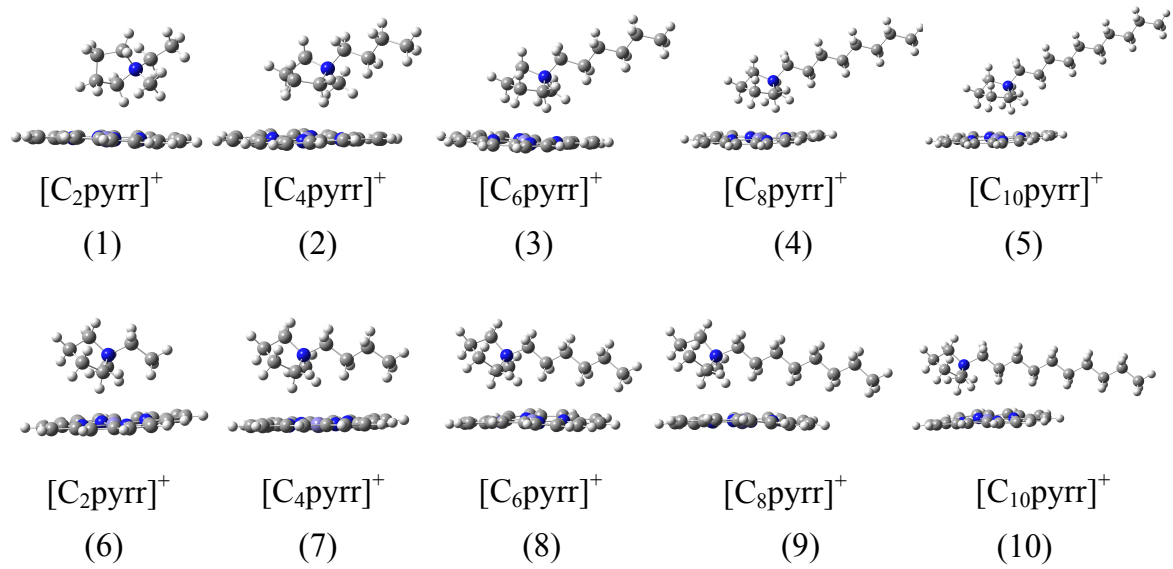


Figure A9: Tail Up (1-5) and Tail Down (6-10) optimized $[C_npyrr]^+$FeP complexes optimized at M06/6-31g(d,p)

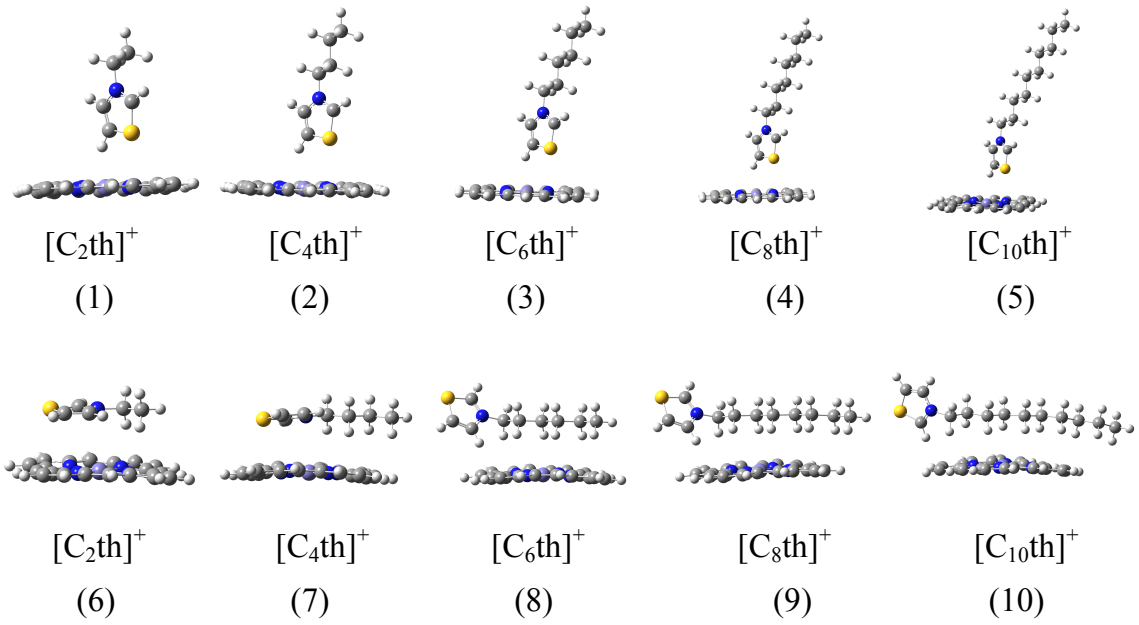


Figure A10: Tail Up (1-5) and Tail Down (6-10) optimized $[C_nth]^+$FeP complexes (ammonium) optimized at M06/6-31g(d,p)

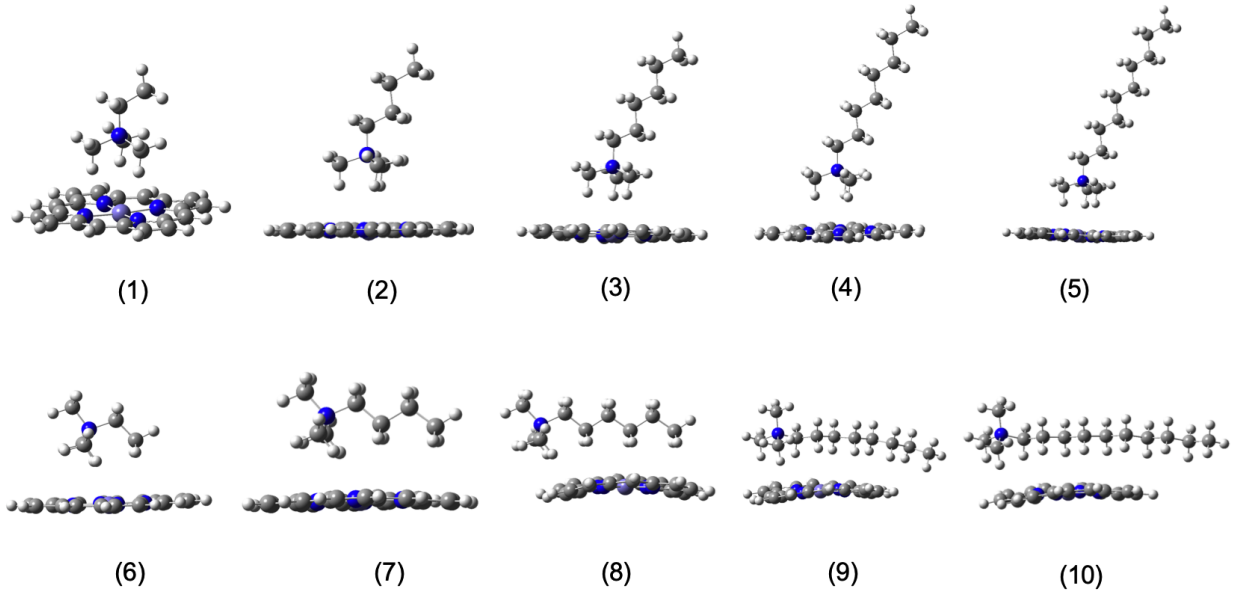


Figure A11: Tail Up (1-5) and Tail Down (6-10) optimized $[C_npho]^+FeP$ (phosphonium) complexes optimized at M06/6-31g(d,p)

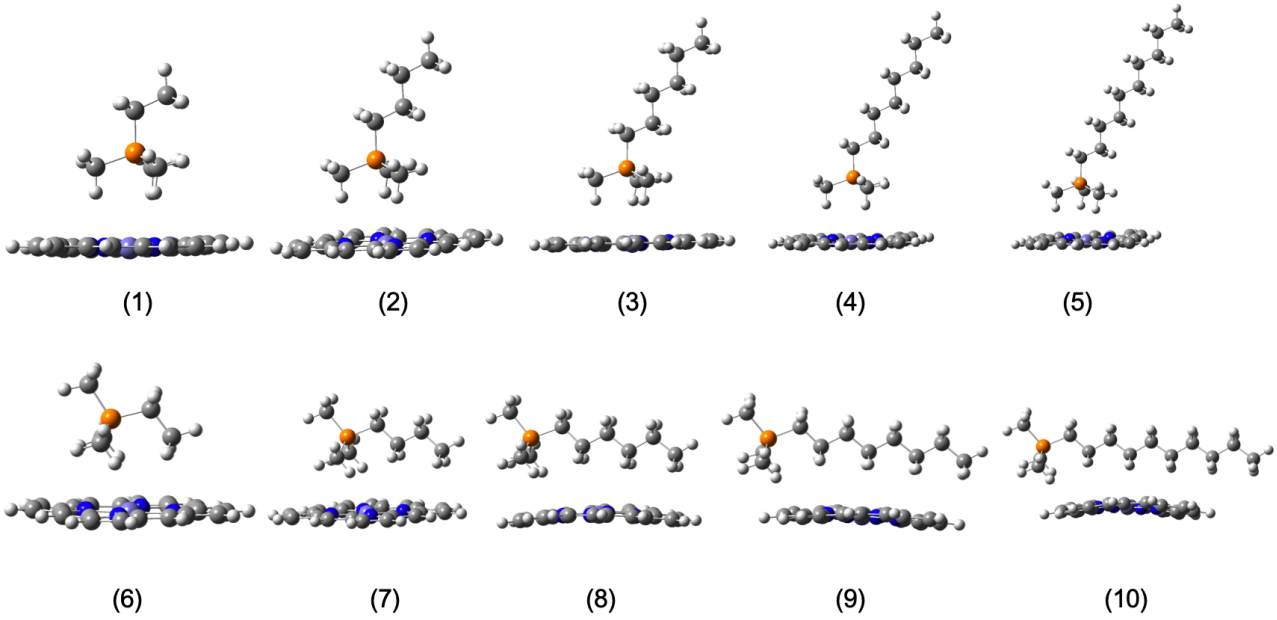


Figure A12: Tail Up (1-5) and Tail Down (6-10) optimized $[C_nsul]^+FeP$ (sulphonium) complexes optimized at M06/6-31g(d,p)

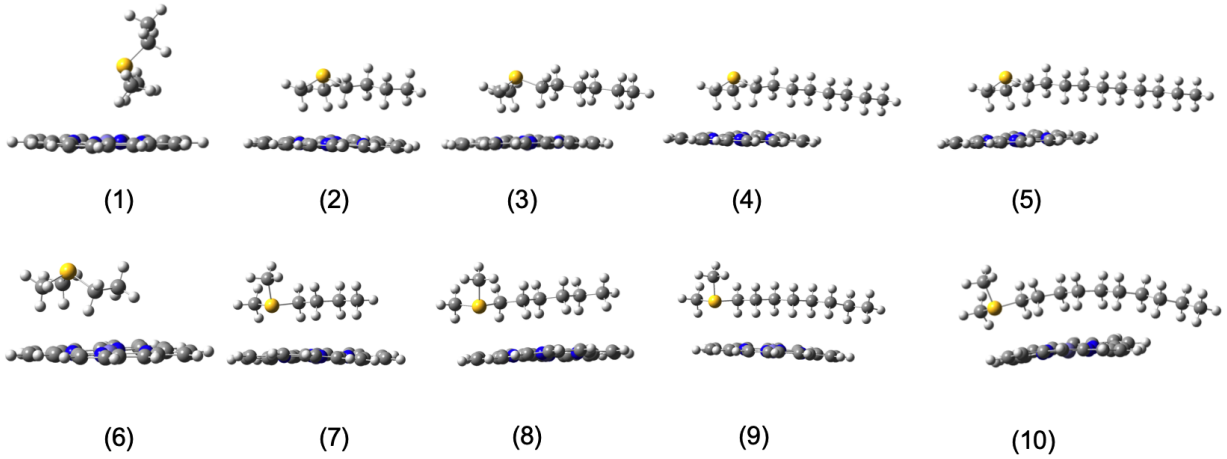


Figure A13: Tail Up (1-5) and Tail Down (6-10) optimized $[C_n amm]^+$FeP complexes optimized at M06/6-31g(d,p)

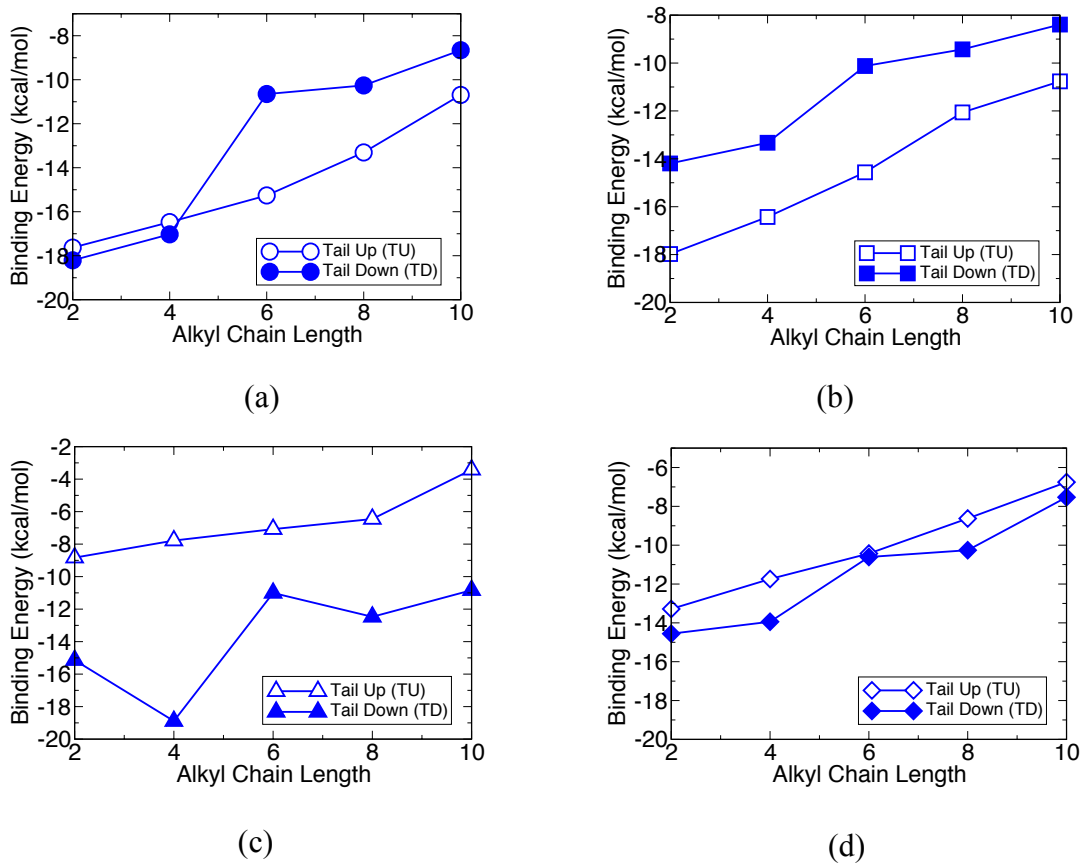


Figure A14: Binding Profiles for (a) $[C_nmim]^+FeP$ (b) $[C_npy]^+FeP$ (c) $[C_npyrr]^+FeP$ (d) $[C_nth]^+FeP$ complexes optimized at M06/6-31+g(d,p) configuration

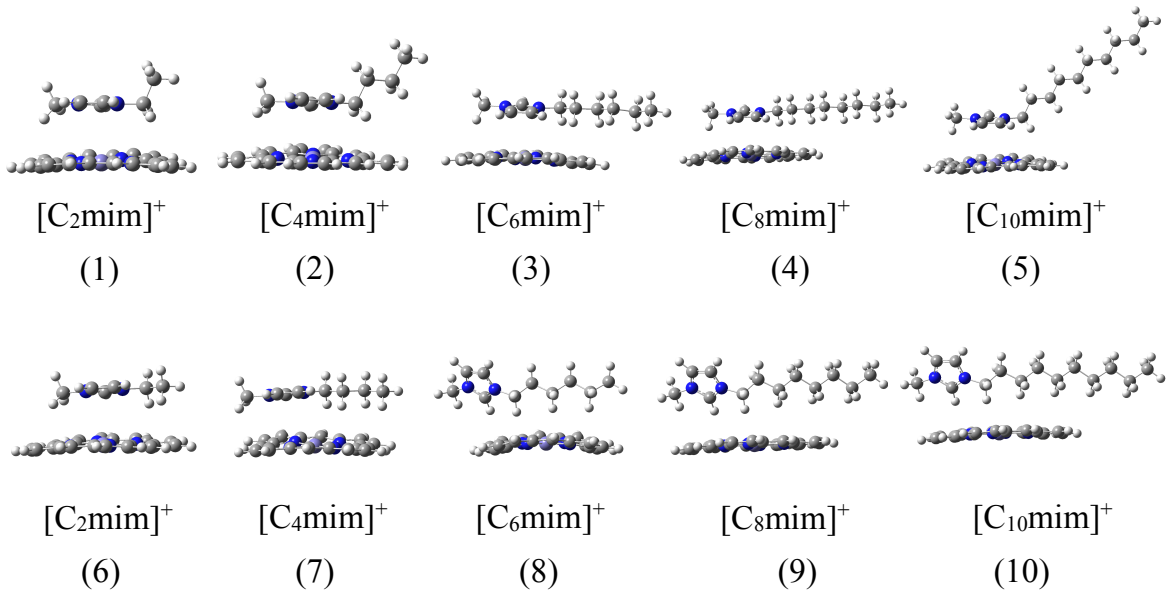


Figure A15: Tail Up (1-5) and Tail Down (6-10) optimized $[C_nmim]^+FeP$ complexes optimized at M06/6-31+g(d,p)

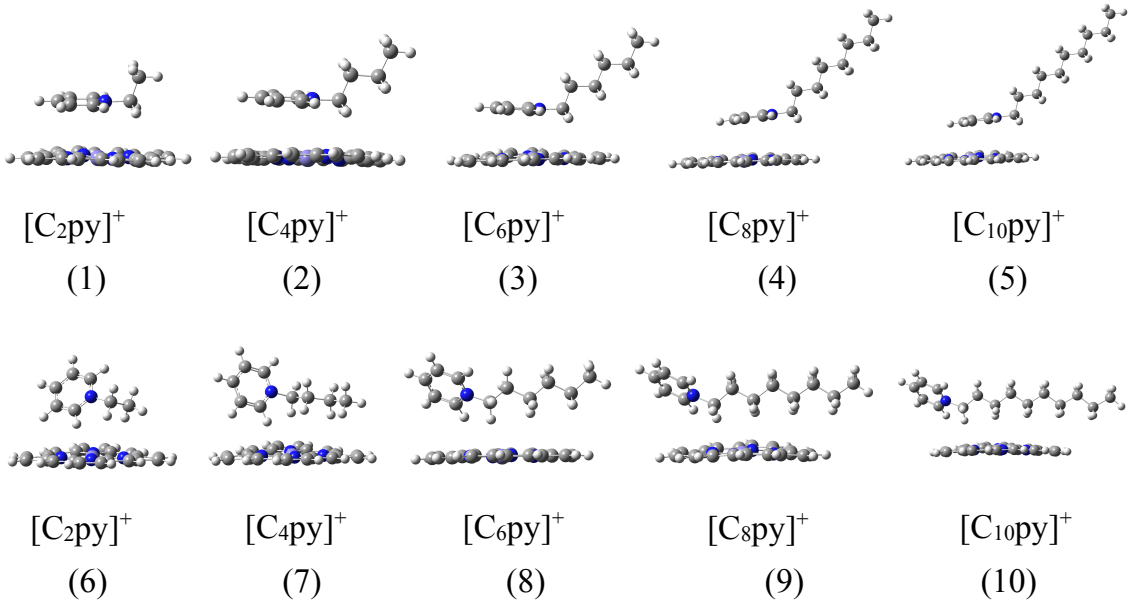


Figure A16: Tail Up (1-5) and Tail Down (6-10) optimized $[C_npy]^+FeP$ complexes optimized at M06/6-31+g(d,p)

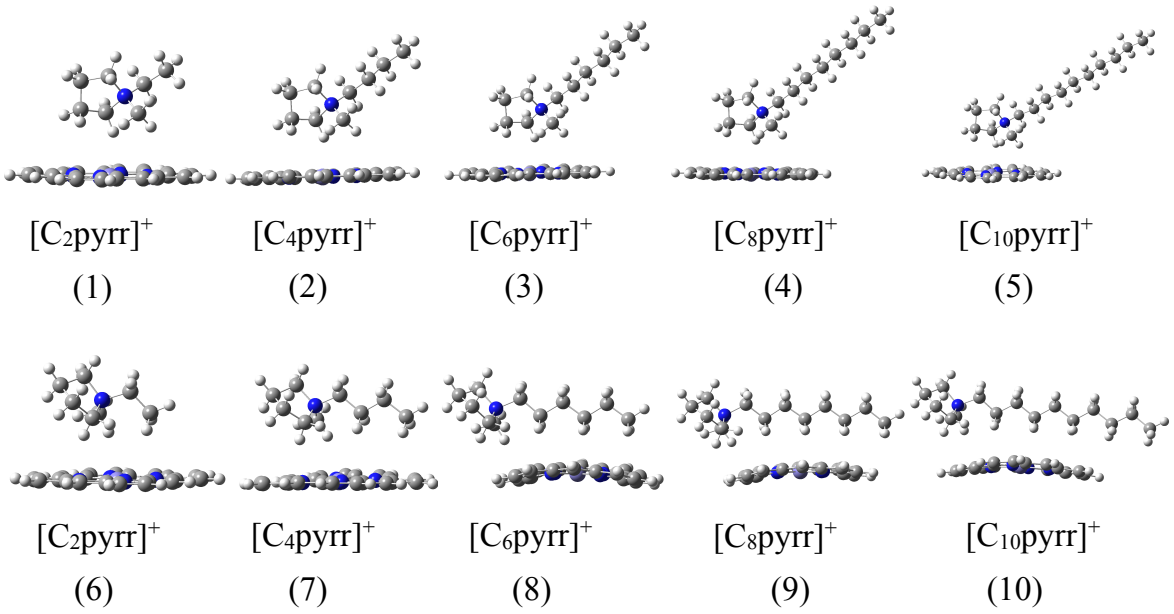


Figure A17: Tail Up (1-5) and Tail Down (6-10) optimized $[C_npyrr]^+$FeP complexes optimized at M06/6-31+g(d,p)

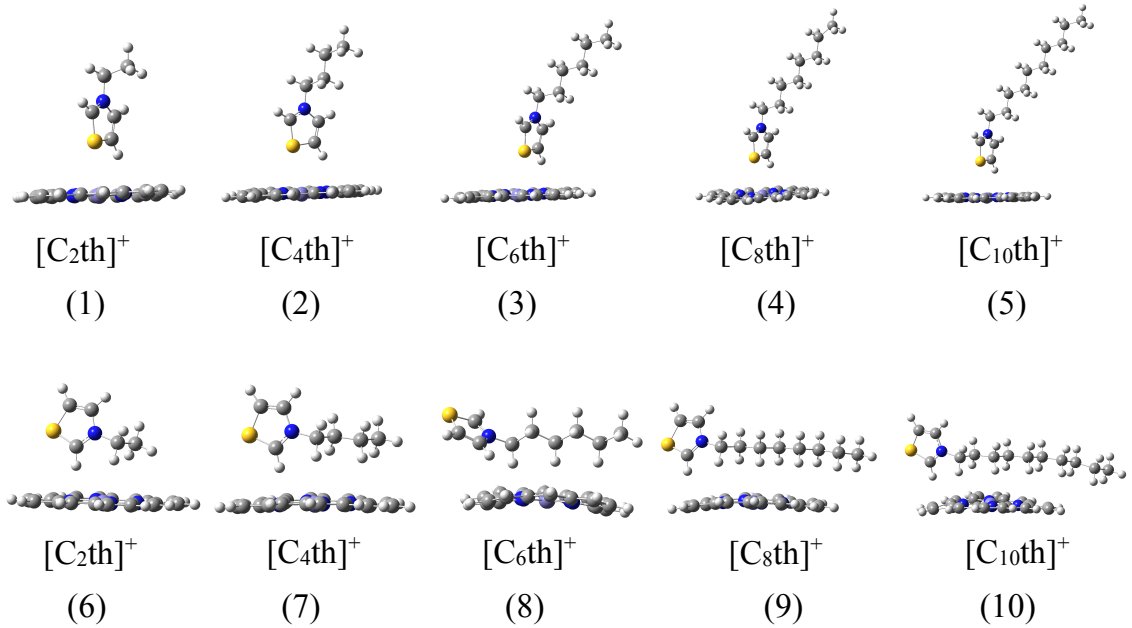


Figure A18: Tail Up (1-5) and Tail Down (6-10) optimized $[C_nth]^+$FeP complexes optimized at M06/6-31+g(d,p)

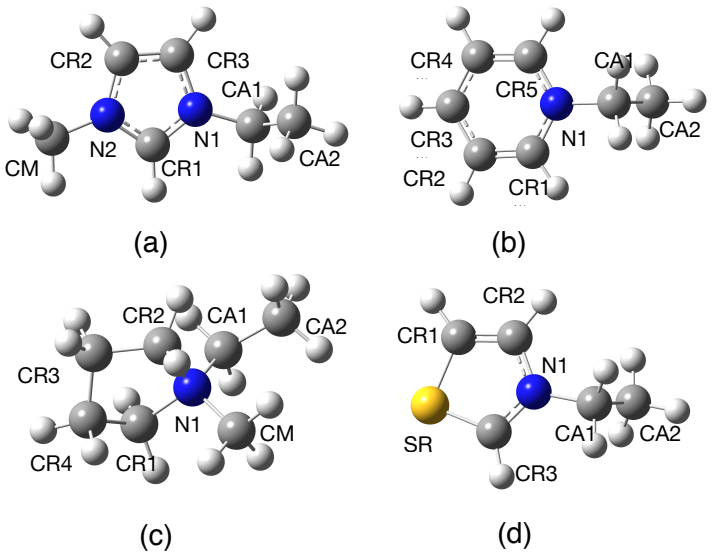
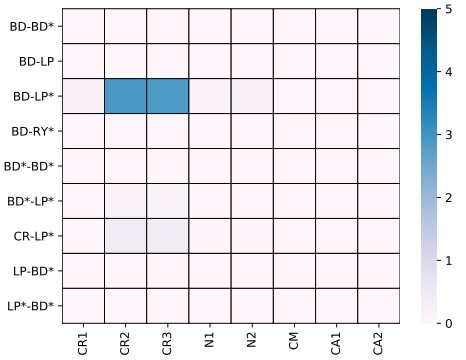
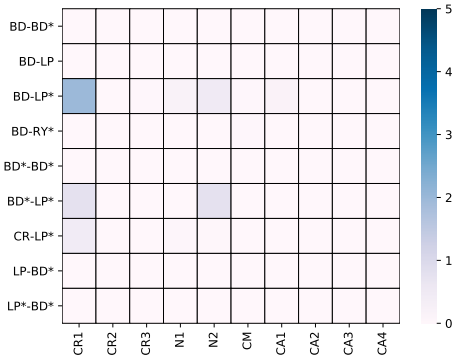


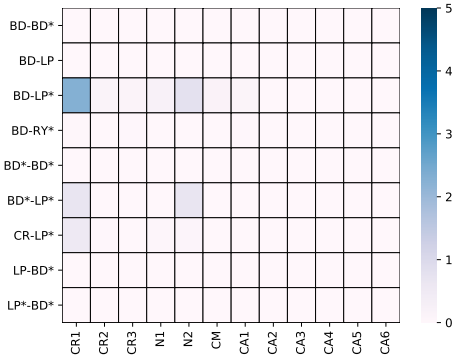
Figure A19: (a) Iron Porphyrin (FeP) and (b) 1-ethyl-3-methylimidazolium $[C_2mim]^+$ (c) 1-ethylpyridinium $[C_2py]^+$ (d) 1-ethyl-1-methylpyrrolidinium $[C_npyrr]^+$ (d) 1-ethylthiazolium $[C_2th]^+$ cations [Atom Color coding: H (White), C (Grey), N (Blue), S (yellow), Fe (purple) The longer alkyl chain cations are labeled as CA3, CA4, etc.



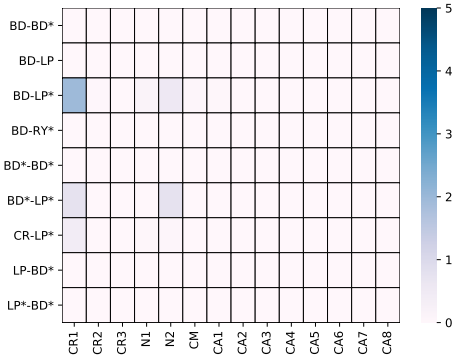
(a)



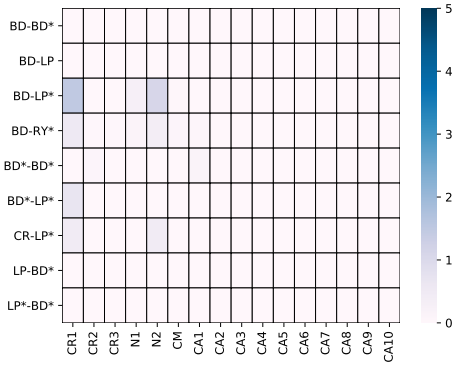
(b)



(c)

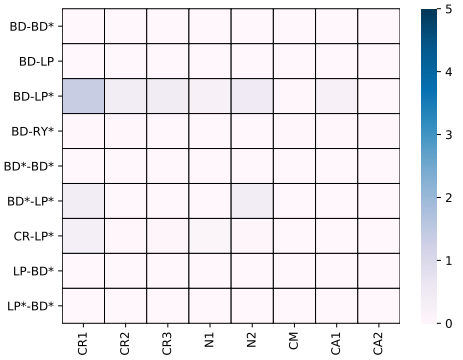


(d)

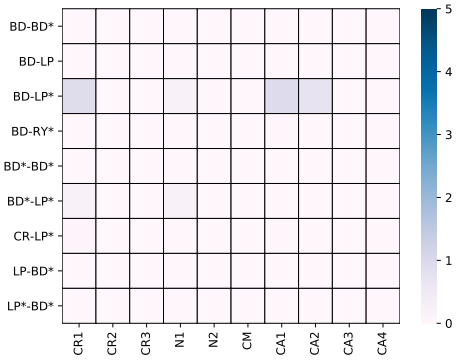


(e)

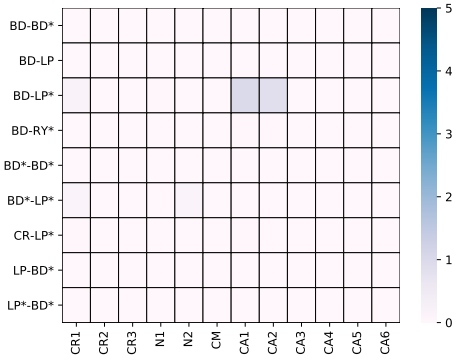
Figure A20: Stabilization energy heatmaps (in kcal/mol) for tail up imidazolium containing complexes having chain length (a) 2 (b) 4 (c) 6 (d) 8 (e) 10 with cation atoms as donor



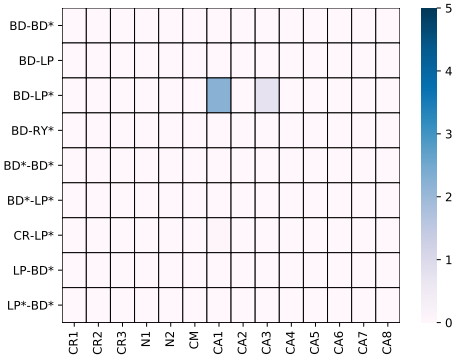
(a)



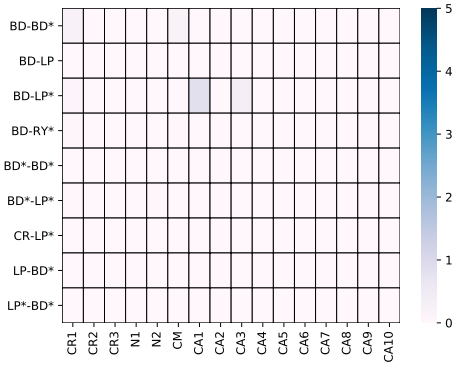
(b)



(c)

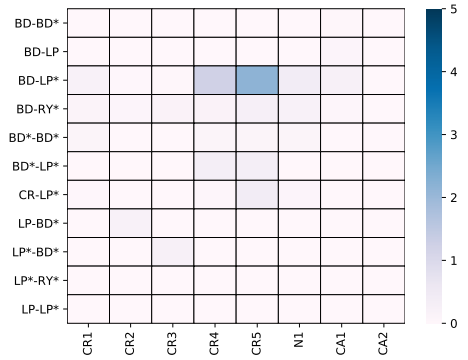


(d)

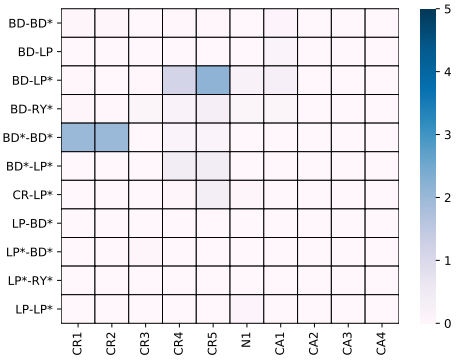


(e)

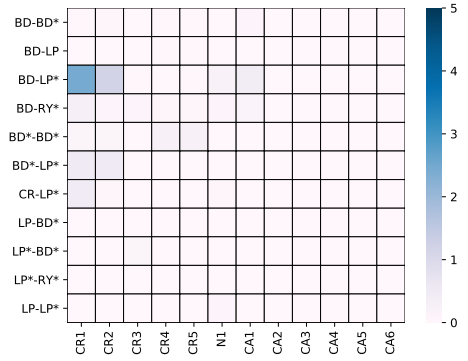
Figure A21: Stabilization energy heatmaps (in kcal/mol) for tail down imidazolium containing complexes having chain length (a) 2 (b) 4 (c) 6 (d) 8 (e) 10 with cation atoms as donor



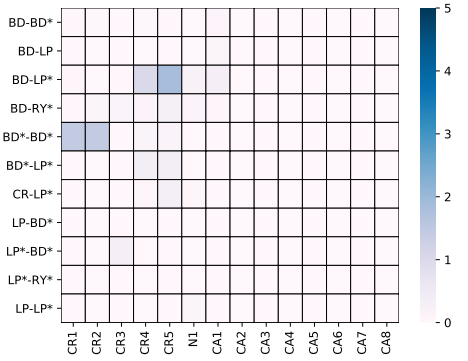
(a)



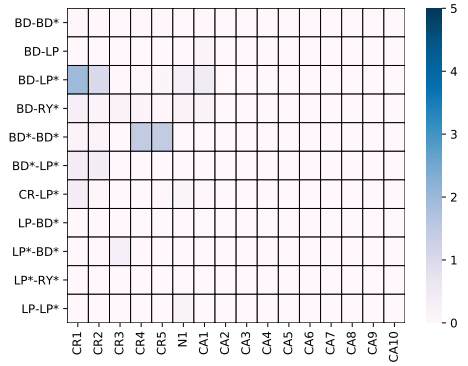
(b)



(c)

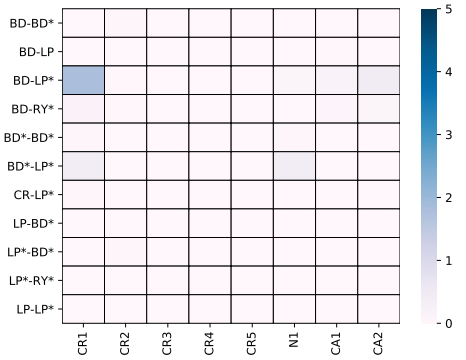


(d)

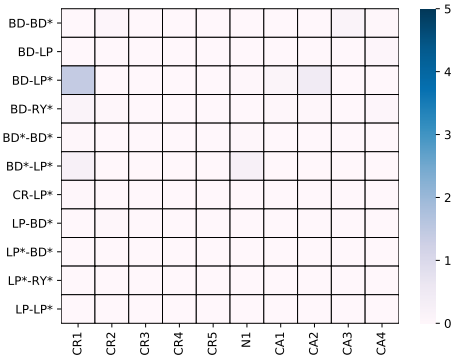


(e)

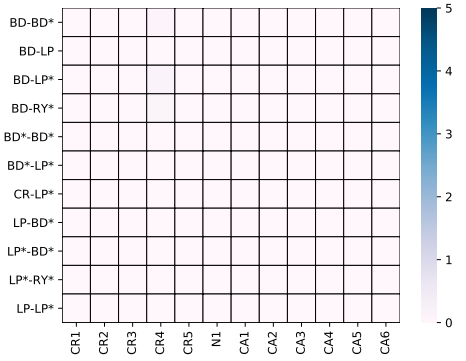
Figure A22: Stabilization energy heatmaps (in kcal/mol) for tail up pyridinium containing complexes having chain length (a) 2 (b) 4 (c) 6 (d) 8 (e) 10 with cation atoms as donor



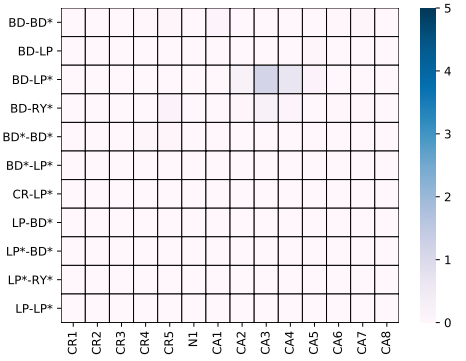
(a)



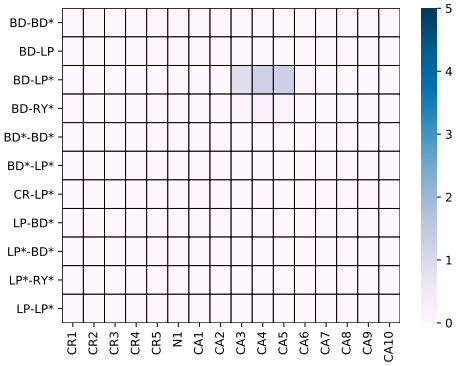
(b)



(c)

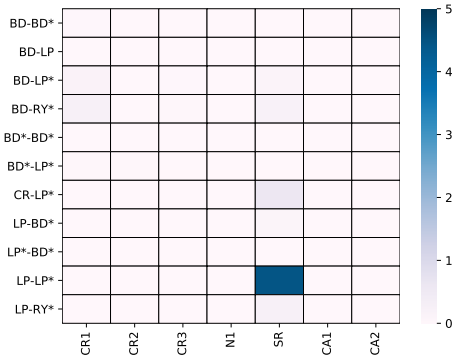


(d)

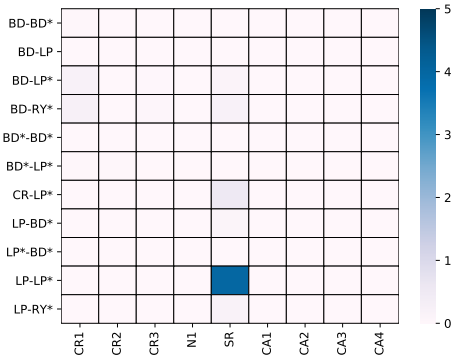


(e)

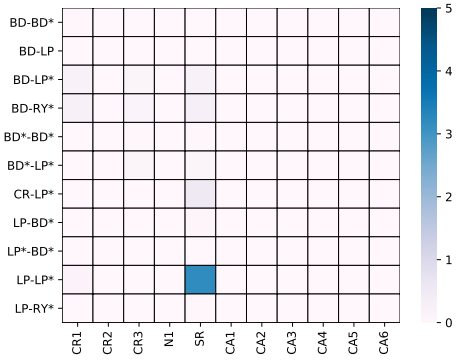
Figure A23: Stabilization energy heatmaps (in kcal/mol) for tail down pyridinium containing complexes having chain length (a) 2 (b) 4 (c) 6 (d) 8 (e) 10 with cation atoms as donor



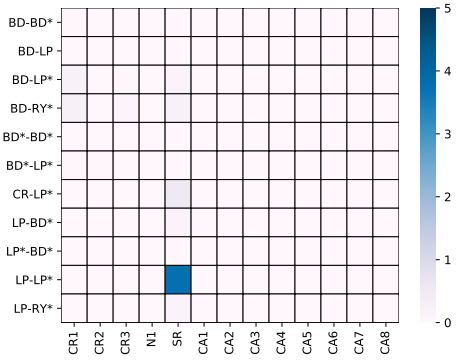
(a)



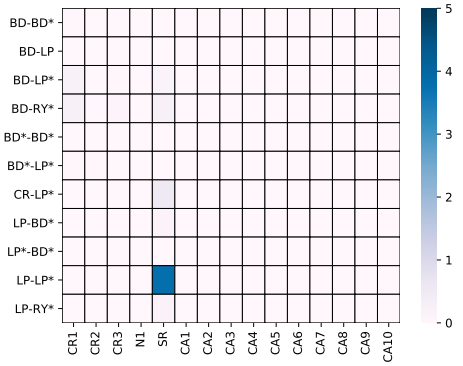
(b)



(c)

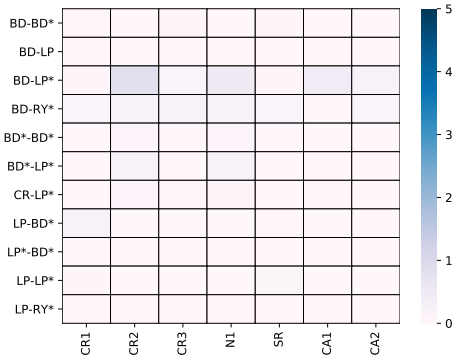


(d)

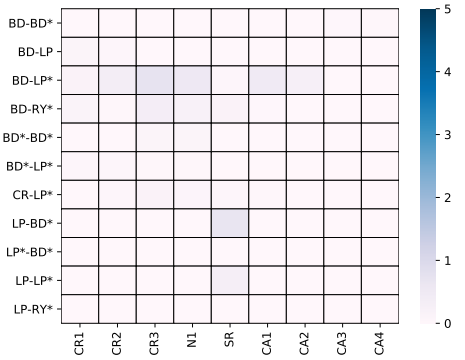


(e)

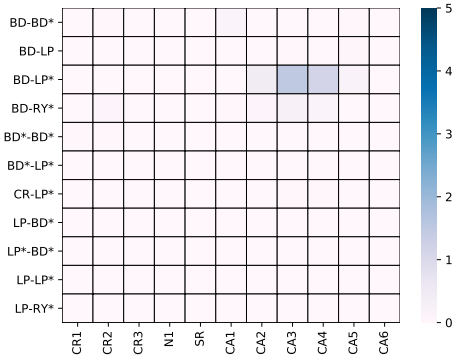
Figure A24: Stabilization energy heatmaps (in kcal/mol) for tail up thiazolium containing complexes having chain length (a) 2 (b) 4 (c) 6 (d) 8 (e) 10 with cation atoms as donor



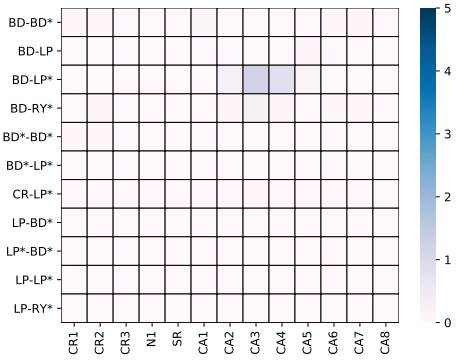
(a)



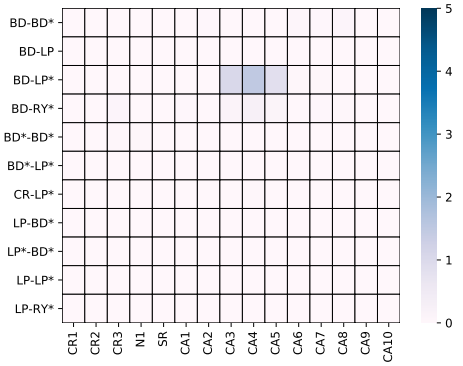
(b)



(c)

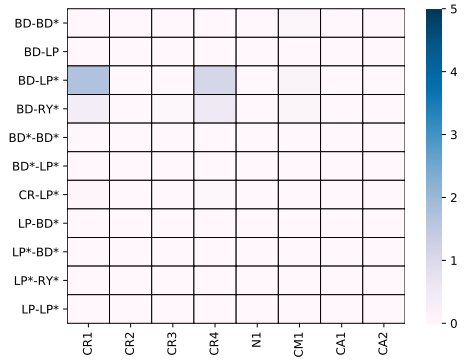


(d)

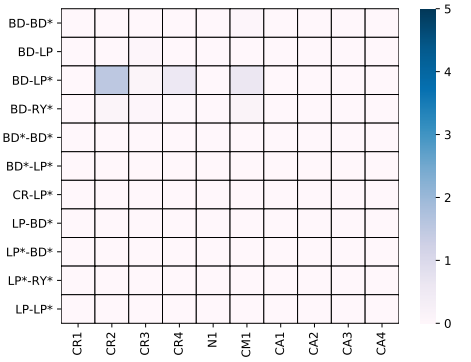


(e)

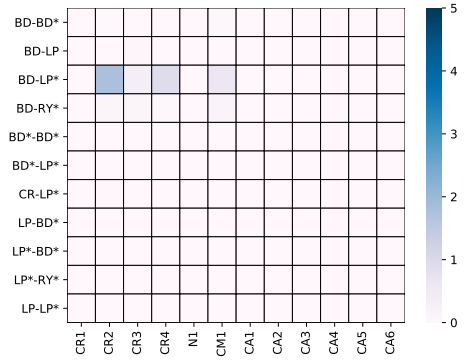
Figure A25: Stabilization energy heatmaps (in kcal/mol) for tail down thiazolium containing complexes having chain length (a) 2 (b) 4 (c) 6 (d) 8 (e) 10 with cation atoms as donor



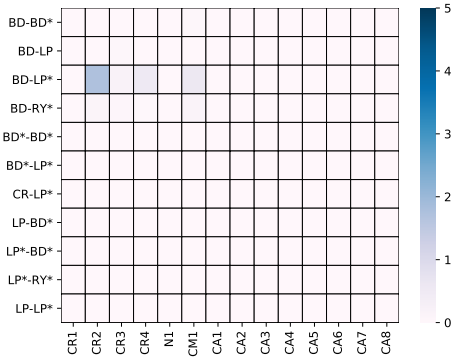
(a)



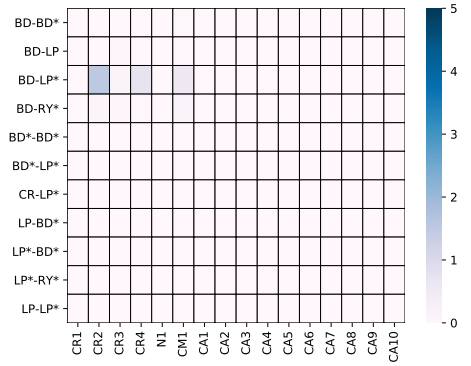
(b)



(c)



(d)



(e)

Figure A26: Stabilization energy heatmaps (in kcal/mol) for tail up pyrrolidinium containing complexes having chain length (a) 2 (b) 4 (c) 6 (d) 8 (e) 10 with cation atoms as donor

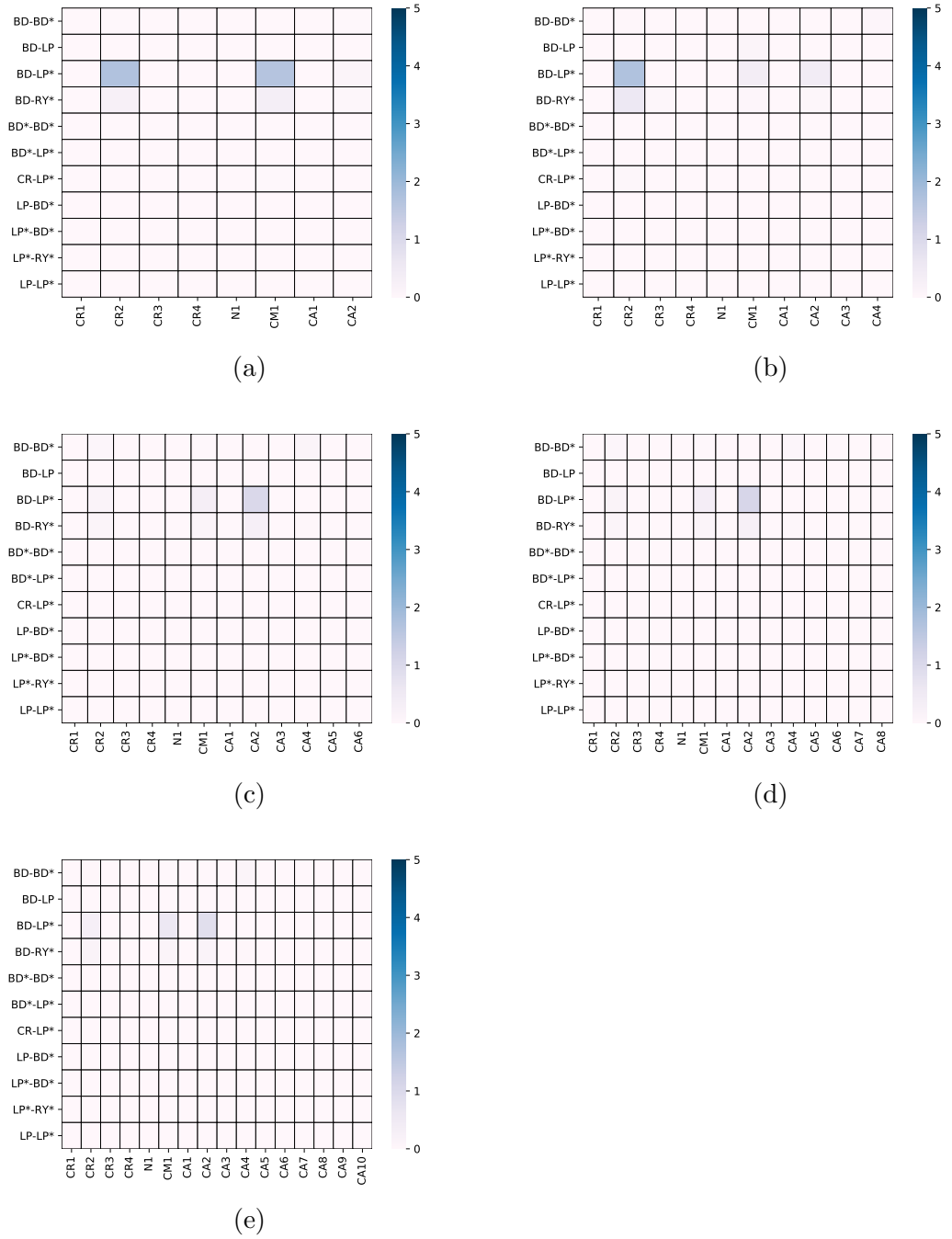
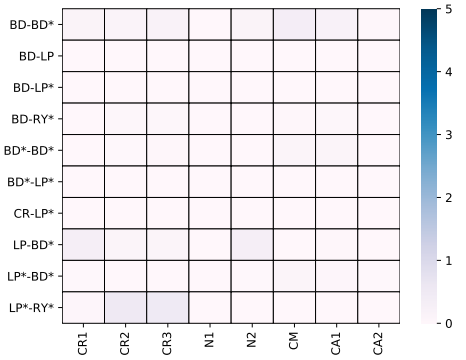
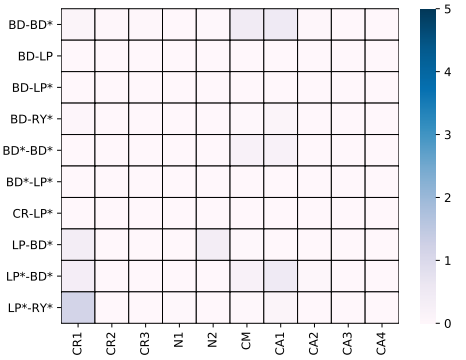


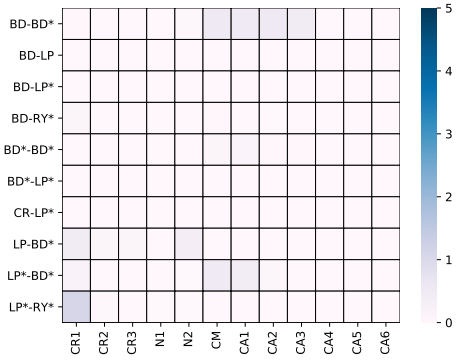
Figure A27: Stabilization energy heatmaps (in kcal/mol) for tail down pyrrolidinium containing complexes having chain length (a) 2 (b) 4 (c) 6 (d) 8 (e) 10 with cation atoms as donor



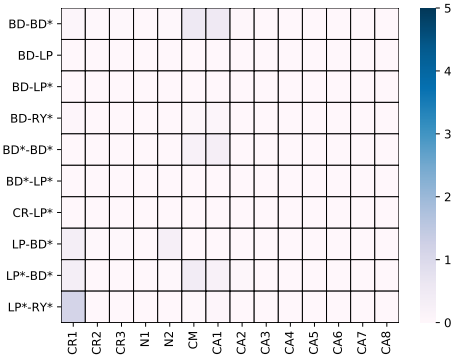
(a)



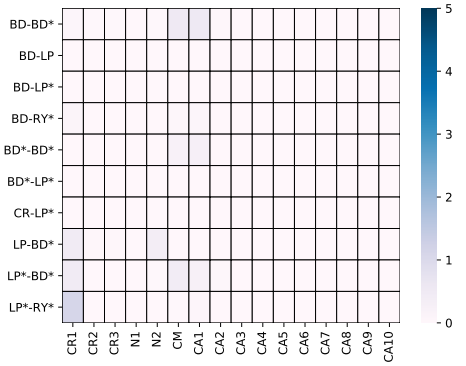
(b)



(c)

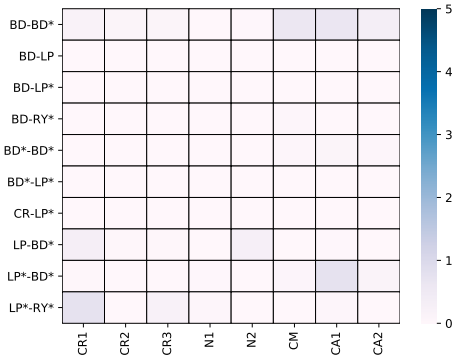


(d)

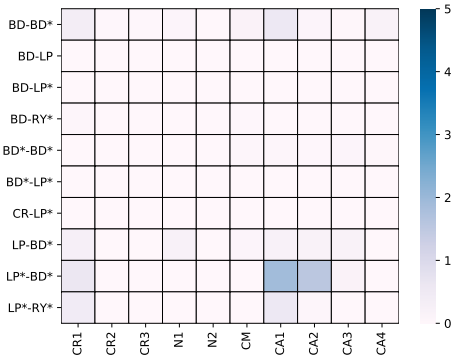


(e)

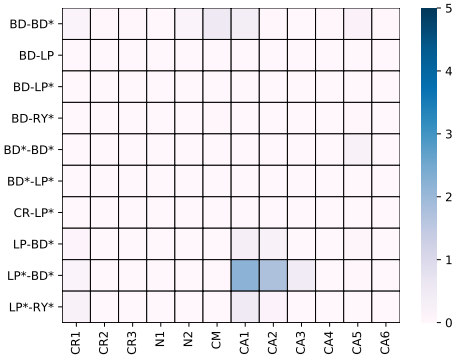
Figure A28: Stabilization energy heatmaps (in kcal/mol) for tail up imidazolium containing complexes having chain length (a) 2 (b) 4 (c) 6 (d) 8 (e) 10 with FeP atoms as donor



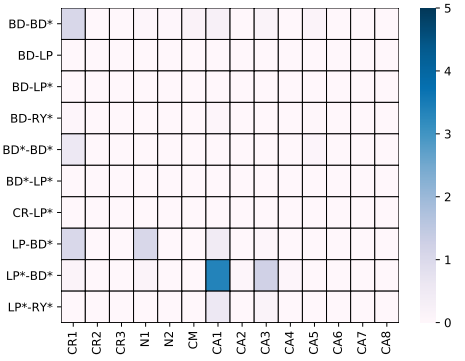
(a)



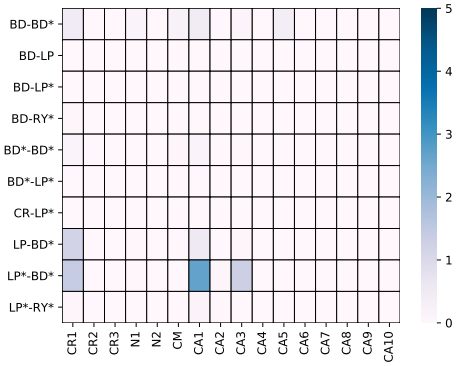
(b)



(c)



(d)



(e)

Figure A29: Stabilization energy heatmaps (in kcal/mol) for tail down imidazolium containing complexes having chain length (a) 2 (b) 4 (c) 6 (d) 8 (e) 10 with FeP atoms as donor

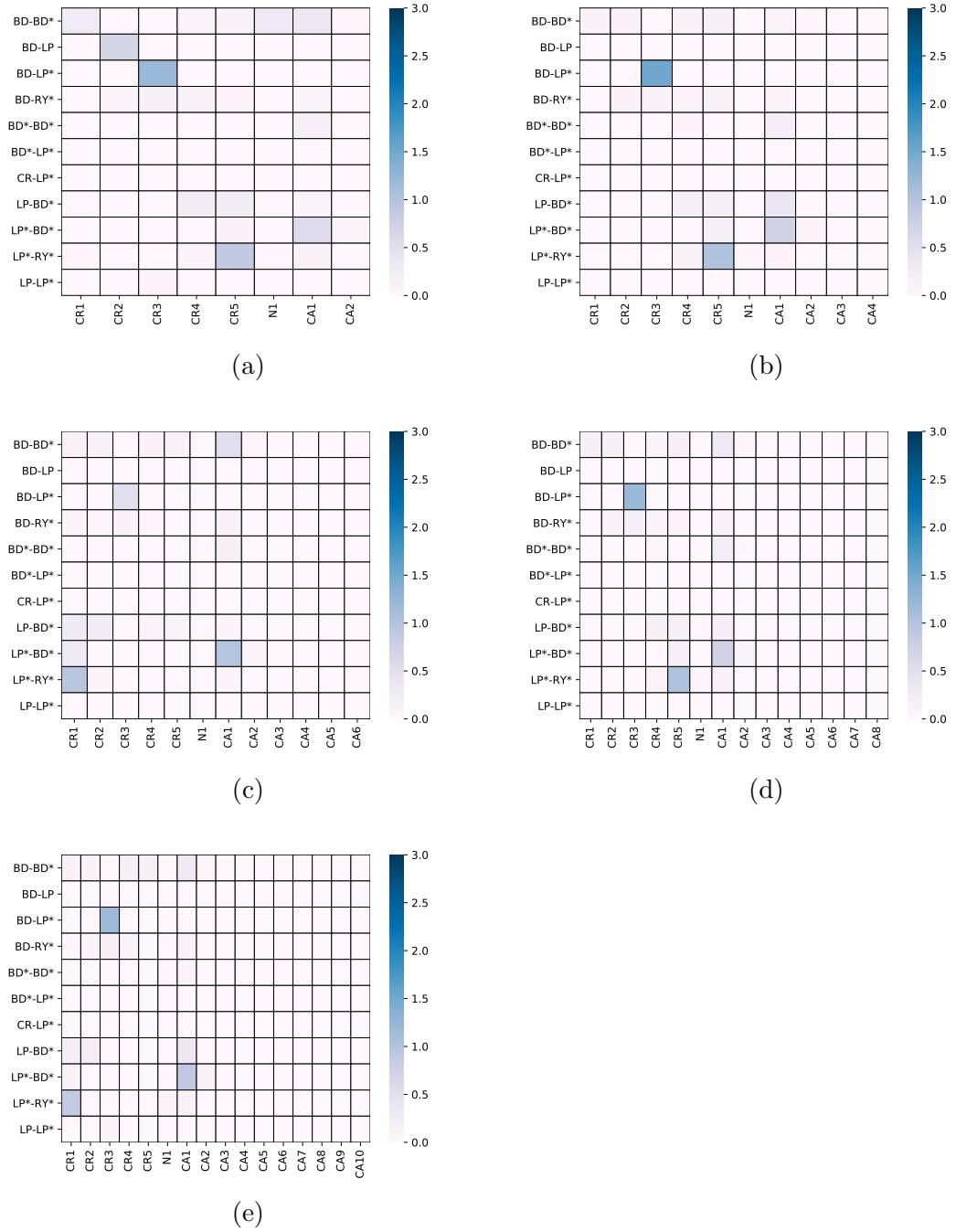


Figure A30: Stabilization energy heatmaps (in kcal/mol) for tail up pyridinium containing complexes having chain length (a) 2 (b) 4 (c) 6 (d) 8 (e) 10 with FeP atoms as donor

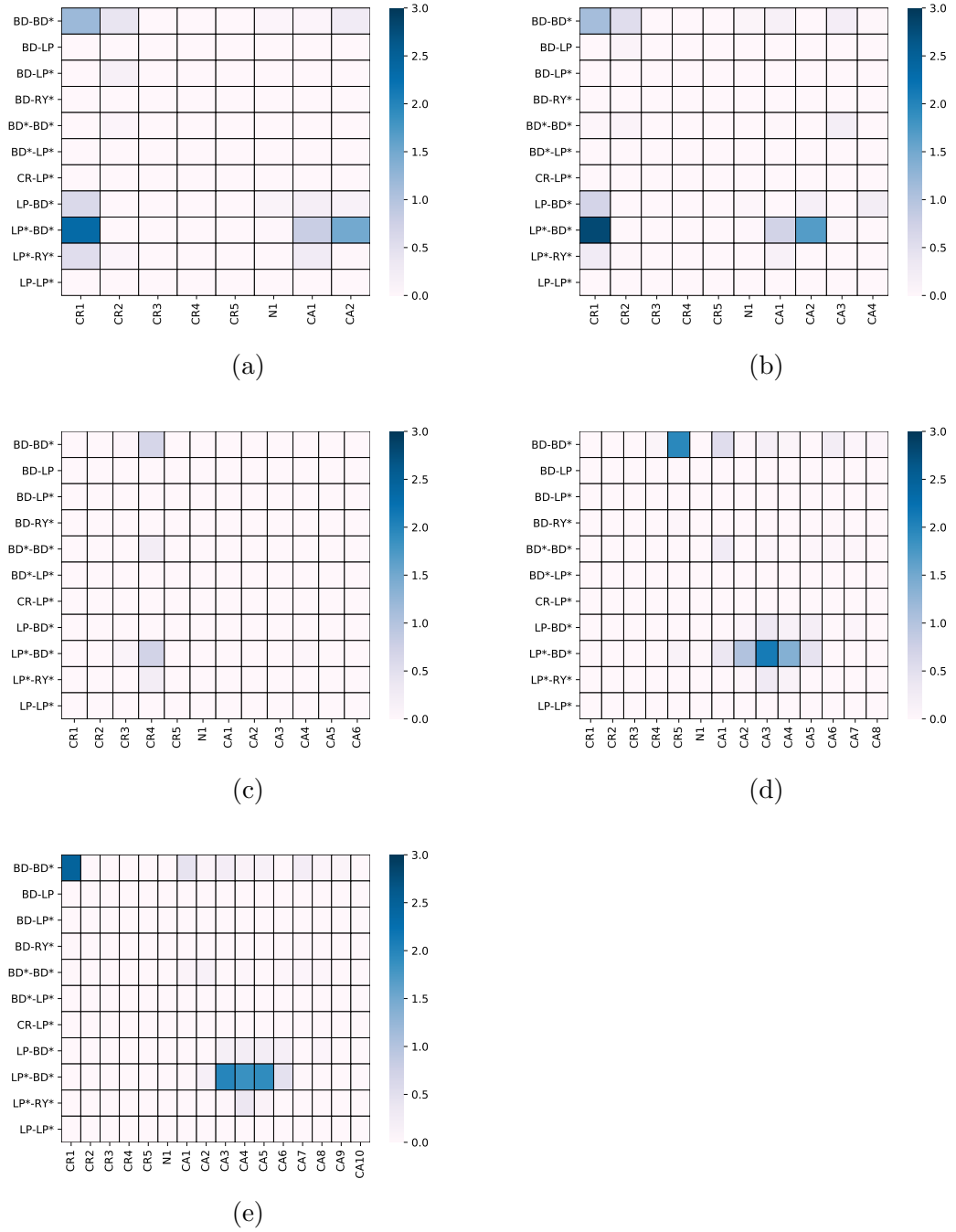
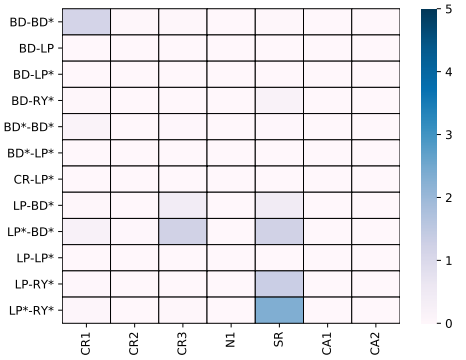
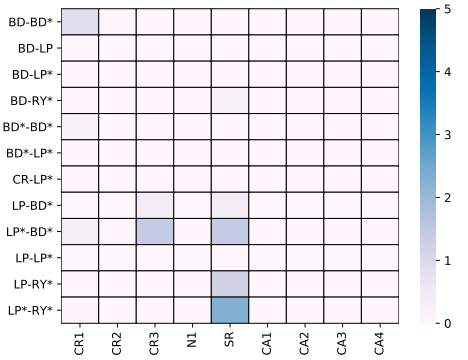


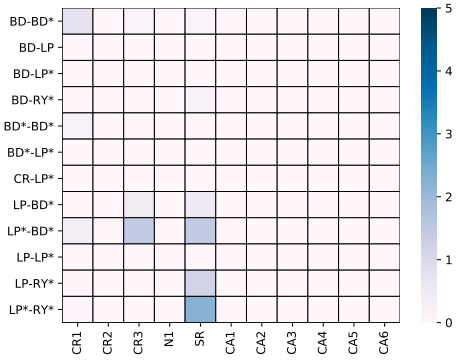
Figure A31: Stabilization energy heatmaps (in kcal/mol) for tail down pyridinium containing complexes having chain length (a) 2 (b) 4 (c) 6 (d) 8 (e) 10 with FeP atoms as donor



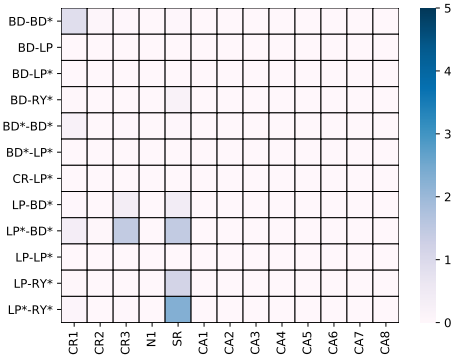
(a)



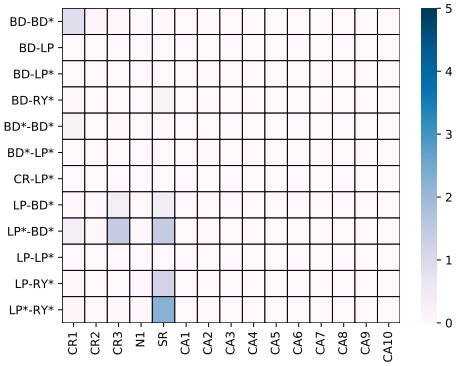
(b)



(c)

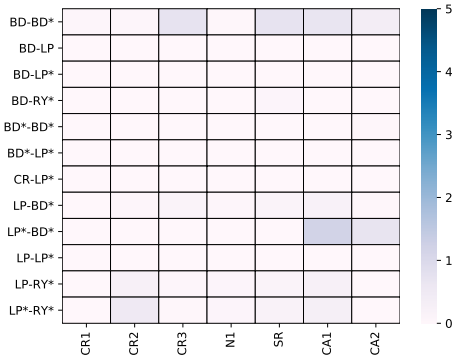


(d)

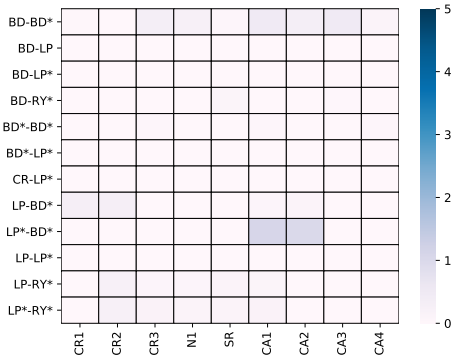


(e)

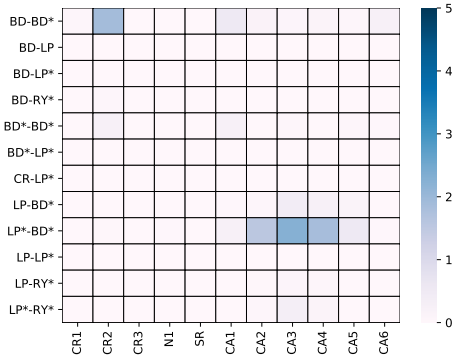
Figure A32: Stabilization energy heatmaps (in kcal/mol) for tail up thiazolium containing complexes having chain length (a) 2 (b) 4 (c) 6 (d) 8 (e) 10 with FeP atoms as donor



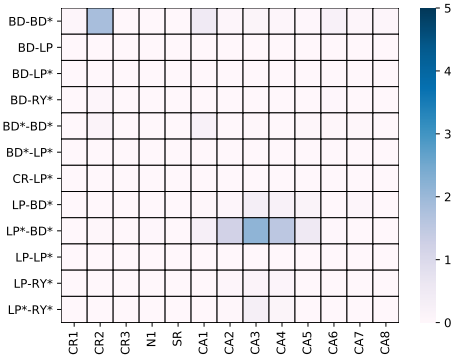
(a)



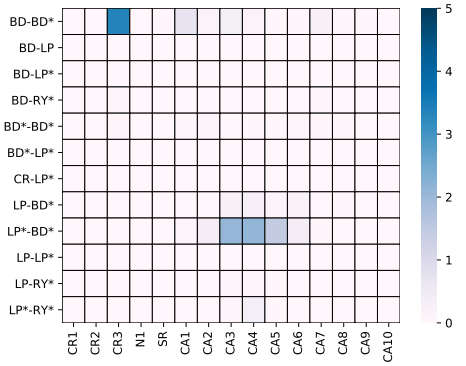
(b)



(c)

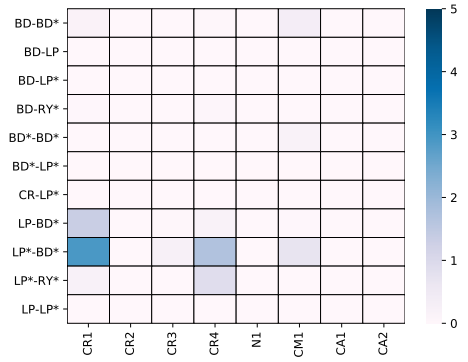


(d)

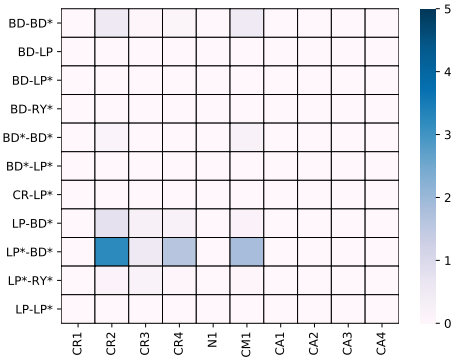


(e)

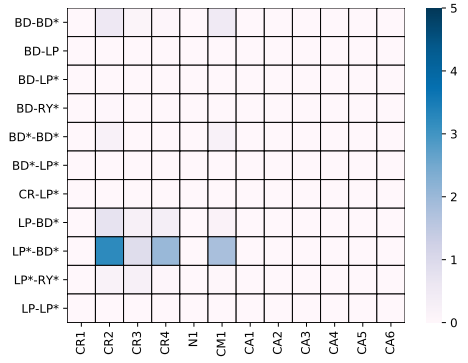
Figure A33: Stabilization energy heatmaps (in kcal/mol) for tail down thiazolium containing complexes having chain length (a) 2 (b) 4 (c) 6 (d) 8 (e) 10 with FeP atoms as donor



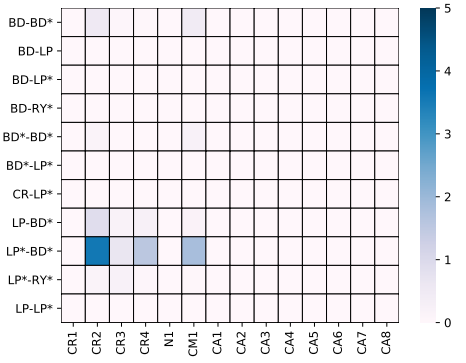
(a)



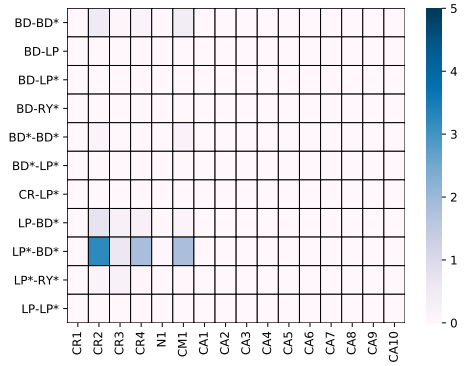
(b)



(c)



(d)



(e)

Figure A34: Stabilization energy heatmaps (in kcal/mol) for tail up pyrrolidinium containing complexes having chain length (a) 2 (b) 4 (c) 6 (d) 8 (e) 10 with FeP atoms as donor

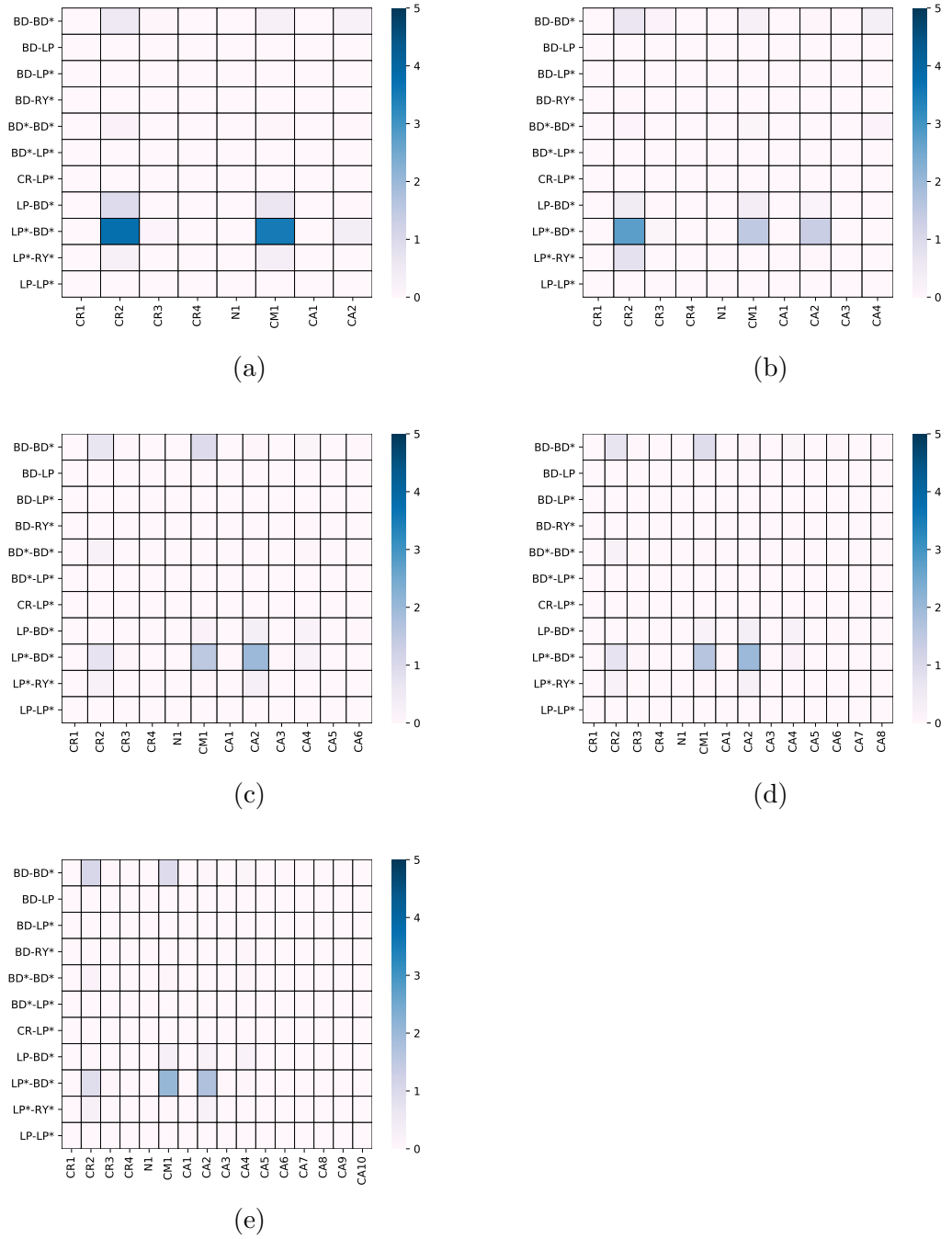


Figure A35: Stabilization energy heatmaps (in kcal/mol) for tail down pyrrolidinium containing complexes having chain length (a) 2 (b) 4 (c) 6 (d) 8 (e) 10 with FeP atoms as donor

APPENDIX B

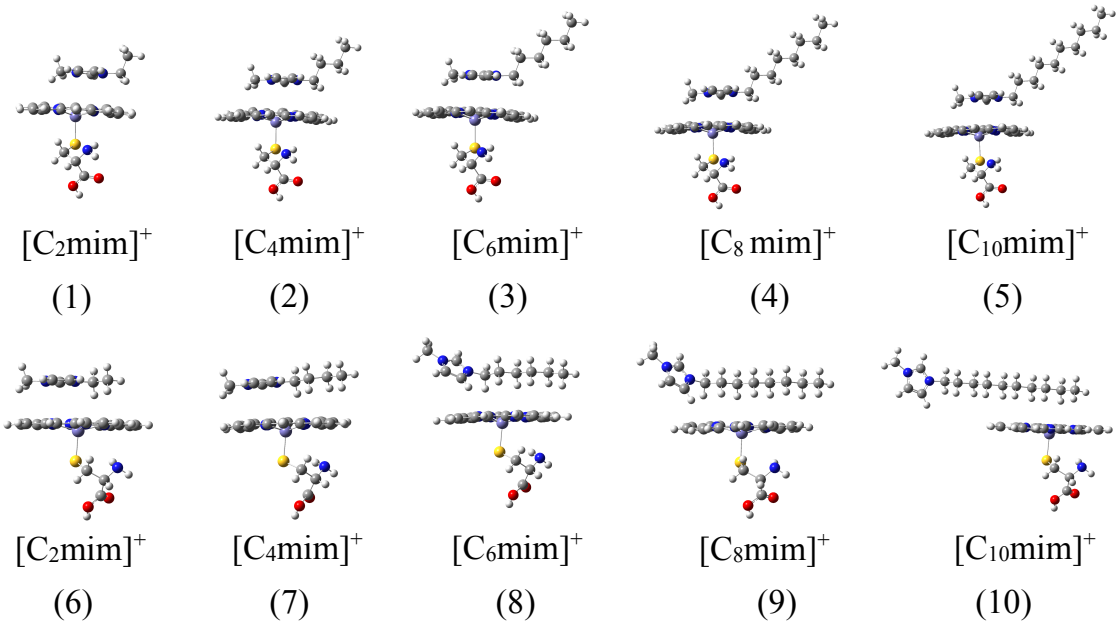


Figure A36: Optimized geometries of S_2 complexes in gas phase; TU (1-5) and TD (6-10)

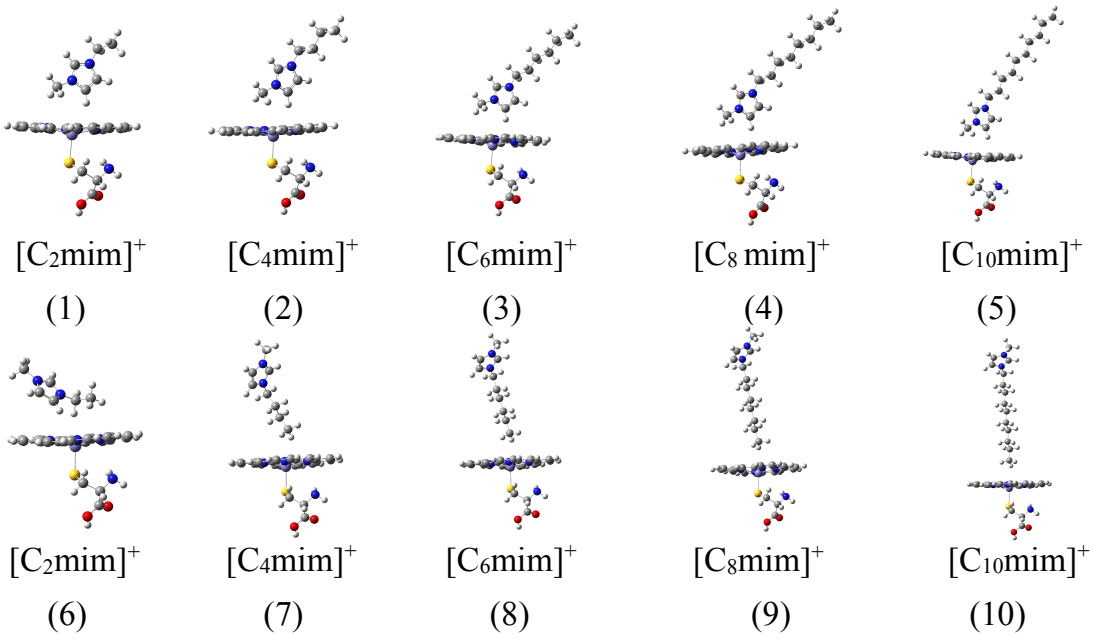


Figure A37: Optimized geometries of S_2 complexes in embedded phase; TU (1-5) and TD (6-10)

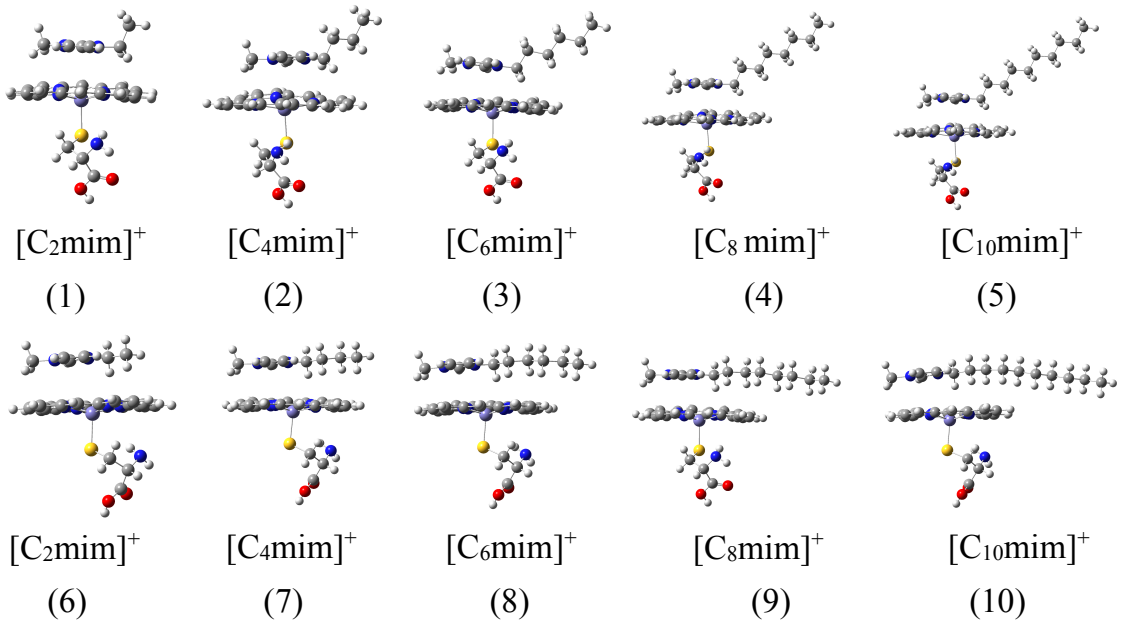


Figure A38: Optimized geometries of S_3 complexes in gas phase; TU (1-5) and TD (6-10)

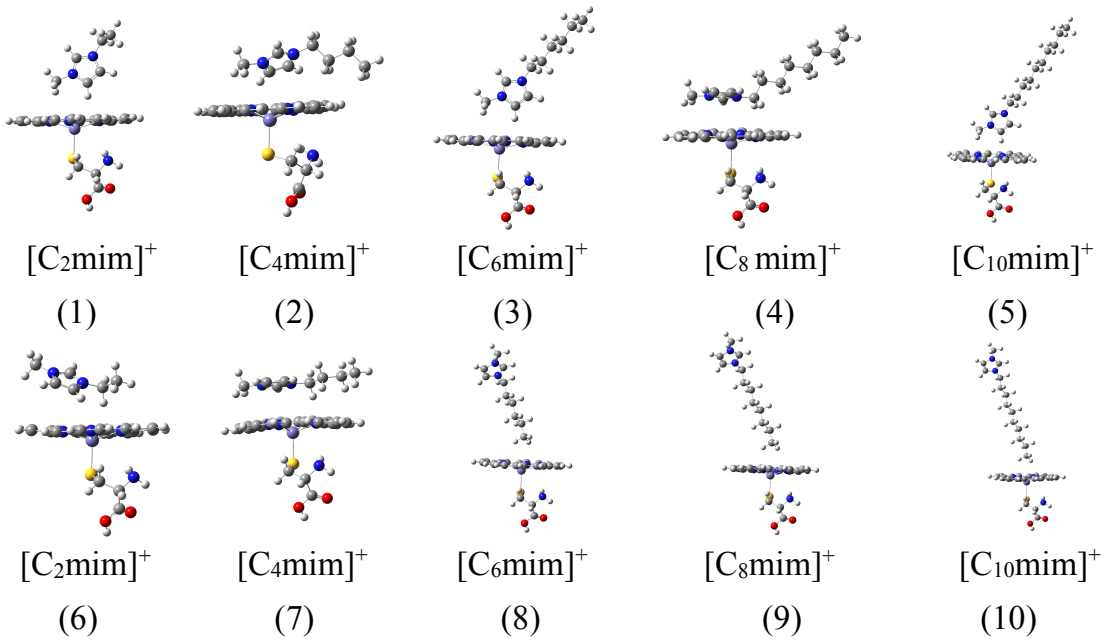


Figure A39: Optimized geometries of S_3 complexes in embedded phase; TU (1-5) and TD (6-10)

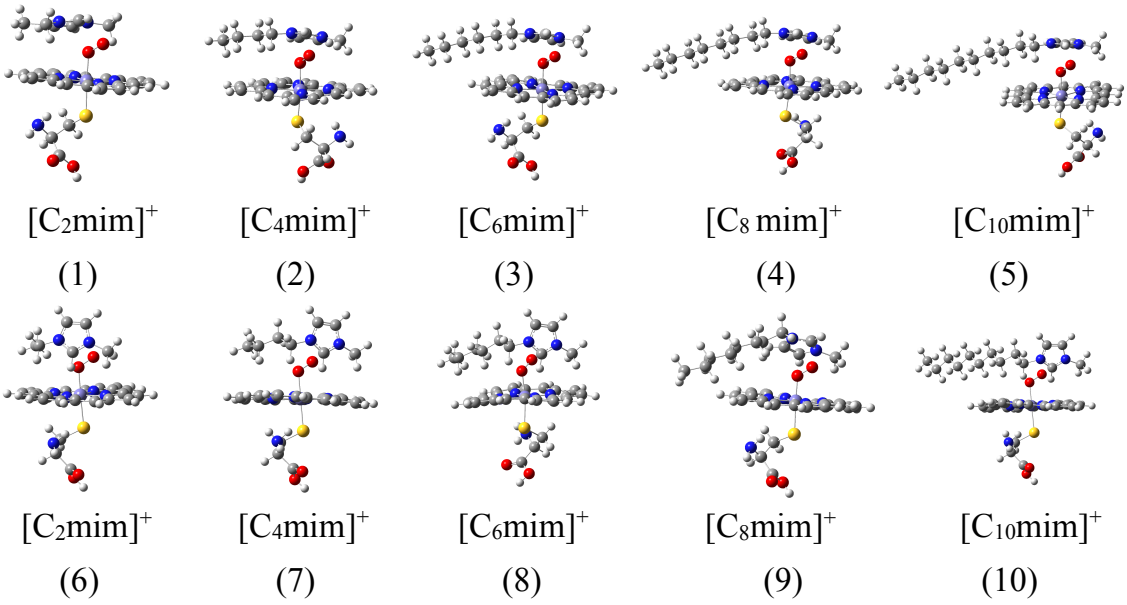


Figure A40: Optimized geometries of S_4 complexes in gas phase; TU (1-5) and TD (6-10)

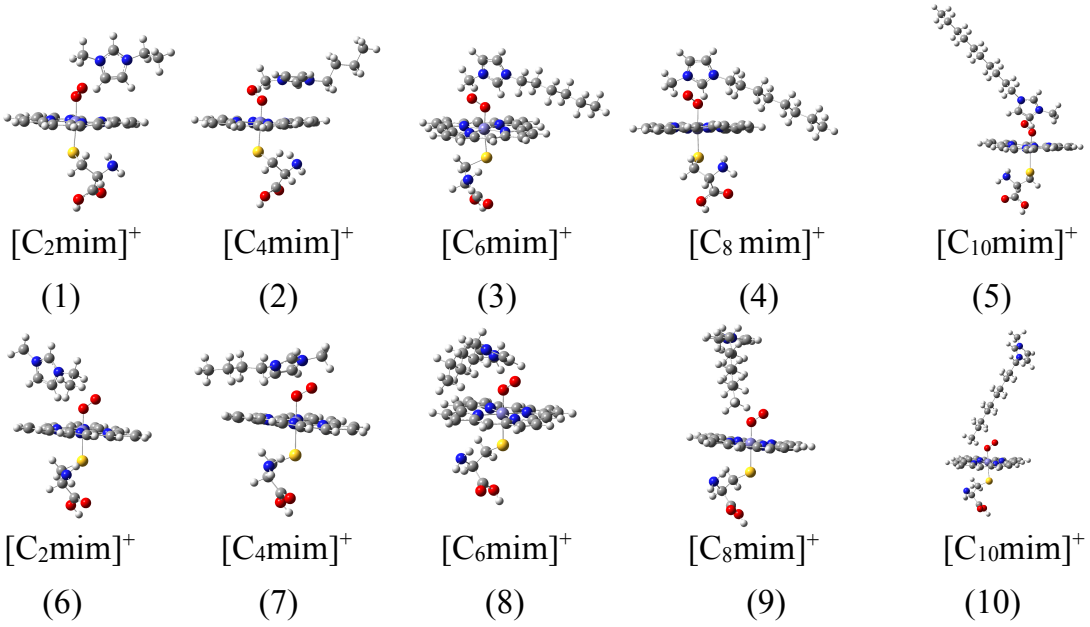


Figure A41: Optimized geometries of S_4 complexes in embedded phase; TU (1-5) and TD (6-10)

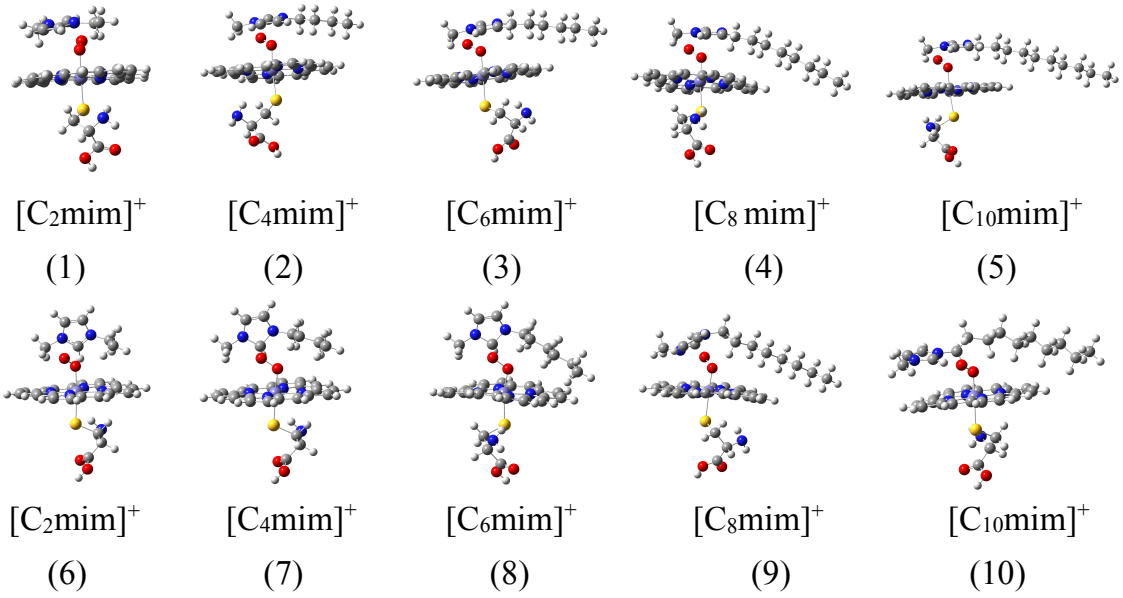


Figure A42: Optimized geometries of S_5 complexes in gas phase; TU (1-5) and TD (6-10)

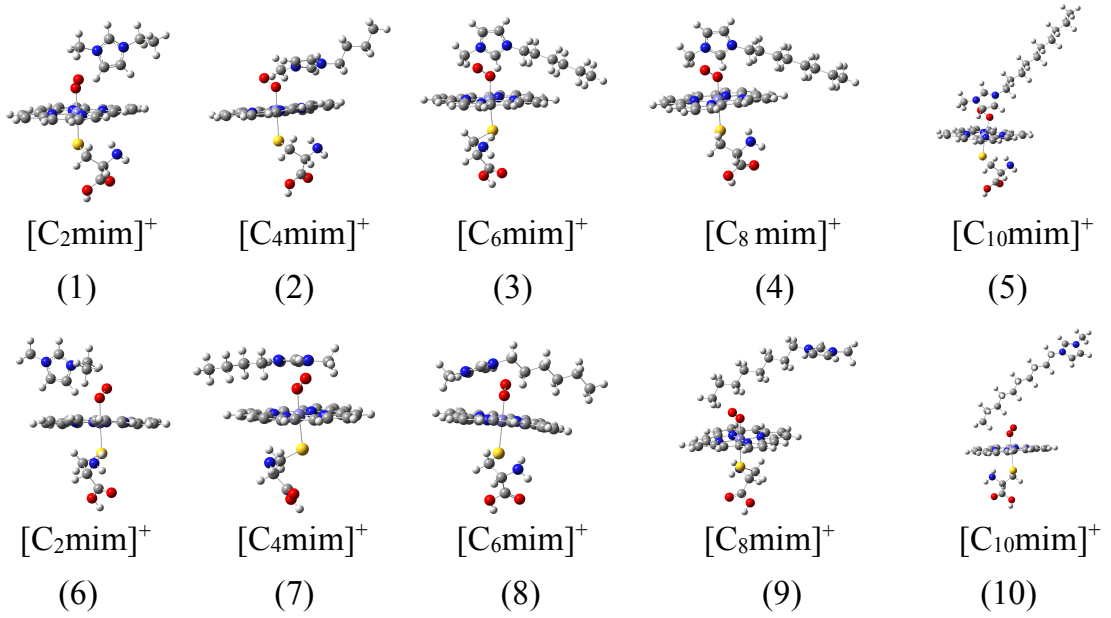


Figure A43: Optimized geometries of S_5 complexes in embedded phase; TU (1-5) and TD (6-10)

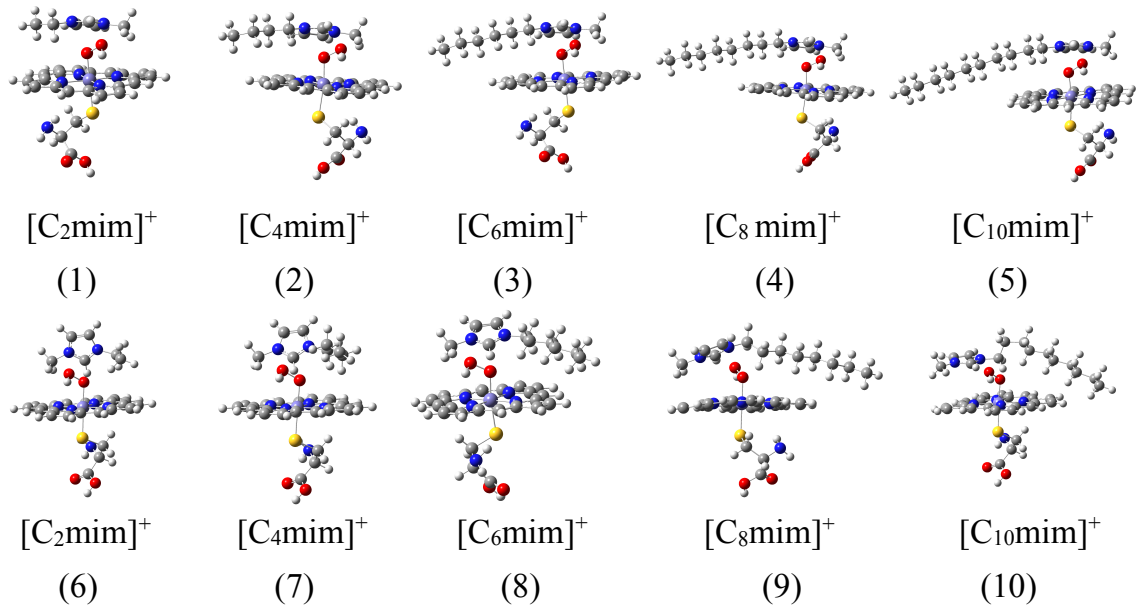


Figure A44: Optimized geometries of S_6 complexes in gas phase; TU (1-5) and TD (6-10)

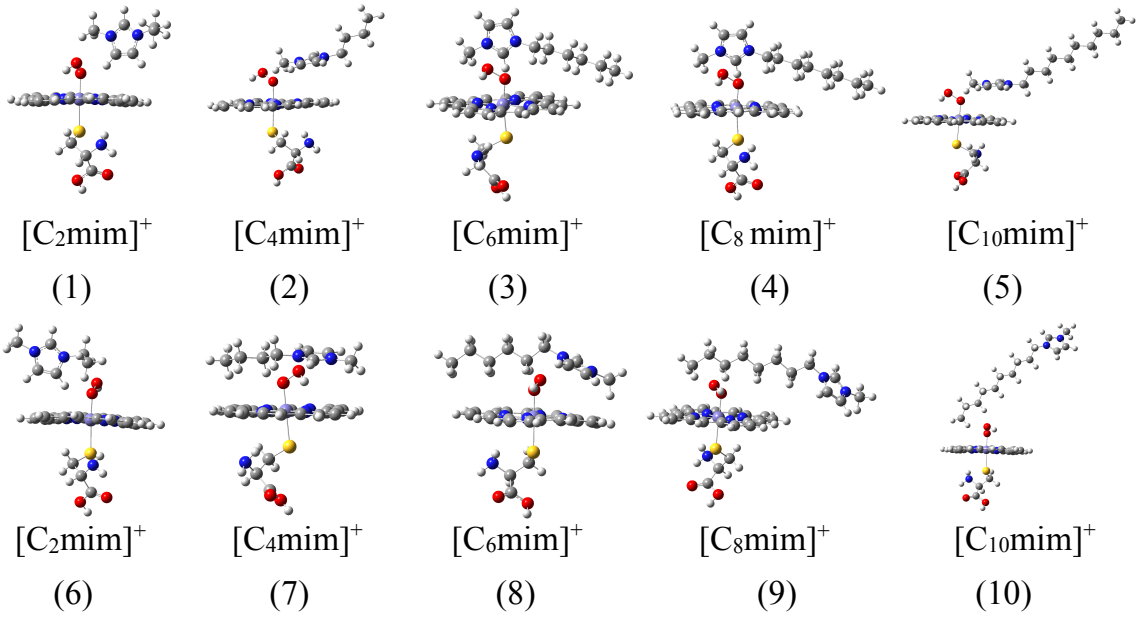


Figure A45: Optimized geometries of S_6 complexes in embedded phase; TU (1-5) and TD (6-10)

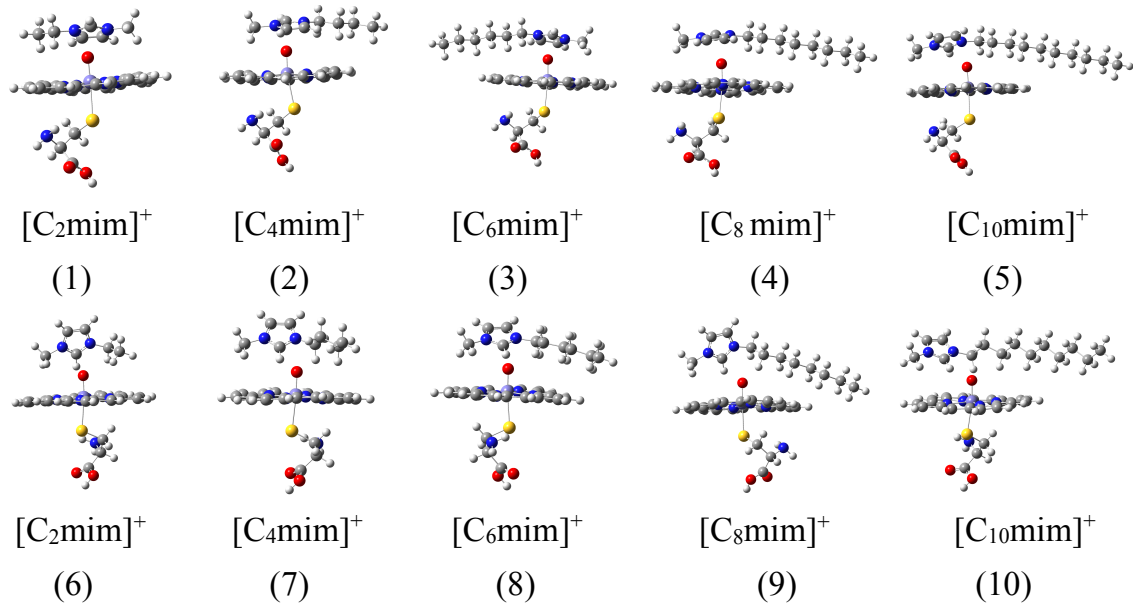


Figure A46: Optimized geometries of S_7 complexes in gas phase; TU (1-5) and TD (6-10)

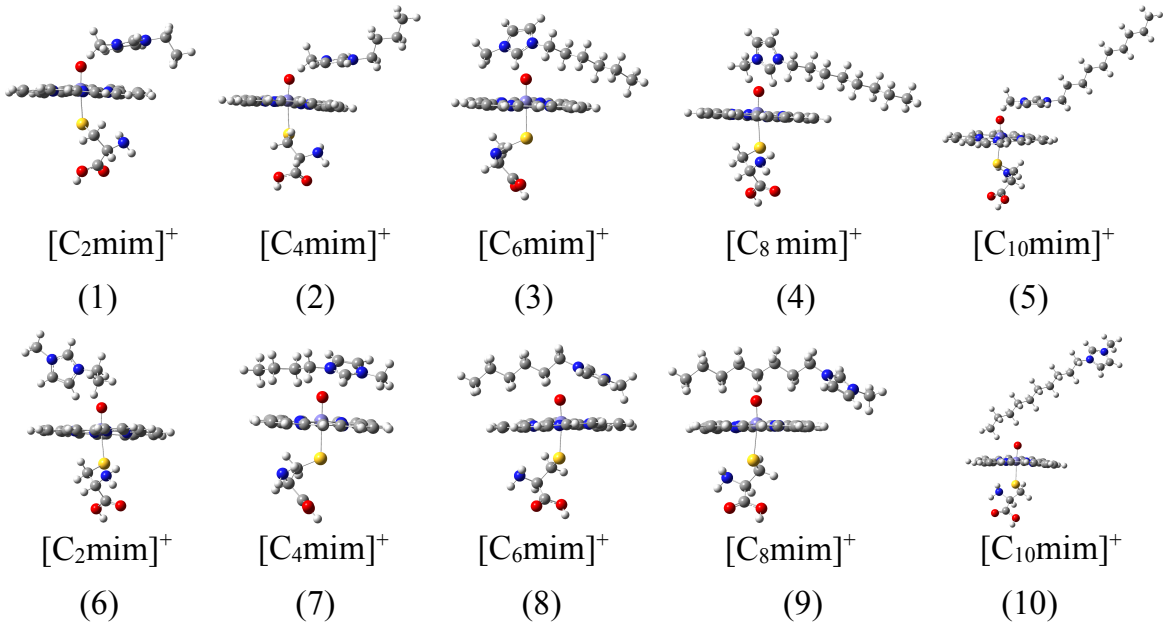


Figure A47: Optimized geometries of S_7 complexes in embedded phase; TU (1-5) and TD (6-10)

APPENDIX C

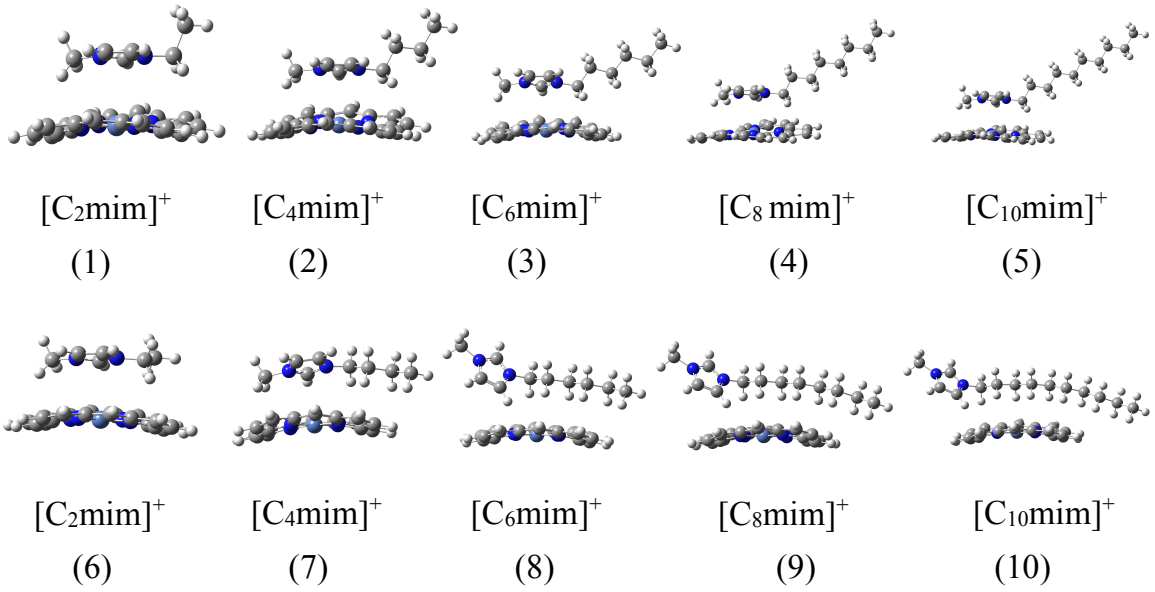


Figure A48: Optimized geometries of $[C_n mim]^+$NiP complexes; TU (1-5) and TD (6-10)

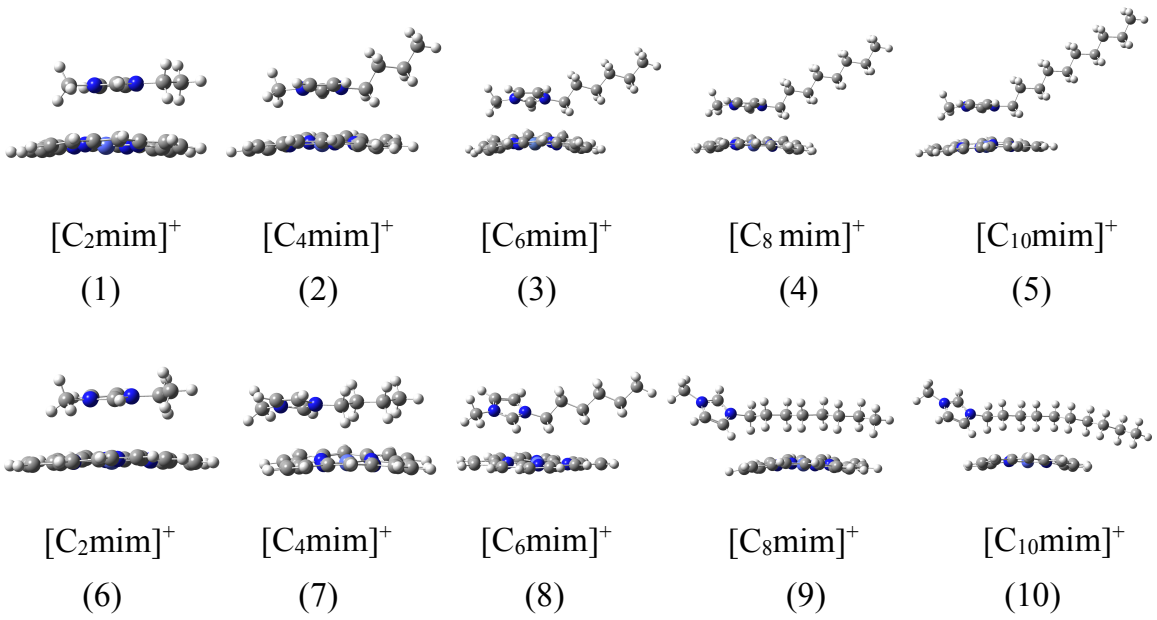


Figure A49: Optimized geometries of $[C_n mim]^+$CoP complexes; TU (1-5) and TD (6-10)

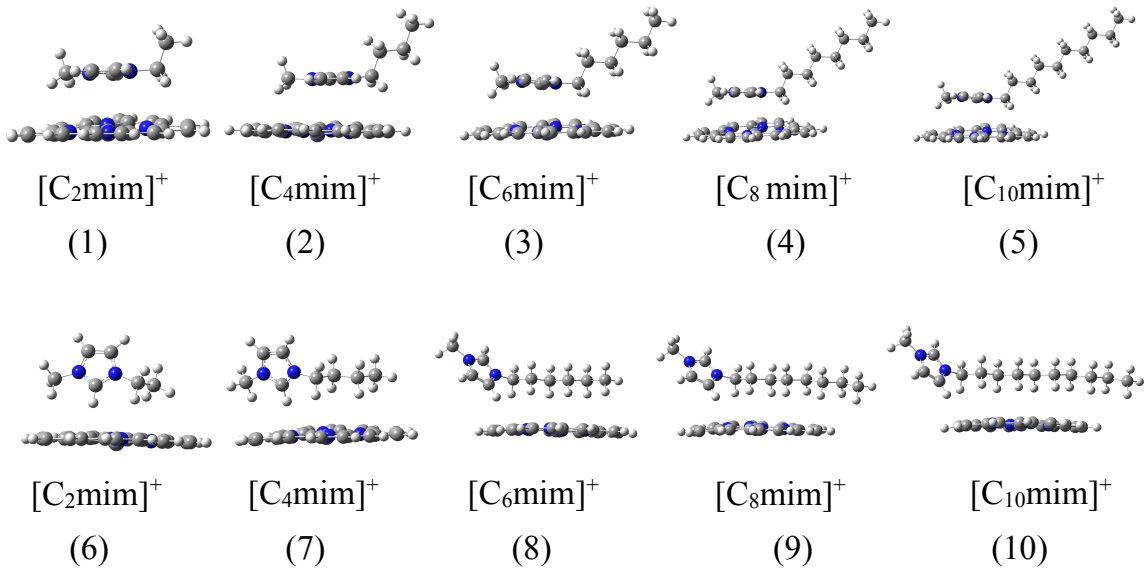


Figure A50: Optimized geometries of $[C_n mim]^+$ZnP complexes; TU (1-5) and TD (6-10)

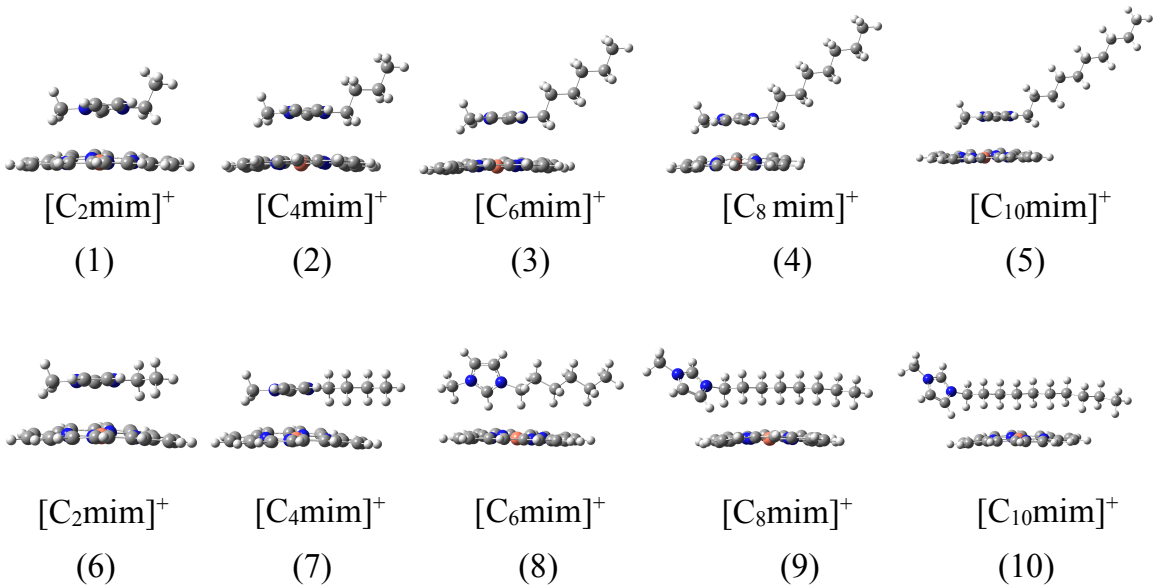


Figure A51: Optimized geometries of $[C_n mim]^+$CuP complexes; TU (1-5) and TD (6-10)

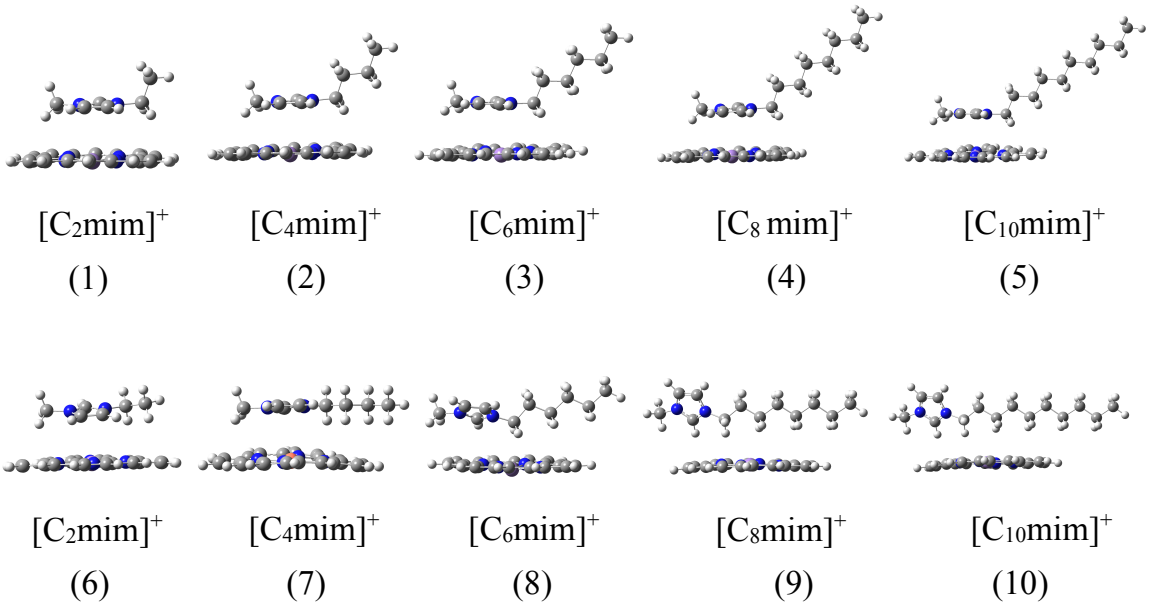


Figure A52: Optimized geometries of $[C_n mim]^+$MnP complexes; TU (1-5) and TD (6-10)

APPENDIX D

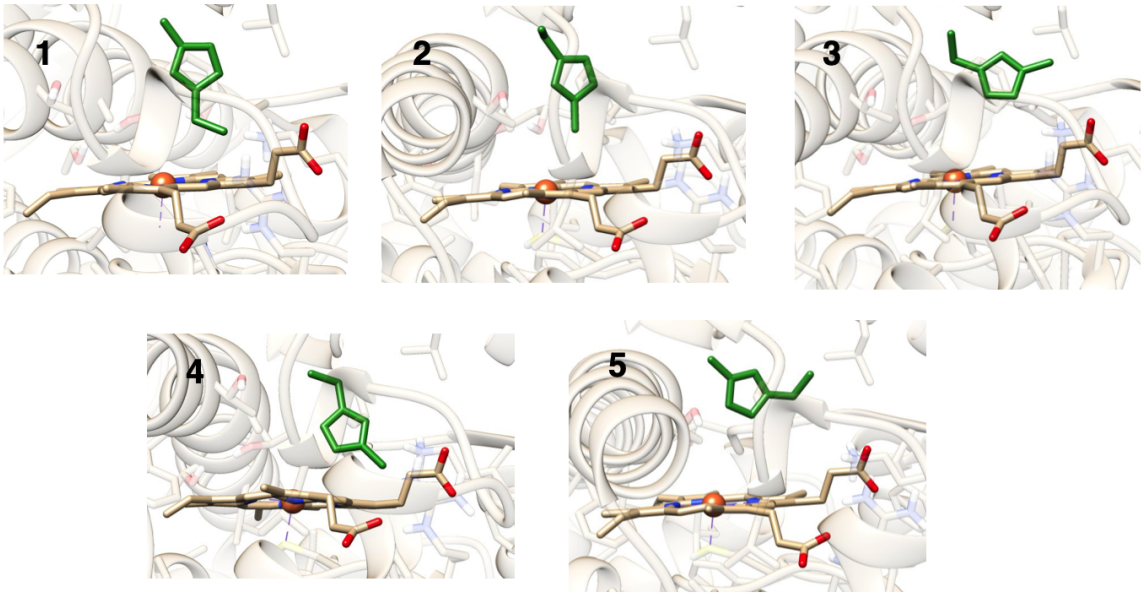


Figure A53: Most favorable docking modes for $[C_2mim]^+$ docked into substrate free cytochrome P-450 BM3 (PDB: 1bvy)

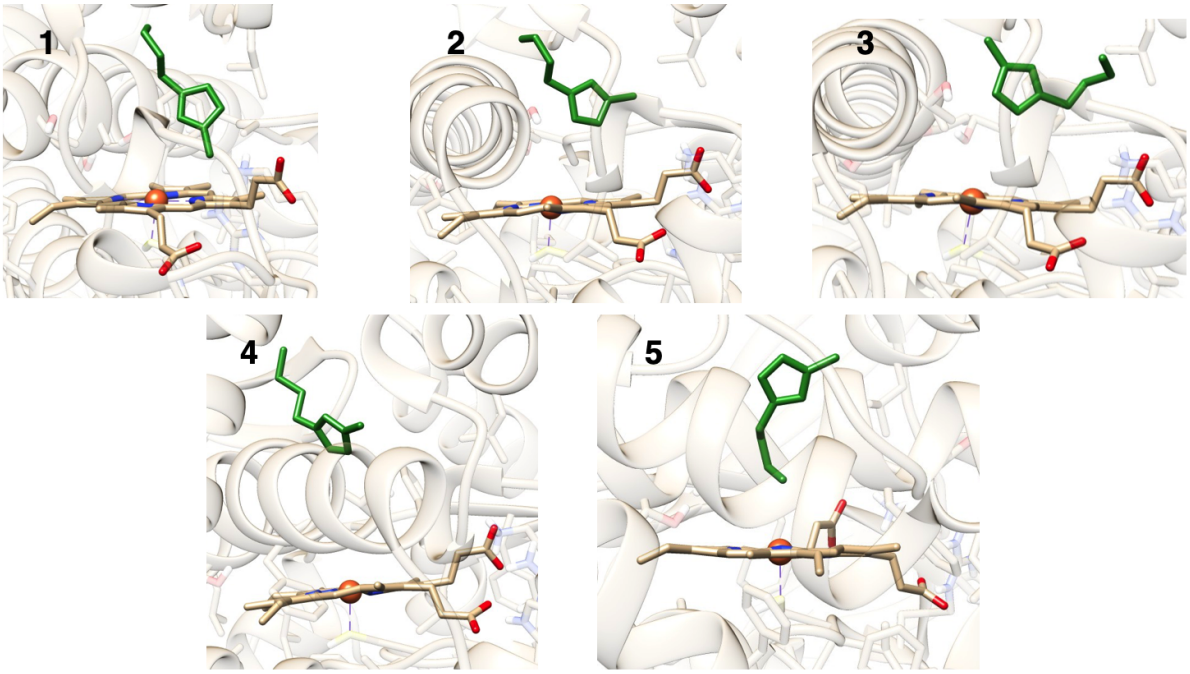


Figure A54: Most favorable docking modes for $[C_4mim]^+$ docked into substrate free cytochrome P-450 BM3 (PDB: 1bvy)

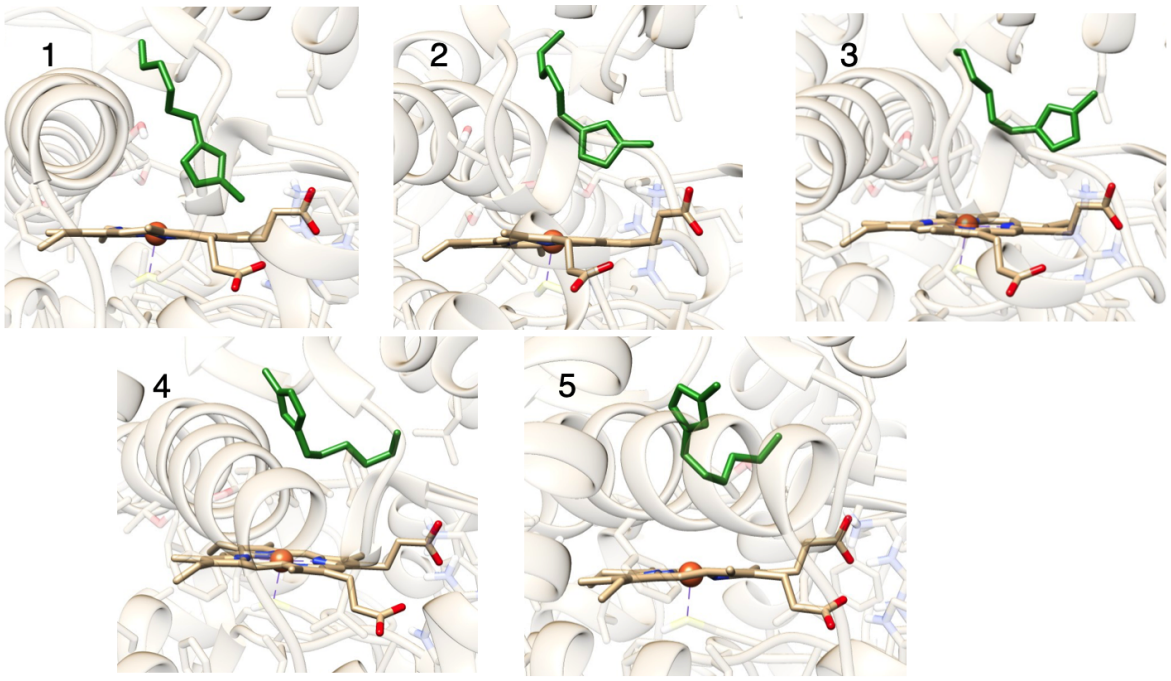


Figure A55: Most favorable docking modes for $[C_6mim]^+$ docked into substrate free cytochrome P-450 BM3 (PDB: 1bvy)

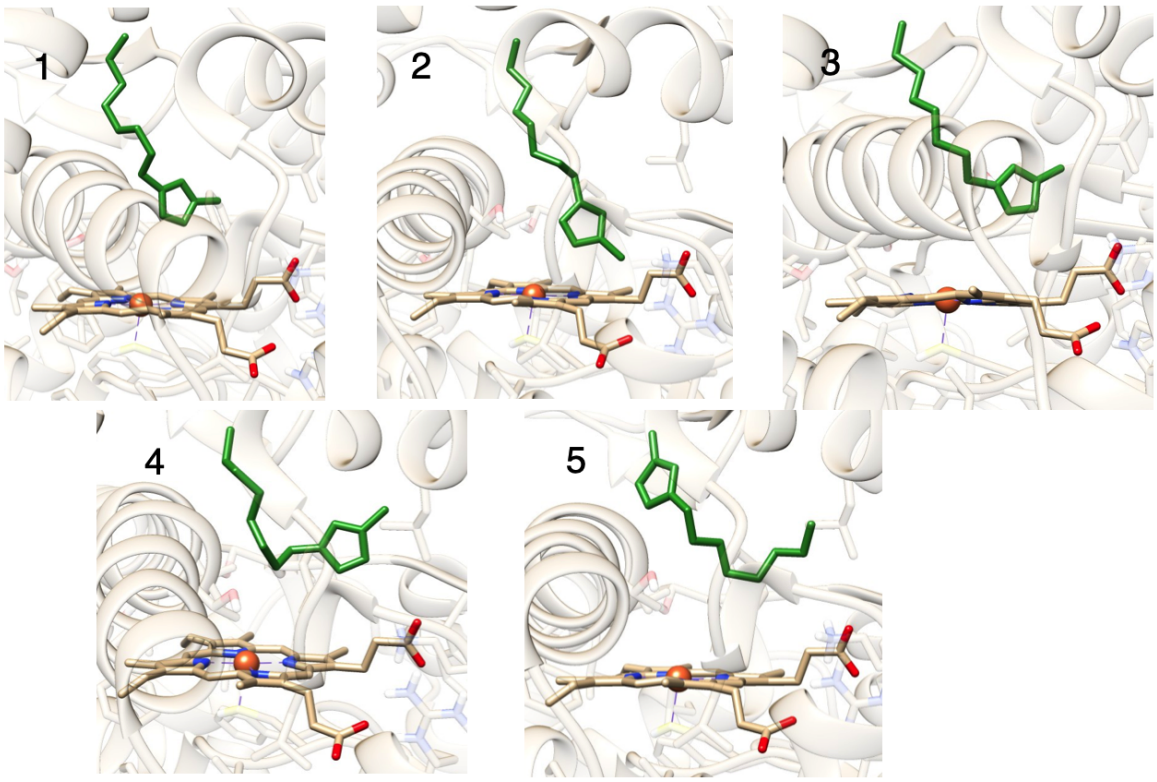


Figure A56: Most favorable docking modes for $[C_8mim]^+$ docked into substrate free cytochrome P-450 BM3 (PDB: 1bvy)

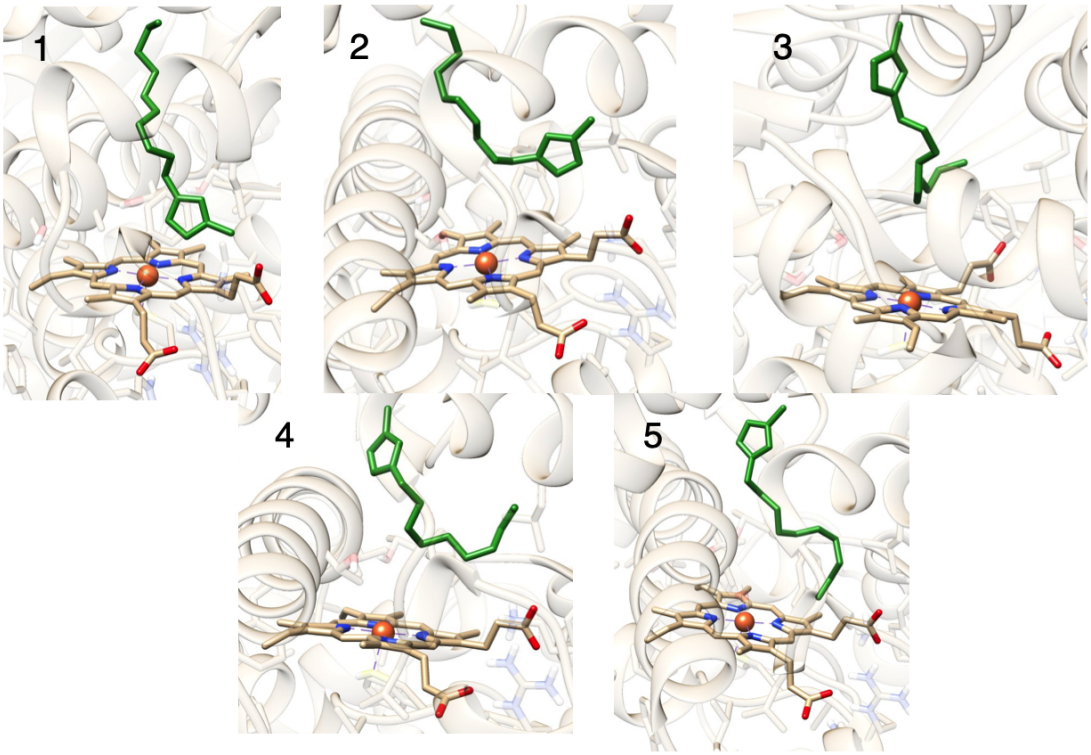


Figure A57: Most favorable docking modes for $[C_{10}mim]^+$ docked into substrate free cytochrome P-450 BM3 (PDB: 1bvy)

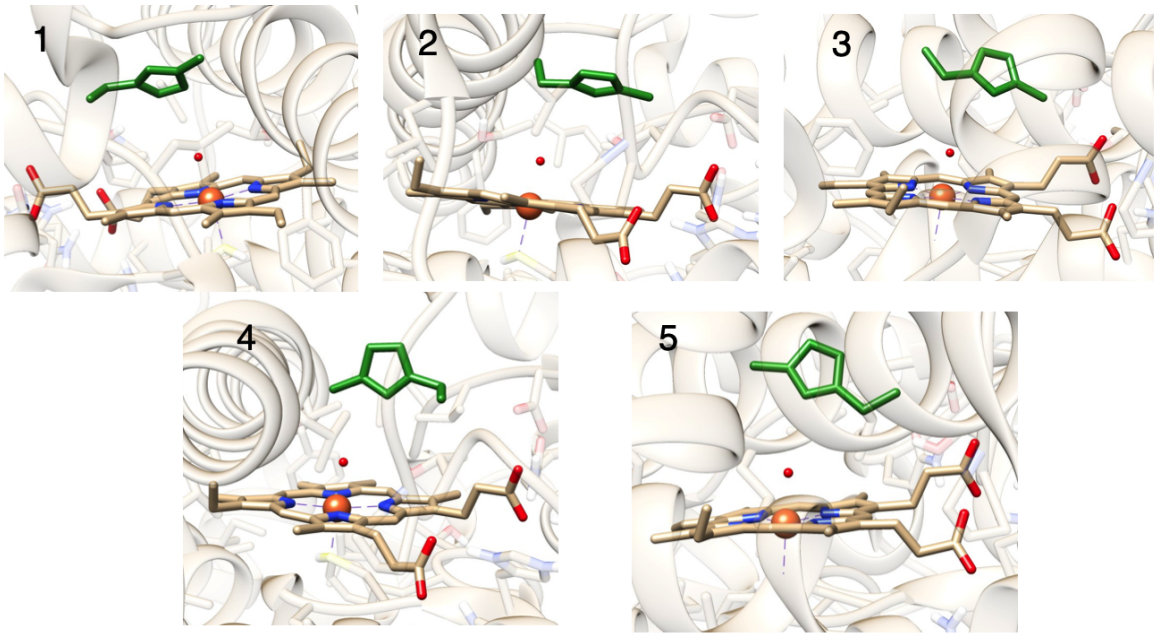


Figure A58: Most favorable docking modes for $[C_2mim]^+$ docked into substrate free cytochrome P-450 BM3 (PDB: 1dz9)

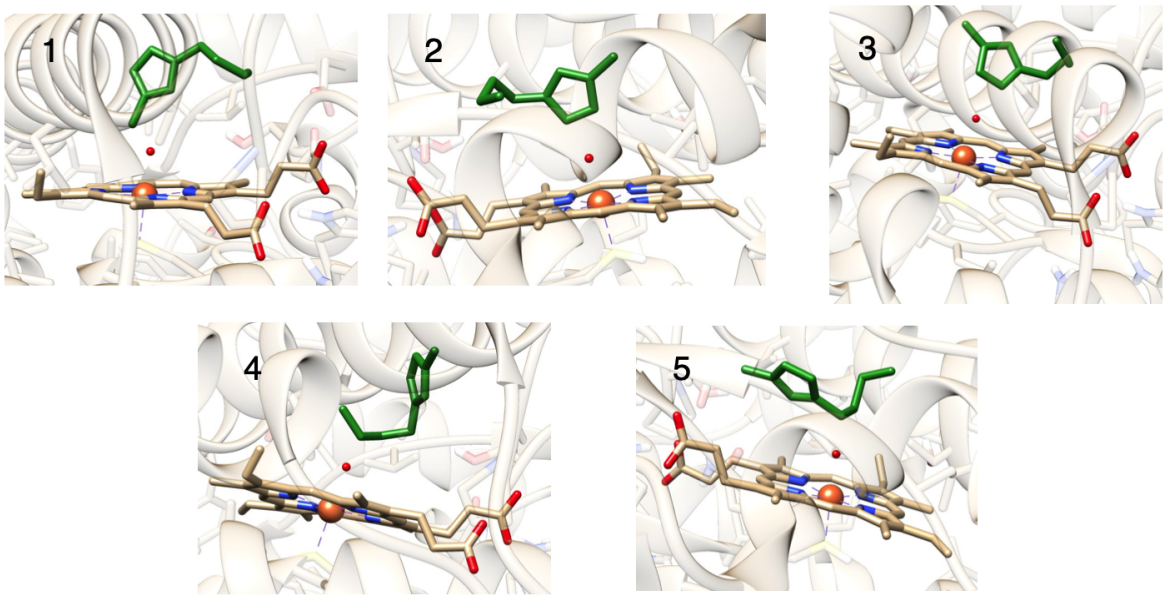


Figure A59: Most favorable docking modes for $[C_4mim]^+$ docked into substrate free cytochrome P-450 BM3 (PDB: 1dz9)

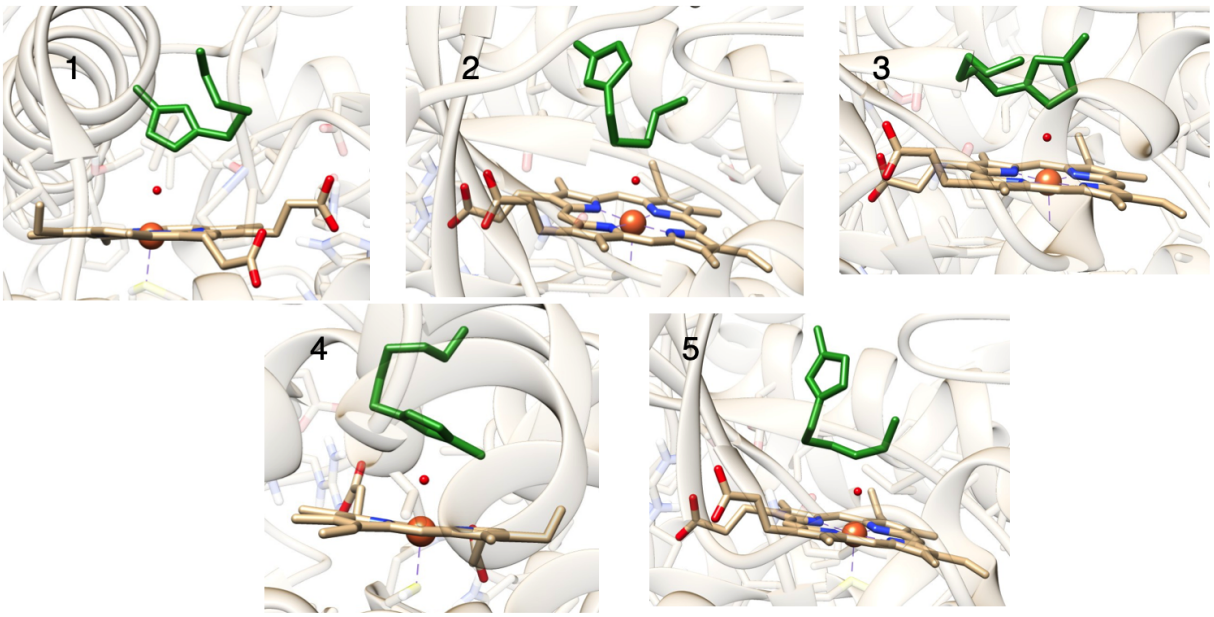


Figure A60: Most favorable docking modes for $[C_6mim]^+$ docked into substrate free cytochrome P-450 BM3 (PDB: 1dz9)

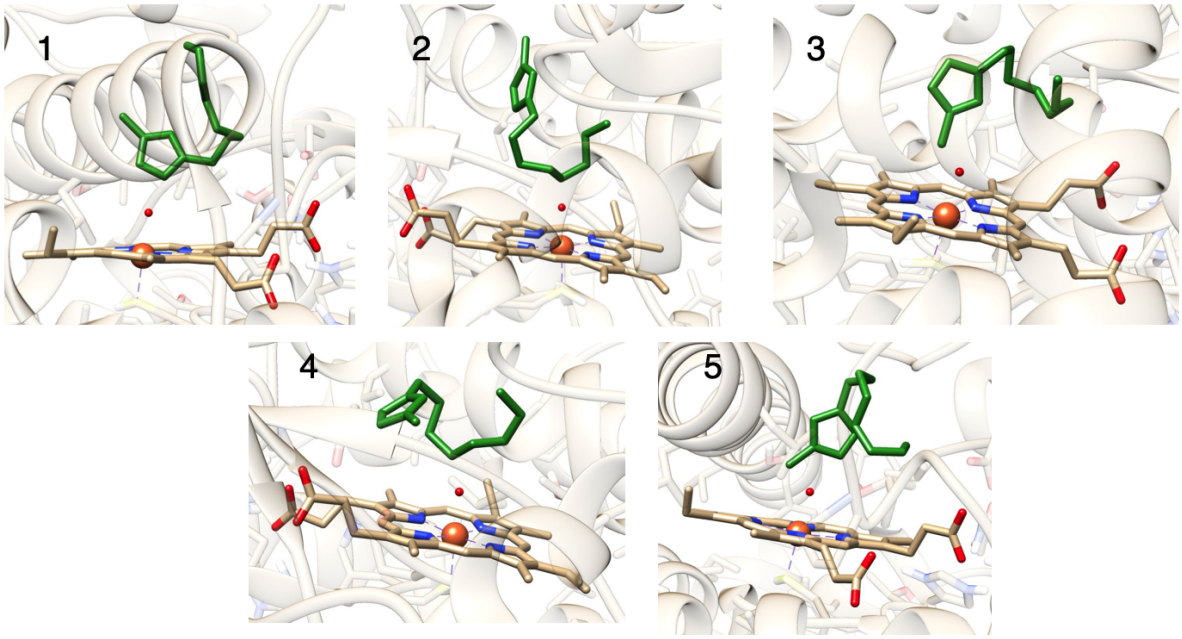


Figure A61: Most favorable docking modes for $[C_8mim]^+$ docked into substrate free cytochrome P-450 BM3 (PDB: 1dz9)

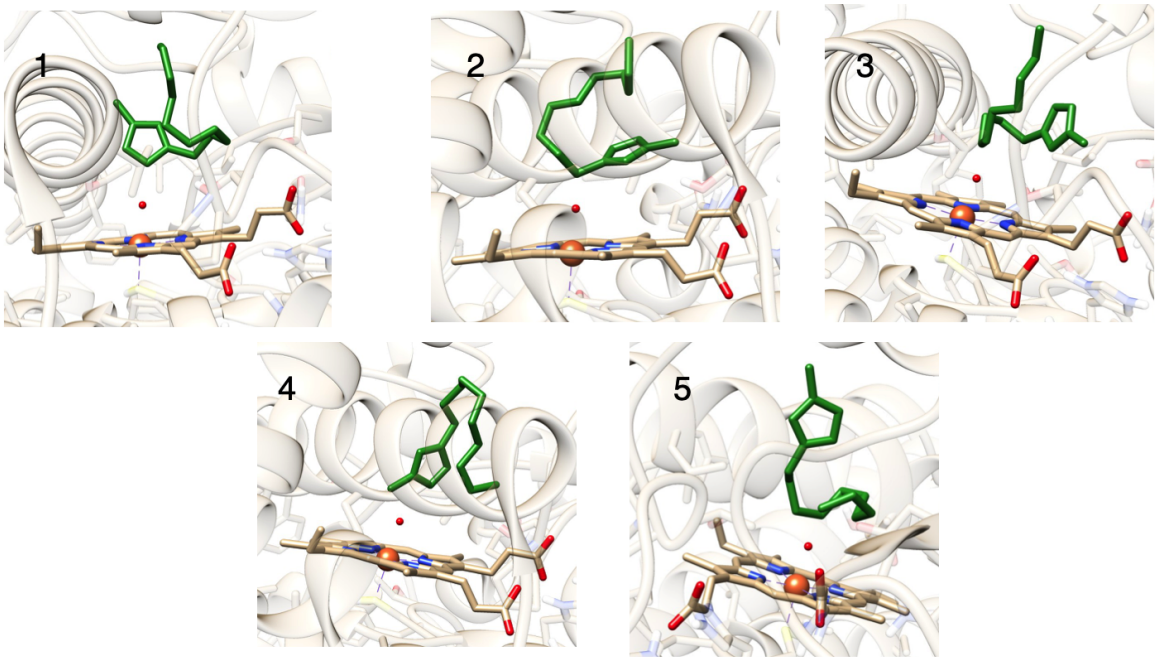


Figure A62: Most favorable docking modes for $[C_{10}mim]^+$ docked into substrate free cytochrome P-450 BM3 (PDB: 1dz9)

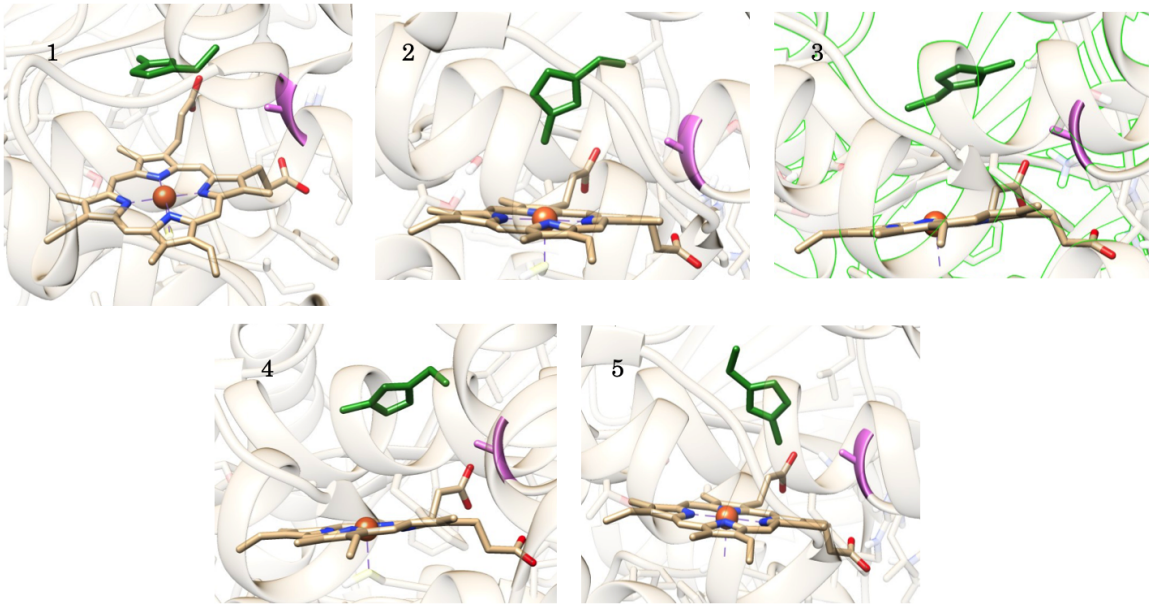


Figure A63: Most favorable docking modes for $[C_2mim]^+$ docked into mutated substrate free cytochrome P-450 BM3 (PDB: 1bvy F87A) ; violet : F87 A

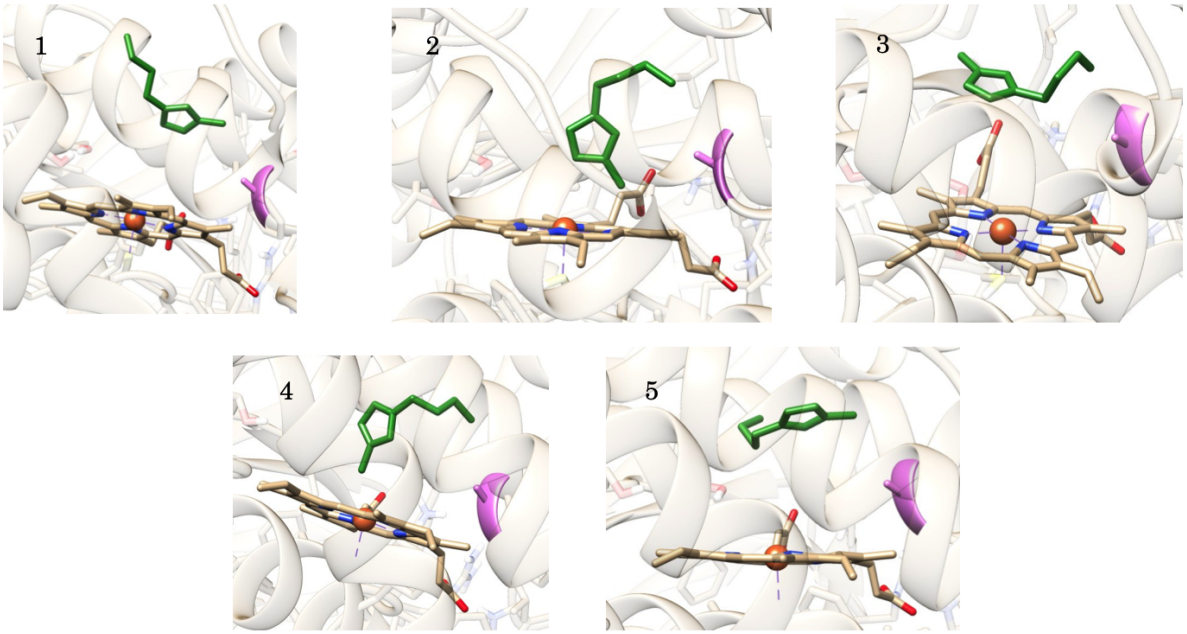


Figure A64: Most favorable docking modes for $[C_4mim]^+$ docked into mutated substrate free cytochrome P-450 BM3 (PDB: 1bvy F87A) ; violet : F87 A

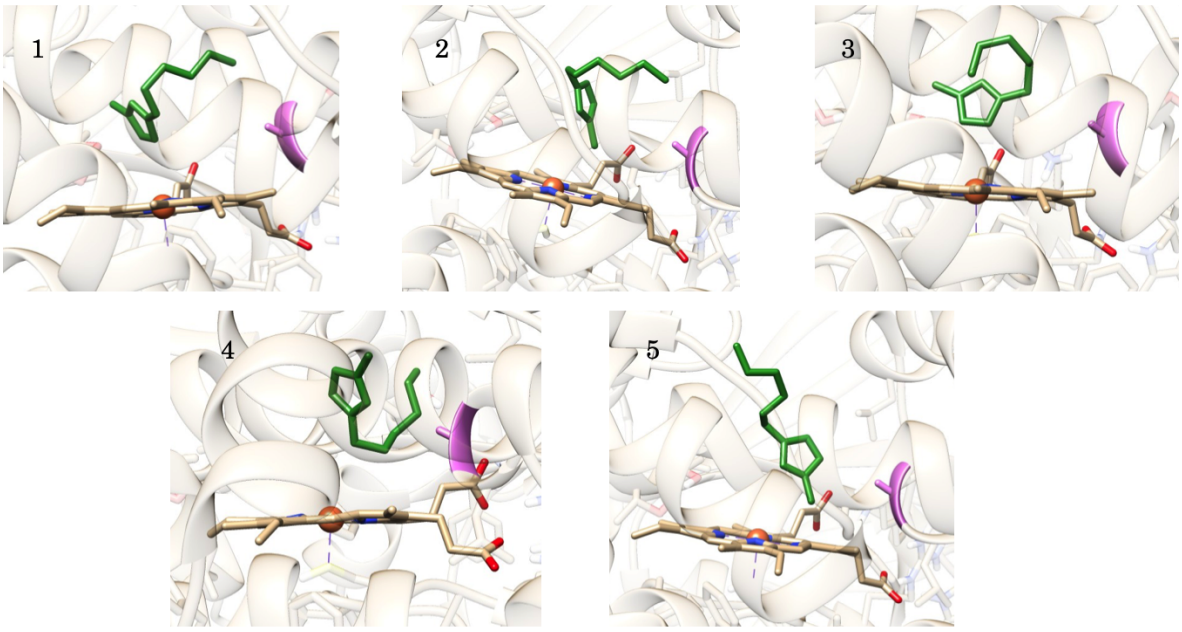


Figure A65: Most favorable docking modes for $[C_6mim]^+$ docked into mutated substrate free cytochrome P-450 BM3 (PDB: 1bvy F87A) ; violet : F87 A

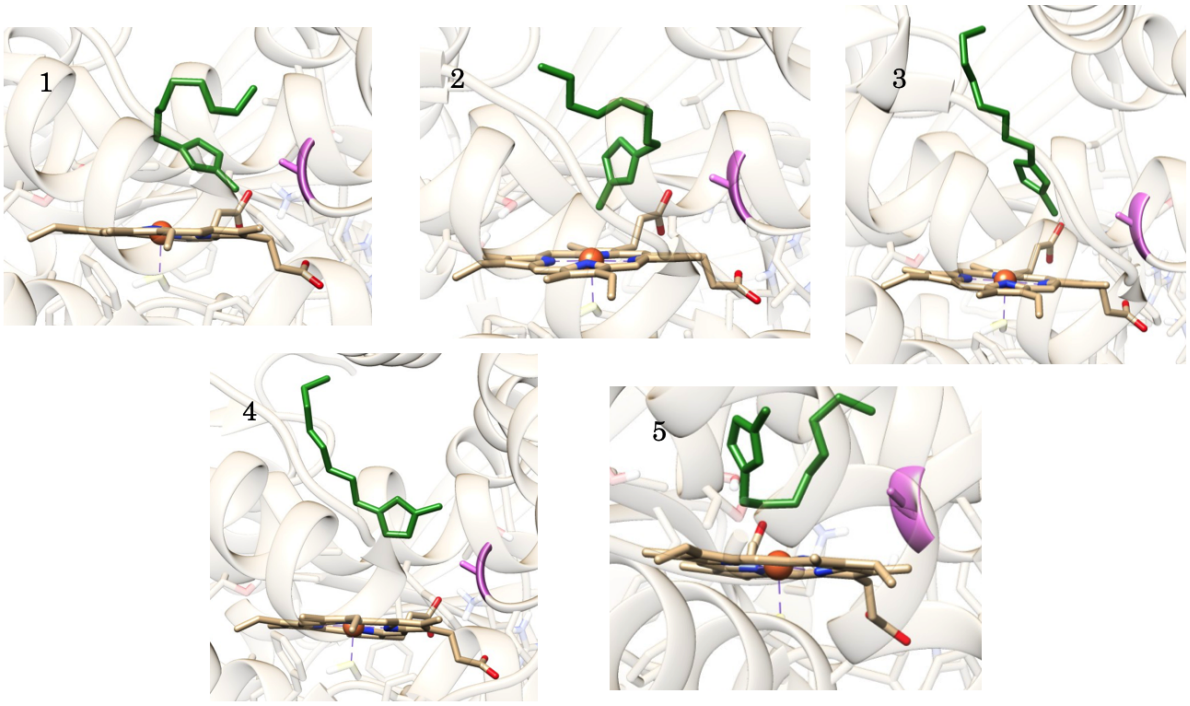


Figure A66: Most favorable docking modes for $[C_8mim]^+$ docked into mutated substrate free cytochrome P-450 BM3 (PDB: 1bvy F87A) ; violet : F87 A

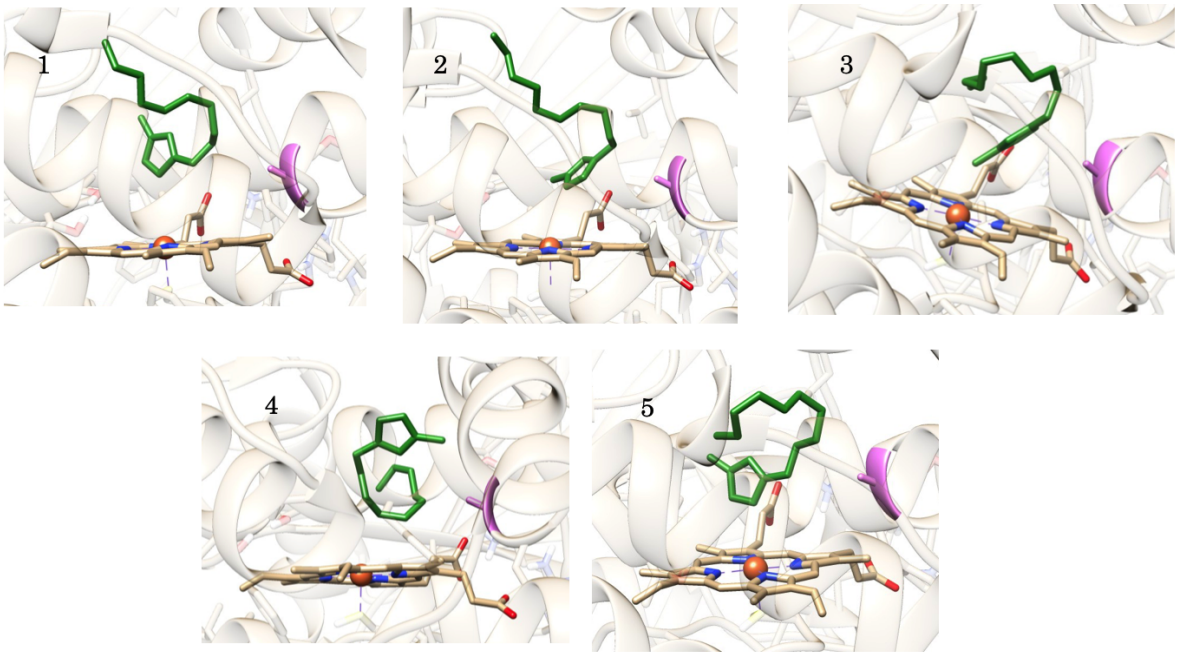


Figure A67: Most favorable docking modes for $[C_{10}mim]^+$ docked into mutated substrate free cytochrome P-450 BM3 (PDB: 1bvy F87A) ; violet : F87 A

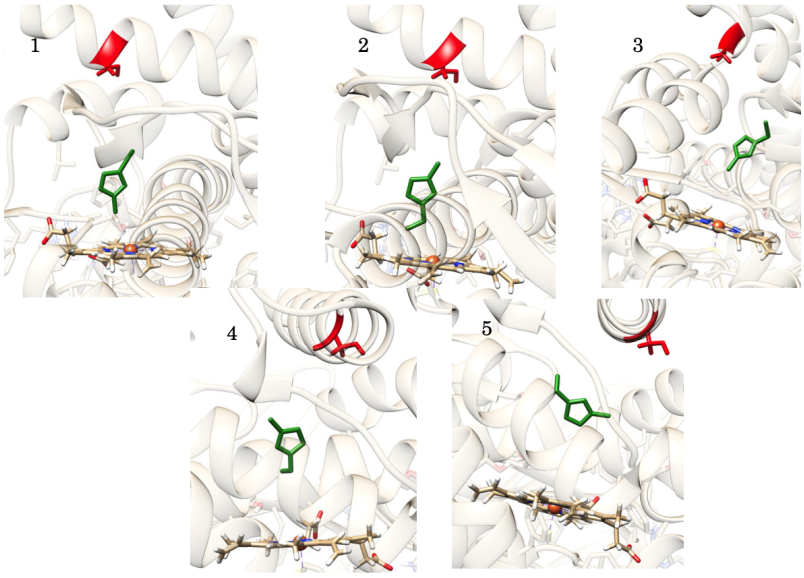


Figure A68: Most favorable docking modes for $[C_2mim]^+$ docked into mutated substrate free cytochrome P-450 BM3 (PDB: 1bvy A184S) ; red : A184S

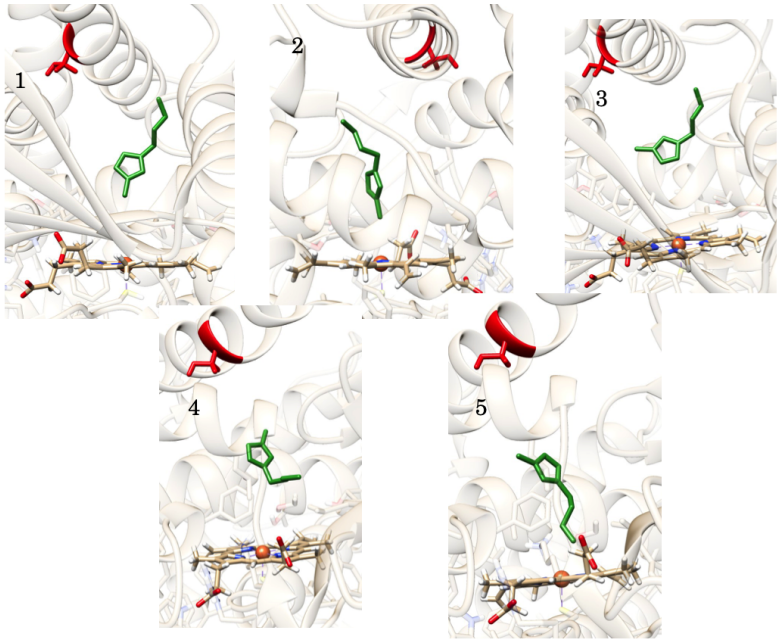


Figure A69: Most favorable docking modes for $[C_4mim]^+$ docked into mutated substrate free cytochrome P-450 BM3 (PDB: 1bvy A184S) ; red : A184S

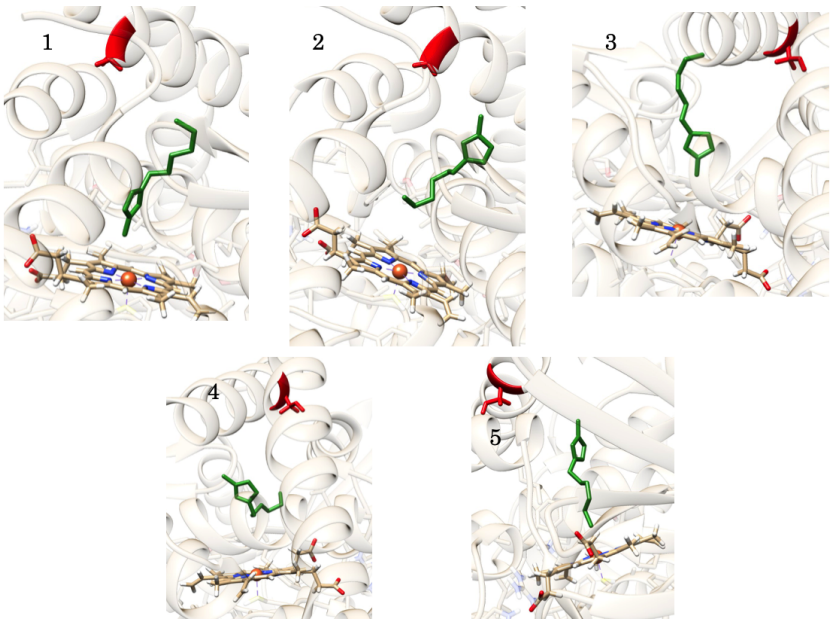


Figure A70: Most favorable docking modes for $[C_6mim]^+$ docked into mutated substrate free cytochrome P-450 BM3 (PDB: 1bvy A184S) ; red : A184S

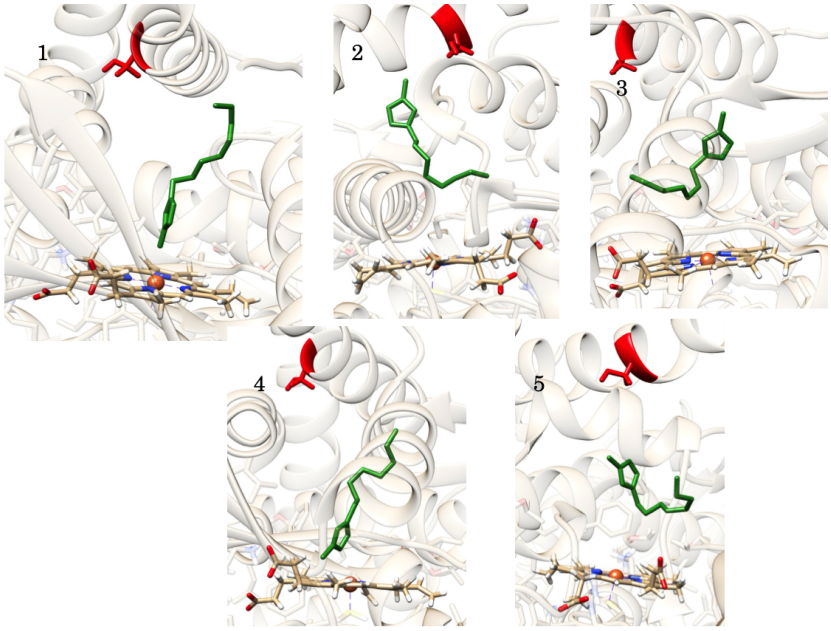


Figure A71: Most favorable docking modes for $[C_8mim]^+$ docked into mutated substrate free cytochrome P-450 BM3 (PDB: 1bvy A184S) ; red : A184S

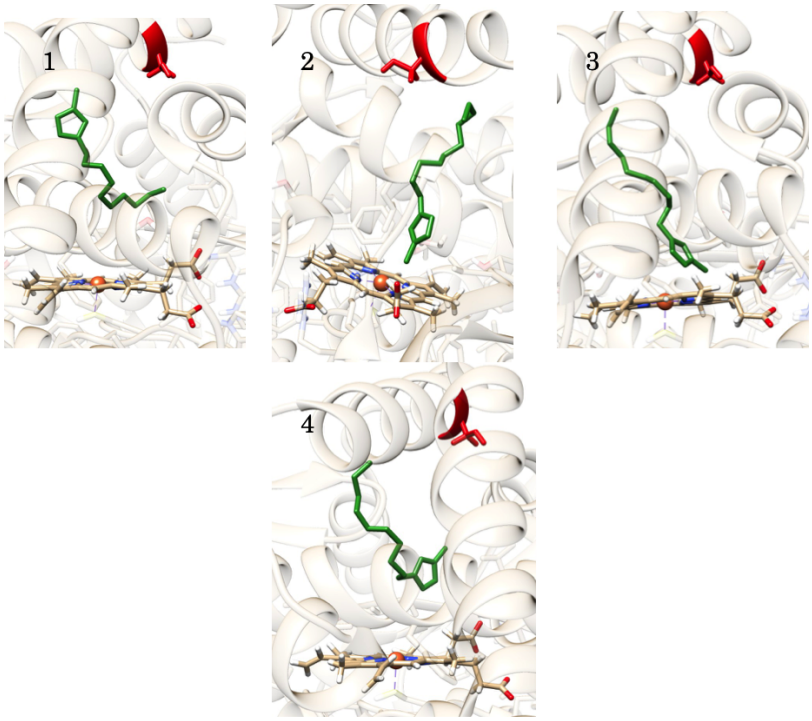


Figure A72: Most favorable docking modes for $[C_{10}mim]^+$ docked into mutated substrate free cytochrome P-450 BM3 (PDB: 1bvy A184S) ; red : A184S

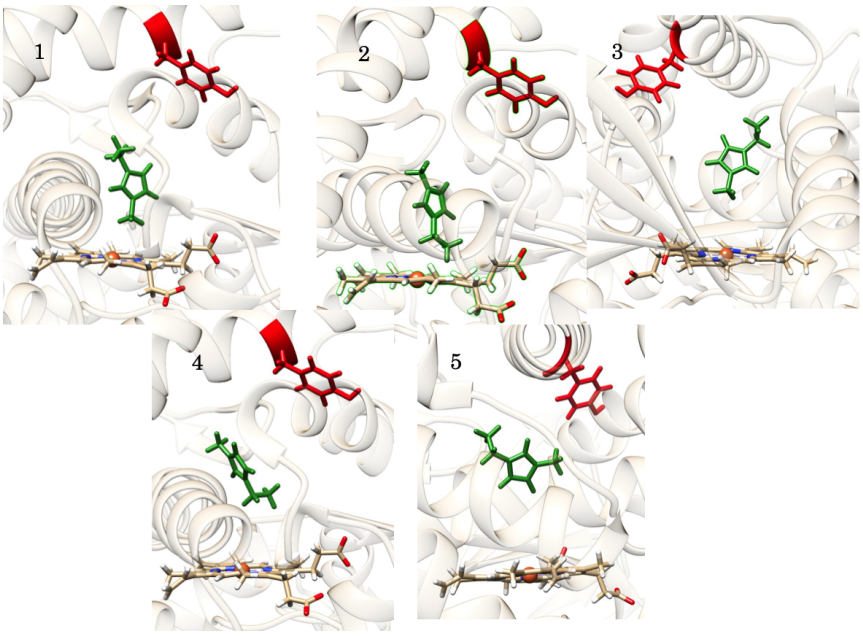


Figure A73: Most favorable docking modes for $[C_2mim]^+$ docked into mutated substrate free cytochrome P-450 BM3 (PDB: 1bvy A184Y) ; red : A184Y)

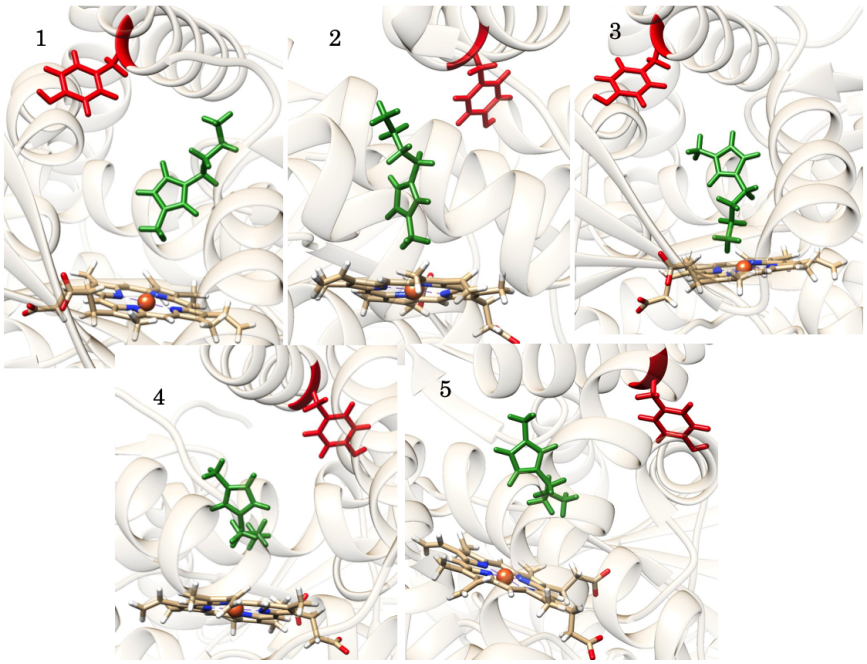


Figure A74: Most favorable docking modes for $[C_4mim]^+$ docked into mutated substrate free cytochrome P-450 BM3 (PDB: 1bvy A184Y) ; red : A184Y)

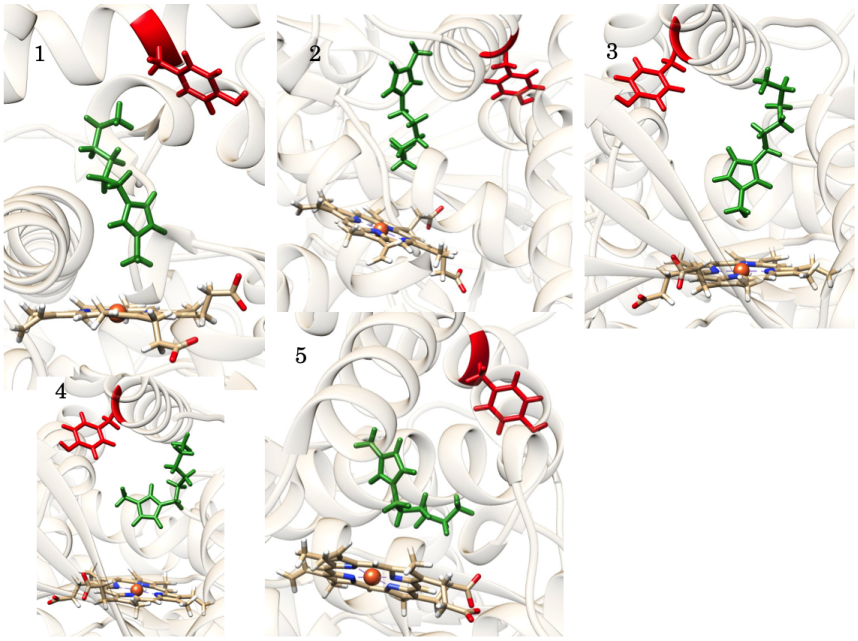


Figure A75: Most favorable docking modes for $[C_6mim]^+$ docked into mutated substrate free cytochrome P-450 BM3 (PDB: 1bvy A184Y) ; red : A184Y)

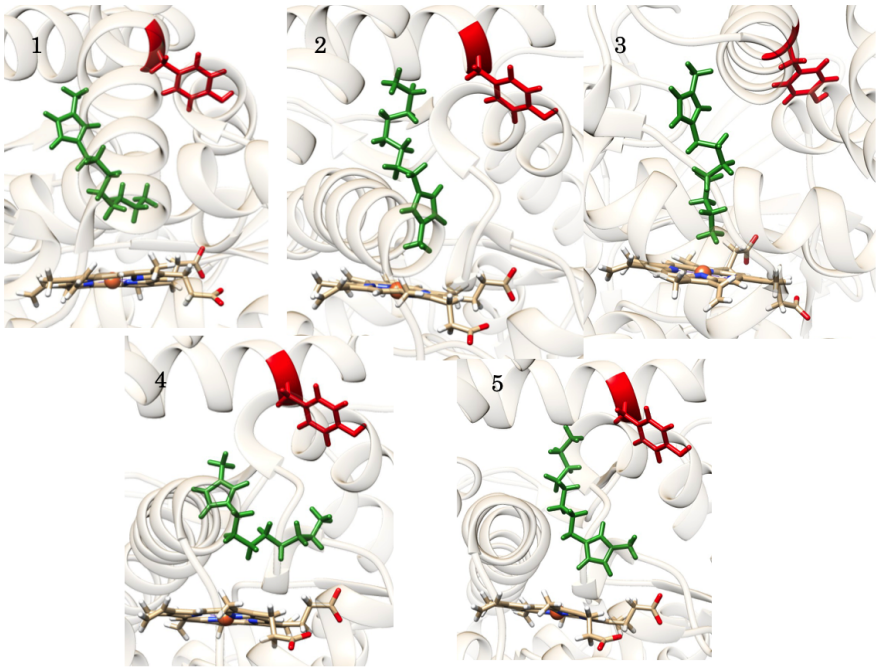


Figure A76: Most favorable docking modes for $[C_8mim]^+$ docked into mutated substrate free cytochrome P-450 BM3 (PDB: 1bvy A184Y) ; red : A184Y)

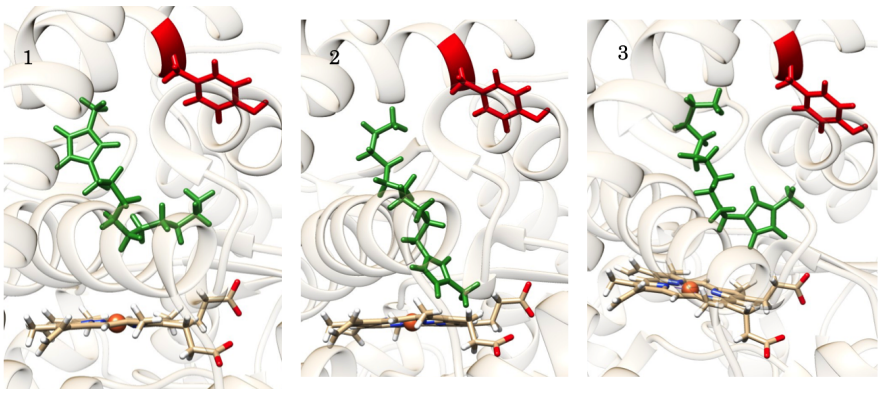


Figure A77: Most favorable docking modes for $[C_{10}mim]^+$ docked into mutated substrate free cytochrome P-450 BM3 (PDB: 1bvy A184Y) ; red : A184Y)

Receptor	Cation	Mode Docking Score(kcal/mol) 1	2	3	4	5
1BVY WT	$[C_2mim]^+$	-3.1	-3.0	-2.8	-2.8	-2.8
	$[C_4mim]^+$	-3.6	-3.5	-3.1	-2.9	-2.7
	$[C_6mim]^+$	-4.1	-3.9	-3.7	-3.7	-3.2
	$[C_8mim]^+$	-4.5	-4.1	-3.8	-3.5	-3.1
	$[C_{10}mim]^+$	-4.5	-4.4	-4.3	-3.4	-2.6
1BVY A328V	$[C_2mim]^+$	-2.6	-2.5	-2.5	-2.4	-2.4
	$[C_4mim]^+$	-3.2	-3.1	-2.7	-2.6	-2.6
	$[C_6mim]^+$	-3.6	-3.6	-3.3	-3.0	-2.9
	$[C_8mim]^+$	-3.6	-3.3	-2.6	-2.6	-2.1
	$[C_{10}mim]^+$	-3.7	-2.6	-2.5		
1BVY F87A	$[C_2mim]^+$	-3.4	-3.3	-3.2	-2.4	-2.3
	$[C_4mim]^+$	-4.0	-4.0	-3.9	-3.8	-3.7
	$[C_6mim]^+$	-4.2	-4.2	-3.9	-3.8	-3.8
	$[C_8mim]^+$	-4.2	-4.1	-3.9	-3.9	-3.9
	$[C_{10}mim]^+$	-4.7	-4.6	-4.5	-4.2	-4.2
1DZ9 WT	$[C_2mim]^+$	-3.3	-3.3	-3.3	-3.2	-3.1
	$[C_4mim]^+$	-3.6	-3.5	-3.4	-3.3	-3.3
	$[C_6mim]^+$	-3.9	-3.7	-3.6	-2.9	-2.9
	$[C_8mim]^+$	-4.1	-3.2	-3.0	-2.7	-2.6
	$[C_{10}mim]^+$	-3.7	-2.8	-2.2	-2.2	-1.9
1BVY A184S	$[C_2mim]^+$	-2.6	-2.5	-2.5	-2.4	-2.3
	$[C_4mim]^+$	-3.2	-3.1	-3.0	-2.7	-2.7
	$[C_6mim]^+$	-3.6	-3.6	-3.1	-3.1	-2.9
	$[C_8mim]^+$	-3.6	-3.6	-3.1	-2.6	-2.1
	$[C_{10}mim]^+$	-3.9	-2.8	-2.6	-2.2	-
1BVY A184Y	$[C_2mim]^+$	-2.6	-2.5	-2.5	-2.4	-2.3
	$[C_4mim]^+$	-3.2	-3.1	-2.7	-2.7	-2.6
	$[C_6mim]^+$	-3.6	-3.5	-3.3	-3.1	-2.9
	$[C_8mim]^+$	-3.7	-3.7	-3.7	-3.0	-2.5
	$[C_{10}mim]^+$	-3.8	-2.7	-2.4	-	-
1BVY (Aq)	$[C_2mim]^+$	-3.0	-2.9	-2.9	-2.8	-2.8
	$[C_4mim]^+$	-3.6	-3.6	-3.6	-3.6	-3.4
	$[C_6mim]^+$	-4.3	-4.2	-4.2	-4.0	-3.6
	$[C_8mim]^+$	-4.7	-4.6	-4.6	-3.9	-3.8
	$[C_{10}mim]^+$	-4.0	-3.8	-3.6	-3.5	-3.2

Table 8.1: Summary of Results obtained from Docking Calculations on Various Receptors

Normal Mode	Frequency (cm^{-1})	IR Intensity	Associated Motion
1	44.4	0	Saddling
2	48.3	0	Ruffling
3	102.0	4.2	Doming
4	151.5	0	Waving
5	151.5	0	Waving
6	214.7	0	Expansion Contraction

Table 8.2: First six non-planar normal modes for Iron Porphyrin

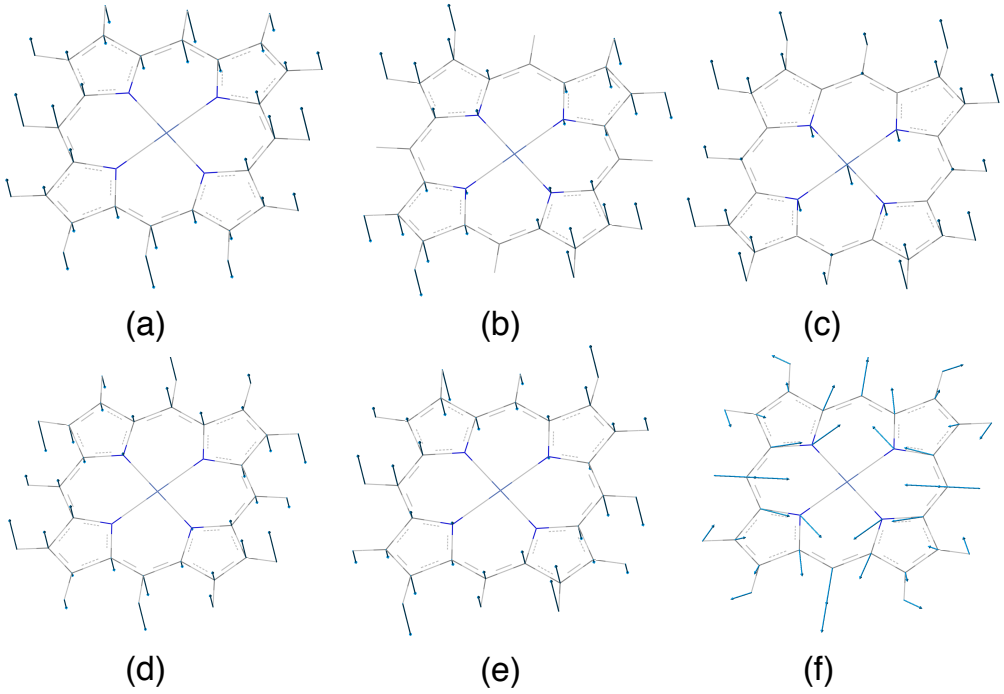


Figure A78: Motions associated with non-planar porphyrin modes with atomic displacement vectors

Normal Mode	Frequency (cm^{-1})	IR Intensity	Associated Motion
1	54.9	0	Saddling
2	62.9	0	Ruffling
3	85.2	2.0	Doming
4	150.9	0	Waving
5	150.9	0	Waving
6	203.5	0	Expansion Contraction

Table 8.3: First six non-planar normal modes for Nickel Porphyrin

Normal Mode	Frequency (cm^{-1})	IR Intensity	Associated Motion
1	39.3	0	Ruffling
2	44.2	0	Saddling
3	100.6	4.1	Doming
4	146.7	0	Waving
5	146.7	0	Waving
6	224.1	0	Expansion Contraction

Table 8.4: First six non-planar normal modes for Cobalt porphyrin

Normal Mode	Frequency (cm^{-1})	IR Intensity	Associated Motion
1	59.8	0	Saddling
2	64.1	0.5	Doming
3	76.9	0	Ruffling
4	146.7	0	Waving
5	146.7	0	Waving
6	224.1	33.2	Doming

Table 8.5: First six non-planar normal modes for Zinc Porphyrin

Normal Mode	Frequency (cm^{-1})	IR Intensity	Associated Motion
1	52.7	0	Saddling
2	54.9	0	Ruffling
3	90.8	3.7	Doming
4	150.3	0	Waving
5	150.3	0	Waving
6	199.2	0	Expansion Contraction

Table 8.6: First six non-planar normal modes for Copper Porphyrin

Normal Mode	Frequency (cm^{-1})	IR Intensity	Associated Motion
1	52.2	0.1	Doming
2	63.8	0	Saddling
3	82.2	0	Ruffling
4	132.3	0	Waving
5	132.2	0	Waving
6	148.7	31.8	Doming

Table 8.7: First six non-planar normal modes for Manganese Porphyrin

VITA

Atiya Banerjee

Candidate for the Degree of

Doctor of Philosophy

Dissertation: AN INVESTIGATION INTO BIODEGRADABILITY OF IONIC LIQUIDS EMPLOYING ACTIVE SITE MODELS

Major: Chemical Engineering

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Chemical Engineering at Oklahoma State University, Stillwater, Oklahoma in July, 2020.

Completed the requirements for the Master of Technology in Chemical Engineering at Indian Institute of Technology, Roorkee, Uttarakhand, India in 2015.

Completed the requirements for the Bachelor of Technology in Biochemical Engineering at Harcourt Butler Technological Institute, Kanpur, Uttar Pradesh, India in 2013.

Experience:

Development of a Computational Framework for understanding the Biodegradability of Ionic Liquids
Oklahoma State University, Aug 2015- Jul 2020.
Analysis of Graphene containing systems to understand Binding Mechanism of di-Cobaltates
Oklahoma State University, Jan 2020 – Mar 2020
Design of a Time Dependent Adaptive Neural Network for a-priori Prediction of UASB Reactor Outputs, I.I.T. Roorkee and M.H.R.D., India, Jun 2014- Jun 2015.
Computer Aided Process Plant Design for dimethyl ether (DME) production and optimization
I.I.T. Roorkee, January -June, 2014

Research Interests:
Protein Structure Prediction, Density Functional Theory, Molecular Modeling of Ionic Liquids, Optimization Algorithms, Empirical Modeling (ANN, Support Vector Machines), Energy Integration

Professional Memberships:

American Institute of Chemical Engineers, Student Member, 2015- Present
American Chemical Society, Student Member, 2017- Present