

---

ELECTROSTATIC AND QUANTUM  
CHEMICAL INVESTIGATION OF THE  
PROTON PUMPING MECHANISM OF  
CYTOCHROME *c* OXIDASE

---

**D I S S E R T A T I O N**

submitted to  
the Faculty of Biology, Chemistry and Geoscience  
of the University of Bayreuth, Germany

for obtaining the degree of  
Doctor of Natural Sciences

presented by  
Punnagai Munusami

Bayreuth 2008



Die vorliegende Arbeit wurde in dem Zeitraum von Januar 2005 bis Mai 2008 an der Universität Bayreuth unter der Leitung von Professor G. Matthias Ullmann erstellt.

Vollständiger Abdruck der von der Fakultät Biologie, Chemie und Geowissenschaften der Universität Bayreuth genehmigten Dissertation zur Erlangung des akademischen Grades Doktor der Naturwissenschaften (Dr. rer. nat.).

Erster Prüfer: Prof. Dr. G. Matthias Ullmann  
Zweiter Prüfer: Prof. Dr. Holger Dobbek  
Dritter Prüfer: Prof. Dr. Andreas Fery  
Prüfungsvorsitz: Prof. Dr. Jürgen Senker

Tag der Einreichung: 09.06.2008  
Kolloquium: 19.09.2008



---

ELECTROSTATIC AND QUANTUM  
CHEMICAL INVESTIGATION OF THE  
PROTON PUMPING MECHANISM OF  
CYTOCHROME *c* OXIDASE

---



---

## SUMMARY

---

Cytochrome *c* oxidase is a crucial enzyme in the respiratory chain. It catalyzes the reduction of oxygen to water and utilizes the free energy of the reduction reaction for proton pumping across the inner-mitochondrial membrane, a process which results in a membrane electrochemical proton gradient. For each oxygen molecule, eight protons are taken up from the matrix of the mitochondria. Four protons together with four electrons are required to reduce oxygen to water at the Fe<sub>a3</sub>-Cu<sub>B</sub> binuclear center and another four protons are translocated across the membrane. Although several high resolution structures have been solved for this enzyme, the molecular mechanism of the proton pumping and electron transfer is not understood.

Recent studies on the cytochrome *c* oxidase Cu<sub>B</sub> center suggested deprotonation of the Cu<sub>B</sub> bound imidazole ring of histidine (His291 in mammalian cytochrome *c* oxidase or His334 in *Rhodobacter sphaeroides* cytochrome *c* oxidase) as a key element in the proton pumping mechanism [1–6]. The central feature of this proposed mechanism is that the p*K*<sub>a</sub> value of the imidazole significantly lowered depending on the redox state of the metals in the binuclear center. The energetic feasibility of this mechanism is tested in this work.

To find a reliable method to calculate effective charges for the p*K*<sub>a</sub> calculations, the charge distribution of the tripeptide Ala-Asn-Ala with different conformations has been analyzed using the hybrid density functional method (B3LYP) with 6-31G\* basis set. Population analysis methods (such as Mulliken and Natural Population Analysis) and electrostatic potential (ESP) methods (such as CHELPG, MK and RESP) are used to analyze the charge distribution of the tripeptide Ala-Asn-Ala. This extensive study provided a better understanding of each method and the parameters which influence the partial atomic charges. The results show that ESP methods like CHELPG and MK give reliable charges when proper sampling points are used for the potential fit.

To comprehend the role of the Cu<sub>B</sub> bound histidines in the reaction mechanism of cytochrome *c* oxidase, density functional theory is used in combination with continuum electrostatics to calculate the p*K*<sub>a</sub> values of these imidazole rings in the aqueous solution as well as in the protein. The p*K*<sub>a</sub> values of His334, His333 and H<sub>2</sub>O molecule are calculated both in oxidized and reduced state of Cu<sub>B</sub> center. The Finite Difference Poisson Boltzmann (FDPB) method and the conductor-like polarizable continuum model (C-PCM) are used to determine the solvation free energies in aqueous solution.

All possible protonation equilibrium reactions in the Cu<sub>B</sub> center are studied to understand the deprotonation reactions of the bound H<sub>2</sub>O molecule, His333 and His334. In aqueous solution, p*K*<sub>a</sub> values of 15.2, 15.9 and 7.4 were obtained for deprotonation of His334, His333 and H<sub>2</sub>O

respectively. These  $pK_a$  values in aqueous solution show that His334 and His333 are likely to be protonated at physiological pH.

The protein environment shifts the  $pK_a$  values of the  $Cu_B$  ligands to even higher values in the range between 15 to 60. These  $pK_a$  values of  $Cu_B$  ligands are significantly higher compared to aqueous solution. The high  $pK_a$  values show that His334 is protonated during all steps of the catalytic cycle and demonstrate that the Fe and Cu ion oxidation states do not lower the  $pK_a$  values of  $Cu_B$  ligands and involved in shifting the  $pK_a$  values of  $Cu_B$  ligands to higher values. These results are incompatible with the proposed role of His334 as a key element in the pumping mechanism. According to the  $pK_a$  values, the proton pumping model as suggested by Stuchebrukhov [1] might not be possible with the involvement of His334. The  $pK_a$  values of the His333 in the  $Cu_B$  center are always shifted to higher values both in the reduced and in the oxidized state of the  $Cu_B$  center. The  $pK_a$  values of His333 show that this residue is likely to be protonated in the protein and an involvement in the reaction mechanism of cytochrome *c* oxidase can therefore be ruled out.

---

## ZUSAMMENFASSUNG

---

Cytochrom *c* Oxidase ist ein wichtiges Enzym in der Atmungskette. Es katalysiert die Reduktion von Sauerstoff zu Wasser und nutzt die freie Energie der Reduktion, um Protonen durch die innere mitochondriale Membran zu pumpen, ein Vorgang, der zu einem elektrochemischen Protonengradienten über der Membran führt. Für jedes Sauerstoffmolekül werden acht Protonen von der mitochondrialen Matrix aufgenommen. Vier Protonen zusammen mit vier Elektronen sind nötig, um Sauerstoff am Binuklearzentrum (Heme-Fe<sub>a3</sub>—Cu<sub>B</sub>) zu Wasser zu reduzieren und weitere vier Protonen werden durch die Membran transportiert. Obwohl für dieses Enzym einige hochaufgelösten Strukturen bestimmt wurden, ist der molekulare Mechanismus des Protonenpumpens und des Elektronentransfers nicht verstanden.

Neuere Studien am Cu<sub>B</sub>-Zentrum der Cytochrom *c* Oxidase legen nahe, dass die Deprotonierung eines am Cu<sub>B</sub> gebundenen Imidazolrings ein Schlüsselement im Protonenpumpenmechanismus darstellt (His291 in der Säugetier Cytochrom *c* Oxidase oder His334 in der *Rhodobacter sphaeroides* Cytochrom *c* Oxidase) [1–6]. Der zentrale Punkt dieses vorgeschlagenen Mechanismus ist eine erhebliche Verschiebung des p*K*<sub>a</sub>-Wertes eines Imidazols im Cu<sub>B</sub>-Zentrum zu niedrigeren Werten in Abhängigkeit vom Redoxzustand der Metalle im binuklearen Zentrum. Die energetische Möglichkeit dieses Mechanismus wird in dieser Arbeit geprüft.

Um eine verlässliche Methode für die Bestimmung effektiver Ladungen für die p*K*<sub>a</sub>-Berechnung zu finden, wurde die Ladungsverteilung im Tripeptids Ala-Asn-Ala in verschiedenen Konformationen mittels einer Hybriddichtefunktionsmethode (B3LYP) mit 6-31G\* als Basisatz analysiert. Populationsanalysemethoden (Mulliken-Analyse und Natural Population Analysis) und Methoden, die das elektrostatischen Potential (ESP) verwenden (CHELP, MK und RESP), wurden benutzt, um die Ladungsverteilung des Tripeptids Ala-Asn-Ala zu analysieren. Diese ausführlichen Untersuchung der Methoden zur Berechnung der Partialladungen liefert ein besseres Verständnis jeder Methode und der Parameter, die die partielle Atomladung beeinflussen. Die Ergebnisse zeigen, dass die ESP-Methoden, wie CHELP und MK, verlässliche Ladungen ergeben, wenn geeignete Probenpunkte für den Potentialfit benutzt werden.

Um die Rolle der an Cu<sub>B</sub> gebundenen Histidine im Reaktionsmechanismus der Cytochrom *c* Oxidase zu verstehen, wurden Dichtefunktionstheorie-Methoden in Verbindung mit Kontinuumselktrostatik-Rechnungen verwendet, um die p*K*<sub>a</sub>-Werte der Imidazolringe in wässriger Lösung und auch im Protein zu berechnen. Die p*K*<sub>a</sub>-Werte von His334, His333 und des gebundenen Wassermoleküls wurden im oxidierten und reduzierten Zustand des Cu<sub>B</sub> Zentrums berechnet. Solvatationsenergien in wässriger Lösung wurden mit Hilfe von finite-difference-

Poisson-Boltzmann-Rechnungen und mit Hilfe der C-PCM (conductor-like polarizable continuum models) Methode bestimmt.

Alle möglichen Protonierungsgleichgewichtsreaktionen im  $\text{Cu}_B$  Zentrum wurden untersucht, um die Deprotonierungsreaktionen des gebundenen Wassers, von His333 und von His334 zu verstehen. In wässriger Lösung erhielten wir  $\text{p}K_a$ -Werte von 15,2, 15,9 und 7,4 jeweils für die Deprotonierung von His334, von His333 und des gebundenen Wasser. Diese  $\text{p}K_a$ -Werte in wässriger Lösung zeigen, dass His334 und His333 bei physiologischen pH-Wert protoniert sind.

Die Proteinumgebung verschiebt die  $\text{p}K_a$ -Werte des  $\text{Cu}_B$ -Liganden zu noch höheren Werten im Bereich zwischen 15 und 60. Diese  $\text{p}K_a$ -Werte des  $\text{Cu}_B$ -Liganden sind deutlich höher als in wässriger Lösung. Die hohen  $\text{p}K_a$ -Werte von His334 zeigen, dass dieser Rest während aller Schritte des katalytischen Zyklus protoniert ist und somit der Oxidationszustand von Fe und Cu den  $\text{p}K_a$ -Wert der  $\text{Cu}_B$ -Liganden nicht erniedrigt. Diese Ergebnisse sind nicht vereinbar mit der vorgeschlagenen Rolle von His334 als Schlüsselement des Pumpmechanismus. Aufgrund der berechneten  $\text{p}K_a$ -Werten für His334 ist das Protonenpumpmodell, wie es Stuchebrukhov et al. [1] vorgeschlagen, nicht möglich. Auch die  $\text{p}K_a$ -Werte von His333 im  $\text{Cu}_B$ -Zentrum sind sowohl im reduzierten als auch im oxidierten Zustand des  $\text{Cu}_B$ -Zentrums immer zu höheren Werten hin verschoben. Die  $\text{p}K_a$ -Werte von His333 zeigen, dass dieser Rest im Protein protoniert ist. Eine Rolle von His333 im Reaktionsmechanismus der Cytochrom *c* Oxidase kann damit ebenfalls ausgeschlossen werden.

---

## ACKNOWLEDGMENTS

---

There are many people to thank for their support and encouragement, without whom this thesis would not have been possible.

First I would like to express my deep gratitude to my advisor, Prof. G. Matthias Ullmann, for his constant support and encouragement which were the driving forces throughout my Ph. D. I am grateful to him for giving me a chance to work in his lab and guided me with suggestions, discussions and ideas. Through all these years, he always had time to meet and discuss about science. His innovative ideas and supervision on all sorts of subjects were ever so inspiring. His suggestions to me were extended from the base like how to give a good talk till how to write good conclusions. I delighted my scientific freedom under his tutelage. His focused and critical direction contributed to the development of my own perspectives on science. Finally, I thank him for his understanding, patience and professionalism.

My sincere thanks to Dr. G. Narahari Sastry for introducing me to the fascinating world of computational chemistry. I wish to express my thanks and gratitude to Dr. Madhavi Sastry for being very supportive and a source of inspiration.

I am grateful to Dr. Timm Essigke for helping me in many aspects. I am indebted to him for everything he has done for me, devoting his valuable time for discussions and explaining the theoretical approaches with patience. I have been impressed by his diligence in work, knowledge about computers above all, his human qualities. I appreciate his efforts in explaining his program `QMPB` and `Perl Molecule` which I used in my projects.

I deeply express my appreciation to Eva-Maria Krammer for extending her support and caring from the beginning when I joined in the lab. She also has been a constant source of books and manuscripts. Many thanks to her for helping me with scientific and non-scientific things.

Frank Dickert helped me a lot, from vacating my apartment to filling forms in German. His timely helps are greatly acknowledged.

Mirco Till provided his program to calculate the protein cavities and hydrogen bond networks. I am thankful to him for patiently answering my questions regarding his program. Thomas Ullmann provided his scripts and programs. I also enjoyed the discussion about the different charge methods with him.

I would like to acknowledge, Edda Kloppmann for providing lot of information about Bayreuth before I arrive here. I also enjoyed watching movie with her. Special thanks to Dr. Astrid Klengen for giving a nice company during the practical courses and also for good comments during scientific presentations. Interacting with Dr. Elisa Bombarda was a pleasant feeling. I enjoyed her company during the lunch time. I really missed her in last few months. Fruitful

discussion with, Dr. Torsten Becker are greatly acknowledged. I enjoyed participating in parties with Silke Wieninger. I remember Siriporn Promsri birthday party, it was one of the best party I had ever enjoyed. Thanks to her for the great thai food. I really had fun in all the X-mas parties and BBQs which I am going to miss.

My stay here at Bayreuth have been great time in my life. For that I must thank my friends and colleagues who filled the days with happiness.

I acknowledge past and present members of the Structural Biology/Bioinformatics group (BISB) for providing me very friendly place to work.

I wish to thank Yangmin for her company and also thankful to her for providing delicious Korean food. She also taught me how to make sushi. I owe lot of thanks to Suja for being my wonderful friend and had some great moments together.

No words to express my thanks to my husband Dr. Semparithi Aravindan for helping me to achieve my goals and gave me a solid support during tough times.

I wish to extend my heartfelt thanks to my parents Mr. K. V. Munusami and Mrs. Visaya Munusami and my brother Mr. Puratchidasan Munusami and my in-laws Mr. G. Aravindan and Mrs. Mangayarkkarasi Aravindan for their care and affection which is leading me in successful way.

I would like to thank BIGSS, the state of Bavaria, Germany for its financial support over the period of my research. I enjoyed all the spring meetings and summer schools organized by the BIGSS and the memories will be cherished by me.

Nature has simplicity and therefore  
a great beauty.

Richard P. Feynman



---

# CONTENTS

---

|   |           |
|---|-----------|
| <b>Summary</b>  | <b>7</b>  |
| Abstract . . . . .  | 7         |
| <b>Zusammenfassung</b>  | <b>9</b>  |
| Zusammenfassung . . . . .   | 9         |
| Acknowledgements . . . . .  | 11        |
| <b>1 Introduction</b>   | <b>19</b> |
| 1.1 Introduction . . . . .  | 19        |
| 1.1.1 Cytochrome <i>c</i> oxidase - A redox-driven molecular machine . . . . .  | 19        |
| 1.1.2 Structure and function of cytochrome <i>c</i> oxidase . . . . .           | 19        |
| 1.2 The general mechanism: Electron transfer and proton translocation . . . . . | 27        |
| 1.3 Outline of the thesis . . . . .   | 33        |
| <b>2 Computational methods</b>  | <b>35</b> |
| 2.1 Acid-base and redox reactions equilibria . . . . .                          | 35        |
| 2.1.1 Fundamental description of acid-base and redox equilibria . . . . .       | 35        |
| 2.1.2 Computation of acid-base and redox equilibria . . . . .                   | 36        |
| 2.2 Continuum electrostatics . . . . .  | 38        |
| 2.2.1 The Poisson-Boltzmann equation . . . . .                                  | 39        |
| 2.2.2 Calculation of protonation and redox equilibria in proteins . . . . .     | 42        |
| 2.3 Quantum Chemistry . . . . .   | 51        |
| 2.3.1 Density Functional Theory . . . . .                                       | 52        |
| 2.3.2 Basis sets . . . . .  | 58        |
| 2.3.3 Solvation free energies . . . . .   | 59        |
| <b>3 Charge distribution in peptide conformations</b>                           | <b>63</b> |
| 3.1 Theoretical background . . . . .  | 64        |

---

|          |   |            |
|----------|---|------------|
| 3.1.1    | Wave function based methods . . . . .   | 64         |
| 3.1.2    | Potential based methods . . . . .   | 67         |
| 3.2      | Structure preparation and Computational Methods . . . . .   | 72         |
| 3.3      | Results and Discussion . . . . .  | 75         |
| 3.3.1    | Mulliken charges . . . . .  | 76         |
| 3.3.2    | NPA charges . . . . .   | 76         |
| 3.3.3    | CHELPG charges . . . . .  | 76         |
| 3.3.4    | MK charges . . . . .  | 76         |
| 3.3.5    | RESP charges . . . . .  | 78         |
| 3.4      | Conclusions . . . . .   | 80         |
| <b>4</b> | <b>Protonation and redox potentials of cytochrome <i>c</i> nitrite reductase</b>                                | <b>87</b>  |
| 4.1      | Hemes and the calcium binding site in cytochrome <i>c</i> nitrite reductase . . . . .                           | 89         |
| 4.2      | Preparation of the crystal structure of cytochrome <i>c</i> nitrite reductase . . . . .                         | 90         |
| 4.2.1    | Density Functional calculations . . . . .   | 90         |
| 4.2.2    | Continuum electrostatic calculations . . . . .  | 91         |
| 4.3      | Results and Discussion . . . . .  | 93         |
| 4.4      | Conclusions . . . . .   | 99         |
| <b>5</b> | <b><math>pK_a</math> calculations of binuclear center of cytochrome <i>c</i> oxidase using DFT calculations</b> | <b>103</b> |
| 5.1      | $pK_a$ calculations . . . . .   | 104        |
| 5.2      | Computational Methods . . . . .   | 105        |
| 5.2.1    | Density Functional calculations . . . . .   | 105        |
| 5.2.2    | Charge fitting . . . . .  | 108        |
| 5.2.3    | Solvation free energy calculations . . . . .  | 109        |
| 5.3      | Results and Discussion . . . . .  | 111        |
| 5.3.1    | $pK_a$ values of the $Cu_B$ center . . . . .  | 112        |
| 5.3.2    | $pK_a$ values of the heme $a_3$ center . . . . .  | 120        |
| 5.4      | Conclusions . . . . .   | 121        |
| <b>6</b> | <b><math>pK_a</math> calculations of <math>Cu_B</math> ligands in cytochrome <i>c</i> oxidase</b>               | <b>125</b> |
| 6.1      | Previous computational work on cytochrome <i>c</i> oxidase . . . . .  | 126        |
| 6.2      | Structure preparation and models . . . . .  | 128        |
| 6.2.1    | Preparation of X-ray structure of protein . . . . .   | 129        |
| 6.2.2    | Redox center models . . . . .   | 130        |

---

|  |            |
|--|------------|
| 6.3 Density Functional calculations . . . . .            | 133        |
| 6.4 Electrostatic calculations . . . . .                 | 134        |
| 6.4.1 Calculation of average $pK_a$ in protein . . . . . | 135        |
| 6.5 Results and Discussion . . . . .                     | 137        |
| 6.6 Conclusions . . . . .                                | 149        |
| <b>7 Concluding Remarks and Outlook</b>                  | <b>151</b> |
| <b>Bibliography</b>                                      | <b>153</b> |



## CHAPTER 1

---

# INTRODUCTION

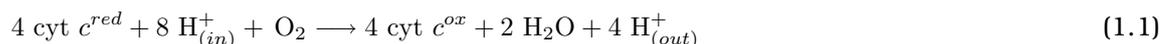
---

## 1.1 INTRODUCTION

### 1.1.1 CYTOCHROME *c* OXIDASE - A REDOX-DRIVEN MOLECULAR MACHINE

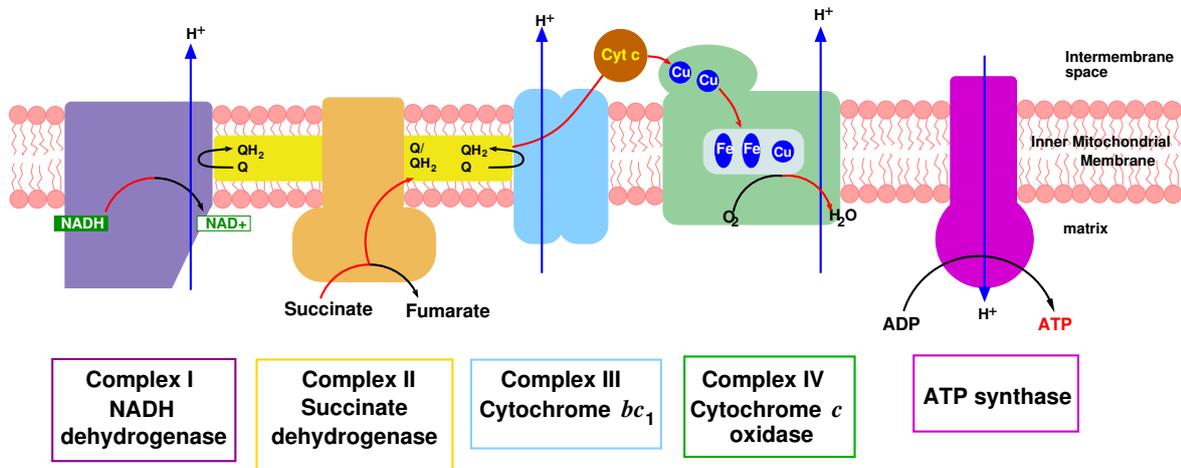
Cytochrome *c* oxidase is the terminal member of the electron transport system of mitochondria and many bacteria. It catalyzes the reduction of molecular oxygen to water and pumps protons across the membrane [7, 8]. In this process the membrane electrochemical proton gradient is generated; the energy stored by the proton gradient is subsequently utilized for ATP synthesis [9].

Cytochrome *c* oxidase is responsible for catalyzing the reduction of more than 95% of the oxygen taken up by aerobically growing higher organisms. Cytochrome *c* oxidase is structurally classified as a member of the superfamily of heme-copper containing terminal oxidases composed of 4 to 13 subunits whose largest and most hydrophobic subunits, I, II and III, are encoded by mitochondrial DNA. These subunits collectively have 18 hydrophobic segments forming membrane-spanning helices similar to those of cytochrome *b* [10]. The protein subunit I of most cytochrome *c* oxidases contain two heme *a* molecules, called heme *a* and heme *a*<sub>3</sub> and a copper B center (Cu<sub>B</sub>). Heme *a*<sub>3</sub> and Cu<sub>B</sub> form a binuclear center where molecular oxygen is reduced into water. Electrons from cytochrome *c* are first transferred to the copper A center (Cu<sub>A</sub>), which is located in the subunit II. The copper A center passes them to heme *a*, which transfers the electrons to the binuclear center, where the molecular oxygen is reduced into water (see Eq. (1.1)). The reduction of oxygen to water in the catalytic center of cytochrome *c* oxidases generates energy that is necessary for proton translocation from the mitochondrial matrix.



### 1.1.2 STRUCTURE AND FUNCTION OF CYTOCHROME *c* OXIDASE

Crystallization and X-ray diffraction analysis of the mitochondrial and of several bacterial cytochrome *c* oxidases have set the stage for understanding of the complex functions of this enzyme [7].



**Figure 1.1. The mitochondrial electron-transport chain.** The electrons are transferred from complex I to complex IV. Electrons are transferred between complexes I and III by the mobile electron carrier coenzyme Q (Q) and from complexes III to VI by the peripheral membrane protein cytochrome *c* (Cyt *c*). Complex II also transfers electrons to Q. The pathways of electron transfer (*red*) and proton pumping (*blue*) are indicated.

CYTOCHROME *c* OXIDASE IN MITOCHONDRIAL RESPIRATION. The mitochondrion is the site of eukaryotic oxidative metabolism. In oxidative phosphorylation, electrons are transferred from NADH (Nicotinamide Adenine Dinucleotide reduced form) or  $\text{FADH}_2$  (Flavin Adenine Dinucleotide, reduced form) to  $\text{O}_2$  via membrane-bound protein complexes. The free energy of the electron transfer reaction is coupled to ATP synthesis [10]. A schematic representation of electron transport chain is shown in Figure 1.1. Protein complexes embedded in the inner mitochondrial membrane catalyze the electron transfer from NADH to oxygen. These protein complexes are commonly referred as respiratory chain. Some of the compounds are highly mobile (coenzyme Q (Q) and cytochrome *c*) which shuttle electrons between the transmembrane protein complexes. The free energy released in the redox reactions is stored as electrochemical gradient of protons across the membrane. This proton gradient is utilized to generate ATP.



The respiratory chain in the inner mitochondrial membrane is commonly organized into four transmembrane complexes, namely NADH dehydrogenase (complex I), succinate dehydrogenase (complex II), cytochrome  $bc_1$  (complex III) and cytochrome *c* oxidase (complex VI). Two mobile electron carriers Q and cytochrome *c* shuttle electrons between the transmembrane protein complexes.

Complex I passes the electrons from NADH to Q. The electron transfer takes place along iron-sulfur (FeS) clusters. During the electron transfer, complex I pumps protons from the matrix to the intermembrane space.

Complex II also transfers electrons to  $\mathcal{Q}$ , however in the case of complex II, the source of the electrons is  $\text{FADH}_2$  produced in the citric acid cycle. Since the standard redox potential of FAD is slightly lower than that of  $\mathcal{Q}$ , complex II does not pump protons from the matrix to the intermembrane space. The electrons from complex I and complex II are shuttled to complex III by reduced  $\mathcal{Q}$  molecules.

Complex III (cytochrome  $bc_1$ ) passes the electrons to the next mobile carrier cytochrome  $c$  which shuttles the electrons to complex IV (cytochrome  $c$  oxidase). Complex IV passes the electrons to diatomic oxygen and thereby reducing it into two water molecules (final electron acceptor). Both cytochrome  $bc_1$  and cytochrome  $c$  oxidase translocate protons from the matrix to the intermembrane space.

#### THE THREE-DIMENSIONAL STRUCTURE OF HEME-COPPER OXIDASES

Heme-copper oxidases are membrane proteins found in the respiratory chain of aerobic organisms. They are the terminal electron acceptors coupling the translocation of protons across the membrane with the reduction of oxygen to water. The heme-copper oxidase superfamily is divided into two main branches based on the identities of the electron donating substrates: cytochrome  $c$  oxidases use a water soluble protein, cytochrome  $c$ , as electron donor whereas ubiquinol oxidases use a membrane soluble ubiquinol molecule as their electron donor. Both the members of the superfamily share important structural and functional features.

**STRUCTURE OF THE CYTOCHROME  $c$  OXIDASE FROM *Paracoccus denitrificans*.** The protein was originally isolated as a two-subunit enzyme complex from the cytoplasmic membrane of the soil bacterium *P. denitrificans* [11]. Michel and co-workers determined the structures of the reduced and oxidized *P. denitrificans* enzyme and found no structural difference [12]. The four subunits of oxidase has been crystallized in the presence of dodecyl maltoside as a complex with a monoclonal antibody fragment ( $F_v$ ) directed against an epitope on the hydrophilic domain of subunit II and its structure was determined at 2.8 Å [12, 13]. Later, the two-subunit complex structure was solved at 2.7 Å again using the  $F_v$  approach to increase the polar surfaces of the protein complex and undecyl maltoside as detergent [14].

**STRUCTURE OF THE CYTOCHROME  $c$  OXIDASE FROM *Rhodobacter sphaeroides*.** The structure of cytochrome  $c$  oxidase from *R. sphaeroides* has been solved by Iwata et al. [15]. The crystal structures were determined for the wild type and a mutant by replacing the glutamate residue 286 of subunit I by glutamine (see Figure 1.2).

**STRUCTURE OF THE CYTOCHROME  $c$  OXIDASE FROM BOVINE HEART.** The crystal structure of bovine heart cytochrome  $c$  oxidase at 2.8 Å was determined by Yoshikawa et al. [16–18] in 1995. The crystal structures of the bovine heart cytochrome  $c$  oxidase in the fully oxidized and the fully reduced states were determined by the same group in 1998 [16]. The protein is composed of 13 different subunits.

**UBIQUINOL OXIDASE.** Ubiquinol oxidases take part in the bacterial electron transport chain oxidizing ubiquinol into ubiquinone and reducing oxygen to water. These enzymes are one

set of the many alternative terminal oxidases in the branched prokaryotic electron transport chain. The overall structure of the ubiquinol oxidase is similar to that of the mammalian cytochrome *c* oxidase, with the addition of a polar ubiquinol-binding site embedded in the membrane.

The cytochrome *c* oxidase contains four redox centers: Cu<sub>A</sub>, heme *a*, heme *a*<sub>3</sub> and Cu<sub>B</sub>. The oxygen reduction takes place in the binuclear center Fe<sub>a3</sub>-Cu<sub>B</sub> and utilizes cytochrome *c* as an electron donor. A second major branch of the *aa*<sub>3</sub>-cytochrome *c* oxidase uses ubiquinol or menaquinol as the reducing substrate, and in these cases the Cu<sub>A</sub> center is absent. Such enzymes are the cytochrome *bo*<sub>3</sub> of *Escherichia coli* [19, 20] and cytochrome *aa*<sub>3</sub>-600 of *Bacillus subtilis* [21–23]. The cytochrome *aa*<sub>3</sub>-type terminal quinol oxidase of *B. subtilis* catalyzes the four-electron reduction of oxygen into water.

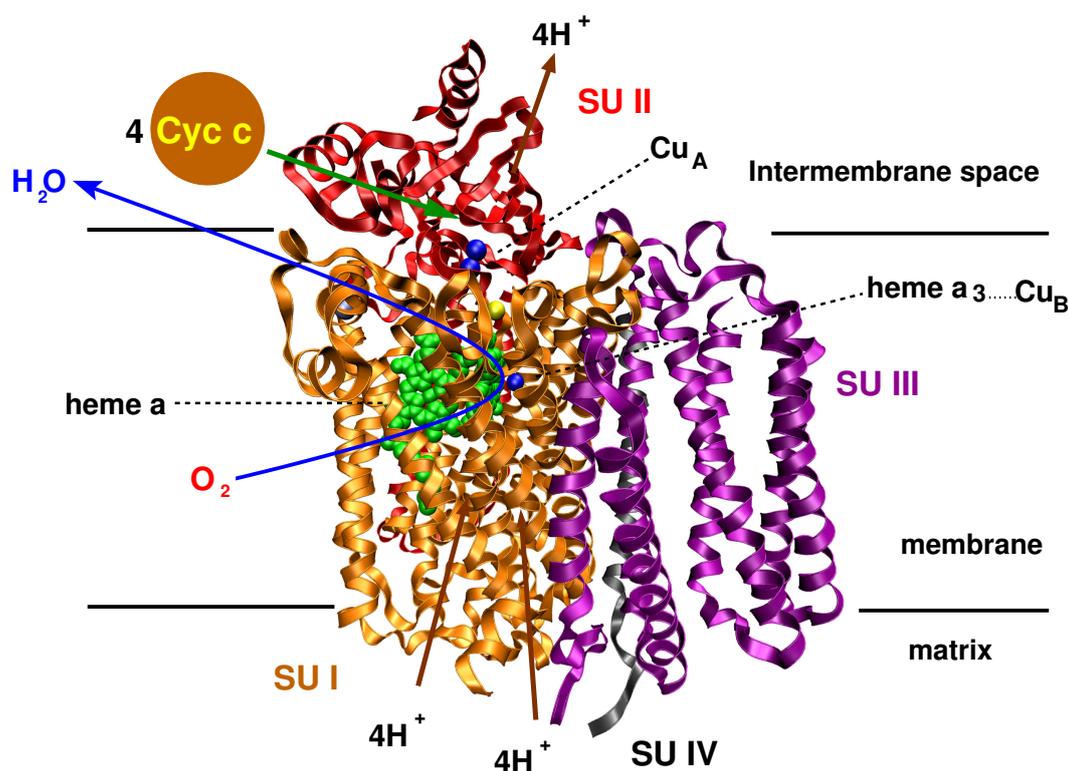
STRUCTURE OF THE *ba*<sub>3</sub>-CYTOCHROME *c* OXIDASE FROM *Thermus thermophilus*. The structure of the *ba*<sub>3</sub>-cytochrome *c* oxidase from *T. thermophilus* has been reported at a resolution of 2.4 Å [24]. The crystal structures of recombinant cytochrome *ba*<sub>3</sub>-cytochrome *c* oxidase from *T. thermophilus* was reported at 2.3 Å by Hunsicker-Wang and co-workers [25]. The model of the *ba*<sub>3</sub>-cytochrome *c* oxidase is composed of three protein subunits I, II and IIa. The main part of the complex is formed by subunit I with 13 transmembrane helices, which binds the heme *b*, heme *as*<sub>3</sub> as well as Cu<sub>B</sub>. The *a*-type heme of the *ba*<sub>3</sub>-oxidase corresponds to the heme *as* present in the SoxB-type terminal oxidases [24]. Heme *b* is the simplest protoheme containing a low-spin iron with two histidine residues as axial ligands. The binuclear center is formed between heme *as*<sub>3</sub> and Cu<sub>B</sub> as in the cytochrome *c* oxidase.

STRUCTURE OF THE *bo*<sub>3</sub>-CYTOCHROME *c* OXIDASE FROM *Escherichia coli*. The structure of cytochrome *bo*<sub>3</sub> oxidase from *E. coli* was reported at 3.5 Å by Abramson et al. [26]. The cytochrome *bo*<sub>3</sub> ubiquinol oxidase is a four-subunit heme-copper oxidase that catalyzes the four-electron reduction of O<sub>2</sub> to water and functions as a proton pump. All the redox centers are located in subunit I, with a low spin protoheme *b* acting as an electron donor to a binuclear center that is composed of an *o*-type heme, heme *o*<sub>3</sub>, and a copper ion Cu<sub>B</sub>. Subunits of I, II and III of ubiquinol oxidase are homologous to the corresponding subunits in the *aa*<sub>3</sub>-type cytochrome *c* oxidase, and the ligands of the two heme groups and the Cu<sub>B</sub> have been identified as histidine residues. In contrast to the cytochrome *c* oxidase, subunit II of the ubiquinol oxidase has neither Cu<sub>A</sub> center, nor a cytochrome *c* binding site. Instead, the heme *b* receives electrons directly from a membrane solubilized ubiquinol molecule.

#### SUBUNITS OF CYTOCHROME *c* OXIDASE

The subunits of cytochrome *c* oxidase from *R. sphaeroides* are discussed in the following section.

Subunit I (see Figure 1.2) of bacterial cytochrome *c* oxidase is largely embedded in the membrane, with its 12 transmembrane helices shaped in a three-winged propeller arrangement [7, 12]. The N-terminus and the long, exposed C-terminus of the polypeptide face the cytoplasmic side. The three redox centers, the two *a*-type hemes and the copper B center, are ligated by amino acid side chains of this subunit. Histidines are the axial ligands to the low-spin heme *a*, whereas a histidine and a presumed hydroxyl or a water molecule are ligands



**Figure 1.2.** The X-ray structure of cytochrome *c* oxidase from *Rhodobacter sphaeroides*. The electrons are transferred from cytochrome *c* to copper A center. Protons are pumped from matrix to intermembrane space and four more protons are delivered to the binuclear center. Oxygen is reduced to water in the binuclear center (heme  $a_3$ ...Cu<sub>B</sub>).

to the high-spin heme  $a_3$  moiety. Both hemes are oriented perpendicular to the membrane plane. Heme  $a_3$ , together with a copper ion (Cu<sub>B</sub>) in its immediate vicinity, forms the binuclear center where oxygen binding and reduction takes place.

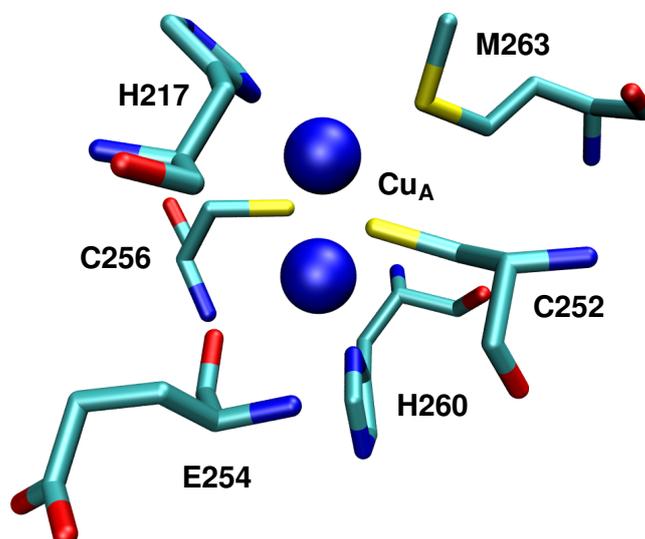
Subunit II (see Figure 1.2) of bacterial cytochrome *c* oxidase has a bipartite structure. The N-terminal has two transmembrane helices followed by a hydrophilic, 10-stranded  $\beta$ -barrel domain extending into the periplasm, comprising the Cu<sub>A</sub> center. The Cu<sub>A</sub> center contains two copper ions in a mixed-valence (Cu<sup>I</sup>.Cu<sup>II</sup>) state and 2.6 Å apart, giving rise to a characteristic EPR (electron paramagnetic resonance) signal in the oxidized state which was observed in the other enzymes as well [27]. The two copper ions are bridged by two cysteine thiolates.

Subunit III (see Figure 1.2) of bacterial cytochrome *c* oxidase is fully embedded in the membrane. No redox cofactors are associated with this subunit. When both subunits III and IV are removed from the *P. denitrificans* cytochrome *c* oxidase there is no loss in the catalytic functions. It has been suggested that subunit III stabilizes the integrity of the binuclear center in subunit I [28]. The cleft of the subunit III may be a binding site for other membrane proteins. Subunit III may be involved in assembly of the oxidase or form the entrance to an oxygen channel leading to the active site [29]. It possesses seven transmembrane helices that are divided by a large V-shaped cleft into two bundles, one formed by the first two helices, and the other by helices III to VII. In this cleft, lipid molecules are found to be firmly bound to the

conserved residues. In mammalian oxidase three hydrophobic channels were proposed and these hydrophobic channels are suggested to be the potential pathways for oxygen to reach the binuclear center [18]. The oxygen channels start at the protein-membrane interface near the center of the lipid bilayer, where oxygen solubility is much higher than in the aqueous phase. One of such channel has also been identified in the bacterial oxidase [29]. It starts in the V-shaped cleft of the subunit III directly above a tightly bound lipid molecule and leads through subunit I into the binuclear site.

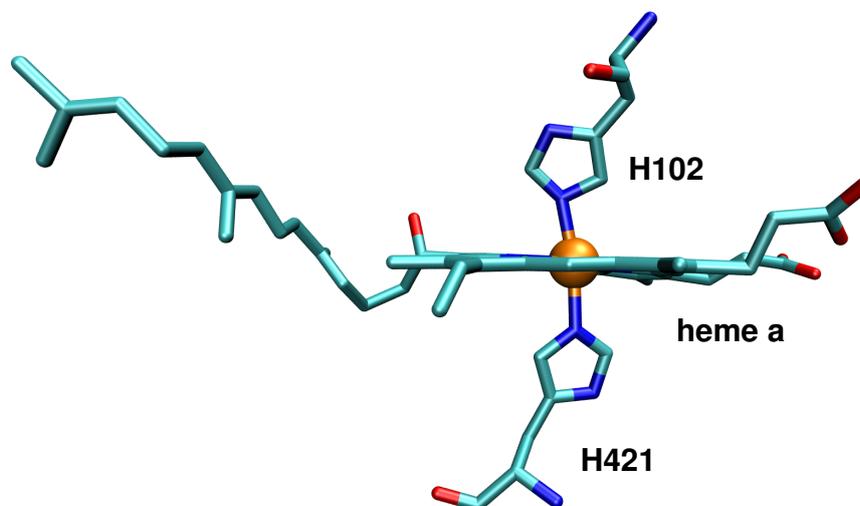
Subunit IV (see Figure 1.2) of the bacterial enzyme of both *P. denitrificans* and *R. sphaeroides* consists of a single transmembrane helix in contact with subunit I and III. The function of this small subunit is unknown [12].

#### THE REDOX CENTERS OF CYTOCHROME *c* OXIDASE



**Figure 1.3. The copper A ( $Cu_A$ ) center of cytochrome *c* oxidase from *R. sphaeroides*.** The copper A ( $Cu_A$ ) center with its ligating residues is shown. The electrons are transferred from cytochrome *c* to the copper A center. The amino acid numbering of *R. sphaeroides* is used throughout this thesis.

**COPPER A ( $Cu_A$ ) CENTER.** The  $Cu_A$  center is located 8 Å above the membrane surface. The  $Cu_A$  center contains two copper ions (see Figure 1.3). These copper ions are bridged by two cysteines sulfur atoms and have additional protein ligands. The two copper ions of the  $Cu_A$  center are coordinated by two His, one Met, a backbone carbonyl oxygen of a Glu and two bridging Cys residues. The spectroscopic measurements indicate that in the reduced form of the  $Cu_A$  center both copper ions are in their  $Cu(I)$  state whereas in the fully oxidized form, the newly acquired electron appears to be delocalized between the two copper ions such that they assume the mixed [ $Cu^{1.5+} \dots Cu^{1.5+}$ ] state.

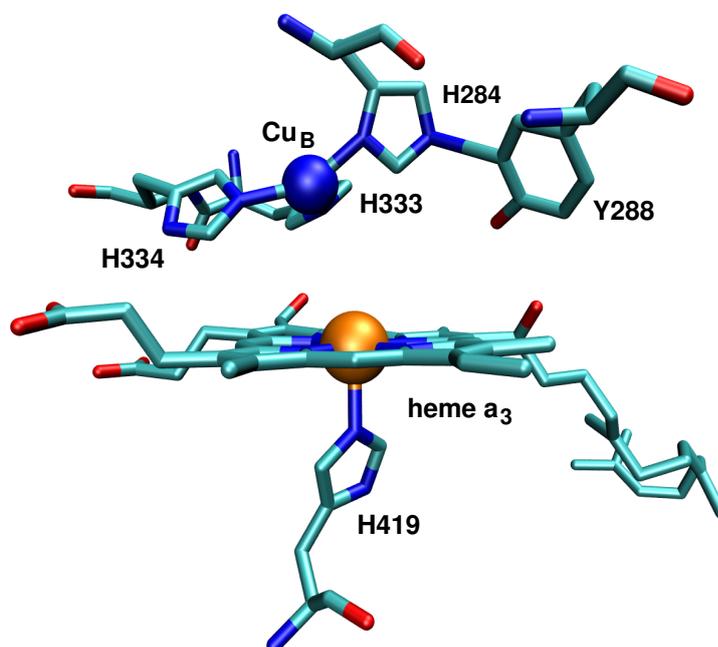


**Figure 1.4. The heme *a* of cytochrome *c* oxidase from *R. sphaeroides*.** The heme *a* porphyrin ring system is shown with two histidine residues as axial Fe-ligands. The amino acid numbering of *R. sphaeroides* is used throughout this thesis.

**HEME *a* CENTER.** The heme *a* center consists of two histidine residues as axial iron-ligands (see Figure 1.4). It transfers the electrons from  $Cu_A$  to the binuclear center. The heme is non-covalently bound to the protein and heme *a* contains a formyl group and a hydrophobic hydroxyethyl-farnesyl group. The heme environments in *P. denitrificans* and bovine heart cytochrome *c* oxidase are very similar.

**HEME *a*<sub>3</sub> AND COPPER B ( $Cu_B$ ) CENTER: THE BINUCLEAR CENTER.** The binuclear center is formed by heme *a*<sub>3</sub> and  $Cu_B$  which is the catalytic center for  $O_2$  reduction (see Figure 1.5). The heme *a*<sub>3</sub> iron appears to be five fold coordinated with a histidine of subunit I. The copper ion in the  $Cu_B$  center is coordinated by  $N_\epsilon$  atoms of His334 and His333 and  $N_\delta$  atom of His284 and the histidines are arranged in an equilateral triangle, centered on  $Cu_B$ . Molecular oxygen is supposed to bind between the heme *a*<sub>3</sub> iron and  $Cu_B$ . The iron of heme *a*<sub>3</sub> is 0.36 Å out of the heme plane in *P. denitrificans* but almost within the plane in the bovine heart. Crystallographic studies on cytochrome *c* oxidase have revealed a unique and unexpected posttranslational modification in the enzyme active site [16]. A tyrosine (Y288) which is located very near to the binuclear center is covalently linked to the  $N_\epsilon$  of histidine residue (H284) which also serves as a ligand to the  $Cu_B$  center (see Figure 1.5). The cross-linked tyrosine and histidine was observed in both bacterial and mammalian cytochrome *c* oxidases.

**AMINO ACIDS AS REDOX CENTERS.** In addition to four metal redox centers, there are other amino acids near the active site of the enzyme that could conceivably form radicals and act as a redox-active centers in cytochrome *c* oxidase. The tyrosine (Y288) is the most proximate amino acid which is cross-linked to the  $Cu_B$  histidine (H284) [30]. The EPR evidence



**Figure 1.5. The heme  $a_3$  and  $Cu_B$  (binuclear center) of cytochrome  $c$  oxidase from *R. sphaeroides*.** Heme  $a_3$  and copper B ( $Cu_B$ ) forming the binuclear center are shown.

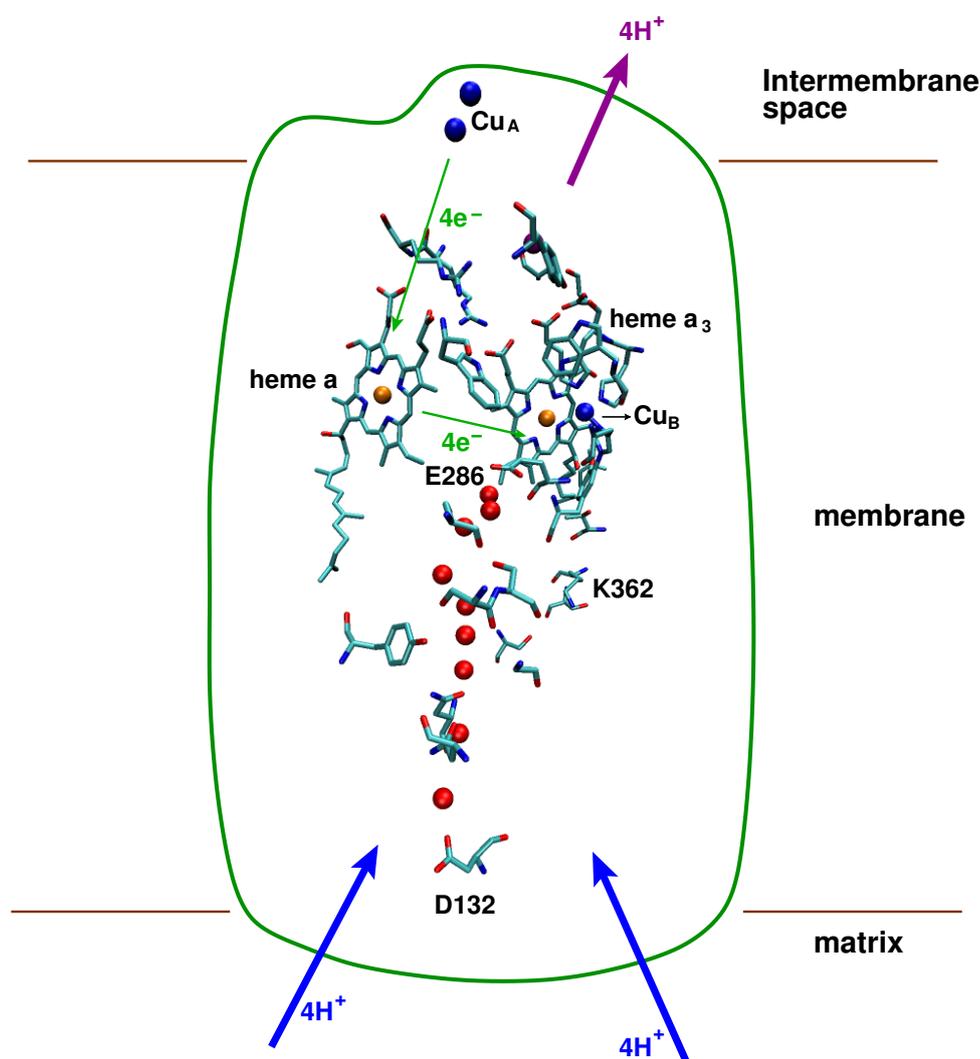
has recently been presented showing the formation of a tyrosyl radical in the enzyme upon oxidation of the binuclear center by peroxide [31]. In addition to tyrosine, there is a very highly conserved tryptophan that could similarly serve as an electron source during catalysis [32, 33].

#### BINDING SITES OF NON-REDOX ACTIVE METAL CENTERS

Non-redox active metal ion binding sites have been characterized in the bacterial and in the mitochondrial oxidases. The enzyme isolated from *P. denitrificans* grown under standard medium conditions contains a manganese ion, giving rise to a characteristic EPR signal [34, 35]. The manganese ion itself is not required for the function of the oxidase and replaced by a magnesium ion. Mutational studies on the magnesium ion binding site [36] located at the hydrophilic interface between subunits I and II in the vicinity of the  $Cu_A$  center show a moderate decrease in electron transfer activity. A similar site is encountered in the *R. sphaeroides* oxidase and in the mammalian enzyme structure. In the structures of bacterial and the mitochondrial cytochrome  $c$  oxidases, another metal ion binding site was observed at the periplasmic side in the transmembrane helix I of subunit I [37–39]. In *P. denitrificans* and *R. sphaeroides* cytochrome  $c$  oxidase, this site was shown to be populated by a tightly bound calcium ion. In bovine heart oxidase the binding site is reasonably assigned to be a sodium ion. Additionally, the mammalian cytochrome  $c$  oxidase contains a zinc ion binding site. The zinc ion is bound by a nuclear encoded subunit on the matrix side of the membrane.

## 1.2 THE GENERAL MECHANISM: ELECTRON TRANSFER AND PROTON TRANSLOCATION

A wide range of experimental studies have been done to understand the mechanism of electron transfer and proton translocation pathways in cytochrome *c* oxidase. Based on a large body of spectroscopic and kinetic data [8, 40–42], it is generally accepted that internal electron transfer in cytochrome *c* oxidase proceeds in four steps (see Figure 1.6). The cytochrome *c* transfers the electrons to the  $\text{Cu}_A$  center which is located in the subunit II and the  $\text{Cu}_A$  center transfers the electrons to the three redox center in subunit I, heme *a*, heme  $a_3$  and  $\text{Cu}_B$ . The protons and electrons are delivered to the binuclear center for oxygen reduction.



**Figure 1.6. The schematic representation of the redox centers of cytochrome *c* oxidase.** The electrons transfer from cytochrome *c* to the  $\text{Cu}_A$  center to heme *a* and heme  $a_3$  to the binuclear center are shown. Protons are translocated from the matrix to the intermembrane space shown in purple arrows and four protons are delivered into the binuclear center shown in brown arrow.

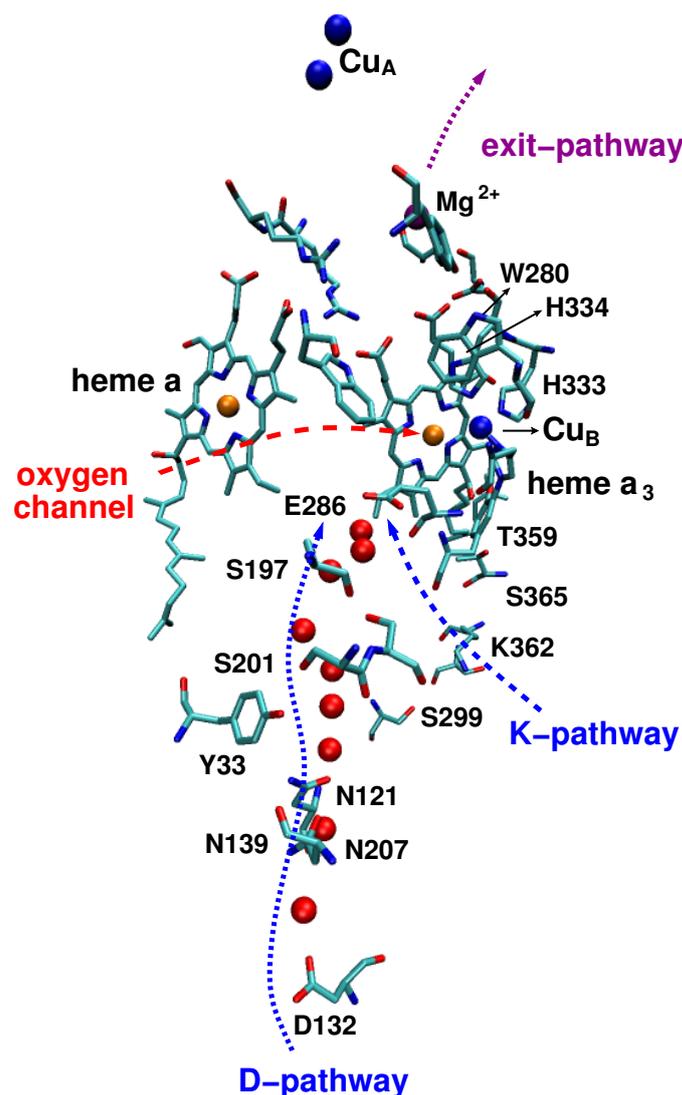
## ELECTRON TRANSFER PATHWAYS

Electrons enter cytochrome *c* oxidase exclusively via the Cu<sub>A</sub> center located in subunit II. It was proven by mutating the residues which serve as ligands for Cu<sub>A</sub> center [43]. The subunit II has been described [44] as the main docking site for cytochrome *c* oxidase, the peripheral protein complex which loosely bound to the inner membrane space and shuttles electrons from the cytochrome *bc*<sub>1</sub> complex (complex III). The interaction of cytochrome *c* with cytochrome *c* oxidase is based on long-range electrostatic preorientation between the highly basic cytochrome *c* and an extended lobe of acidic residues on the surface of subunit II close to the Cu<sub>A</sub> site of cytochrome *c* oxidase.

A surface exposed tryptophan (in subunit II) is considered to be crucial for electrons to enter the cytochrome *c* oxidase complex from cytochrome *c*. A series of specific mutations in this conserved tryptophan in *P. denitrificans* and *R. sphaeroides* cytochrome *c* oxidases showed drastically diminished presteady state [44–46]. Electrons from cytochrome *c* travel to the Cu<sub>A</sub> center and then to the heme *a*. The distance between Cu<sub>A</sub> and heme *a* is almost 19.5 Å. The heme *a* transfers the electron to the binuclear center where the water formation takes place. Both heme ring systems are only separated by a distance of 4.5 Å. An alternative pathway, a direct transfer of electrons to heme *a*<sub>3</sub> is strongly unfavorable due to the long distance of 22.1 Å from Cu<sub>A</sub> to the iron in heme *a*<sub>3</sub>; However a mutant (R54M in *P. denitrificans*) significantly lowers the redox potential of the heme *a* and this pathway has been suggested to yield a turnover activity in the mutant of 2% compared to the WT [47]. Several suggestions have been made for discrete electron pathways [12, 18] but still the electron transfer pathways are not clear.

## PROTON TRANSLOCATION PATHWAYS

For every molecular oxygen a total of eight protons are consumed from the matrix side of the mitochondrion: four protons are delivered to the binuclear center for water formation and another four protons are translocated across the membrane (see Figure 1.7). The pumped protons are required to travel about 50 Å through the protein. Proton transfer through the proteins take place by hopping along a series of hydrogen bonds which can be considered to be a proton wire [48, 49]. Such proton wire however, need not to be a stable structural element that can be clearly deciphered from the X-ray structures. The hydrogen bonds might be transient and the internal water molecules are almost certainly the critical elements of the proton wire. Therefore, the connectivity of the proton wires as they appear in the X-ray models may not give the complete picture. Mutagenesis studies and electrostatic calculations support that atleast two separate pathways should exist in cytochrome *c* oxidase which are called the K-channel and D-channel (see Figure 1.7). The names are derived from conserved amino acid residues within each of the respective channels. When a widely conserved aspartate residue present in the D-channel (D132), located at the periplasmic face of the subunit I was mutated, the enzymes electron transfer activity was diminished and partly uncoupled from proton pumping. Another mutation of a conserved lysine in the K-channel (K362), located within the membrane section of the subunit I led to a complete loss of the electron transfer.



**Figure 1.7. The K- and D-pathways for the proton uptake in cytochrome *c* oxidase.** The K-channel starts at the mitochondrial matrix and extends to the region of the binuclear center. The key residue lysine (K362) is shown. The D-channel extends from the matrix entrance aspartate 132 to a buried glutamic acid 286 and to the binuclear center. The redox centers  $\text{Cu}_A$ , heme *a*, heme *a*<sub>3</sub> and  $\text{Cu}_B$  are also shown.

**THE K-CHANNEL.** This channel starts from mitochondrial matrix and extends to the region of the binuclear center. The key residue lysine (K362 in the *R. sphaeroides*) and Thr359 lead to the hydroxy group of Tyr288 at the active site [14, 16]. The Tyr288 is cross-linked to one of the histidine which serves as a ligand to  $\text{Cu}_B$  and therefore in a position to deliver protons directly into the active site.

**THE D-CHANNEL.** The D-pathway (see Figure 1.7) extends from the matrix entrance Lys132 to a buried glutamic acid 286. From the glutamic acid it is not clear how the protons reach the binuclear center. Site-directed mutagenesis at several positions along the pathway has pro-

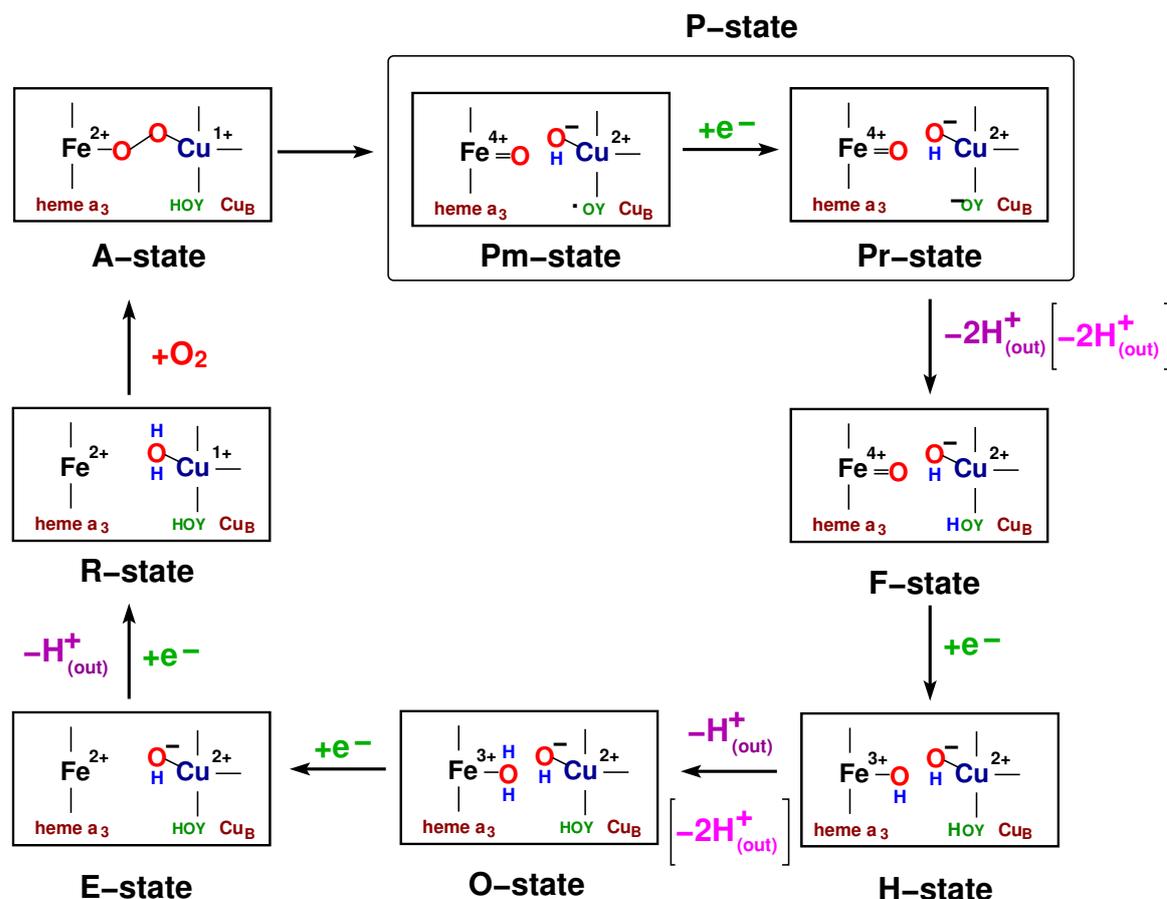
vided a functional evidence for the proton conduction [8, 50–52]. The importance of the highly conserved residue Glu286 in proton conduction has been acknowledged for many years and its role in redox-dependent shuttling of protons into the binuclear center has been suggested recently [42]. A mutant crystal structure (E286Q) of cytochrome *c* oxidase from *R. sphaeroides* shows a loss of a hydrogen bond [15], along with several other subtle conformational changes and relocation of water molecules. There are number of water molecules within the protein which provide a reasonable pathway for proton transfer between residues Lys132 and Glu286 that are separated by 25 Å.

Many details of the K- and D-pathways have been studied over the years using mutational and spectroscopic techniques but the precise mechanism for delivering and pumping of protons is not yet clear. The role of these two pathways (K and D), the proton pumping in relation to the oxygen reduction cycle and how they are correlated with structure of oxygen intermediates are still remain unknown.

**THE OXYGEN CHANNEL.** The area nearest to  $\text{Cu}_B$  comprises Phe237, Trp288, Trp236 and His291 residues and the side chains of these residues as well as other side chains nearby provide a hydrophobic environment. This region has been proposed to be a possible oxygen channel [18]. The next proposed oxygen channel starts from His151 of subunit I and leads to the lipid pool of subunit III. A network of hydrophobic amino acids Leu199, Phe67, Phe238, Phe235, Trp126 and Trp236 [18] in this region are well conserved. The third candidate for the proposed oxygen channel includes the hydroxyethyl-farnesyl group of heme  $a_3$ , with the terminal being on the surface of subunit I between the two attaching point of the two trans-membrane helices of subunit II. The proposed oxygen channel in *R. sphaeroides* is shown in Figure 1.7.

Inspection of the electron density map of cytochrome *c* oxidase from *P. denitrificans* reveals a highly hydrophobic channel from the binuclear site to the edge of subunit I where it appears to exit between trans-membranous helices IV and V. This channel bears an interesting relationship to the structure of subunit III and its positioning relative to subunit I. Riistame et al. performed a systematic mutagenesis study to test whether this channel might form the path for oxygen diffusion into binuclear site [29]. They performed site-directed mutagenesis of the quinol-oxidizing cytochrome  $bo_3$  of *E. coli*. The mutation studies on conserved valine (V279I mutation) located in this oxygen path has shown to dramatically increase the  $K_M$  of oxygen compared to WT enzyme. These results suggest that the open structure of subunit III and its position relative to the channel may have a special function in  $\text{O}_2$  transport.

**THE EXIT CHANNEL.** The exit pathway for protons which are pumped to the inner membrane space is less understood. It most likely comprises the area above both heme propionates. There is an extensive hydrogen bond network involving several arginine residues and the heme propionates which can also in principle provide a proton storage area and a part of the exit pathway [53]. Mutations of some of these arginines clearly perturb proton pumping in the cytochrome  $bo_3$  quinol oxidase [54].



**Figure 1.8. The qualitative overview of the molecular steps taking place in the active site of cytochrome *c* oxidase during catalysis.** In the initial state, all four redox centers ( $Cu_A$ , heme  $a$ , heme  $a_3$  and  $Cu_B$ ) are oxidized (**O**). Single electron reduced (**E**), reduced (**R**), adduct after oxygen binding (**A**), “peroxy” state (**P**), oxo-ferryl state (**F**) and **H** state are shown. The intermediate states are further detailed in their electronic configuration at the binuclear center: heme  $a_3$ ,  $Cu_B$ , and a tyrosine side chain of subunit I are shown. The heme  $a_3$  is represented by the central Fe atom and by a line defining the porphyrin ring. YO $^{\bullet}$  stands for the neutral radical of the His284-Y288 cross-link. YO $^-$  for the tyrosinate. The  $H^+_{(out)}$  show only the delivery of the last two electrons being coupled to the proton pump as proposed by Wikström [49] and  $H^+_{(out)}$  show coupling during the reductive part of the catalytic cycle as proposed by Michel [49].

#### PROPOSED CATALYTIC CYCLE OF CYTOCHROME *c* OXIDASE

The catalytic cycle of cytochrome *c* oxidase is traditionally viewed as consisting of two asymmetric halves. The first part includes the initial reduction of both heme  $a_3$  and  $Cu_B$  in the binuclear center, the initial binding of oxygen and its reduction to the intermediate called P state. The second part of the catalytic cycle includes two single electron transfer steps, first to convert P state to F state and then converting O state via H state. Intermediates P and F were first discovered and characterized by Wikström [8].

**A STATE.** The proposed catalytic cycle for the molecular steps taking place in the active site of cytochrome *c* oxidase during catalysis is given in Figure 1.8. Oxygen binds to the reduced binuclear center and forming the A state (adduct after oxygen binding).

**P STATE.** A state leads to the so-called peroxy state P. The P state appears to exist in two forms,  $P_m$ , a two electron reduced and the  $P_r$  state. The  $P_m$  state resulted from the internal electron redistribution and the tyrosine is proposed to be neutral radical (see Figure 1.8). Weng and Baker [55] were the first to suggest that the P state is not a peroxy state but the oxo-ferryl state with hydroxyl group bound to the  $Cu_B$  center. They suggested that a tryptophan is the source of the missing electron in analogy to cytochrome *c* peroxidase. Later Kitagawa and co-workers [56] provided evidence by Raman spectroscopy that P state is a hydrogen-bonded oxo-ferryl state. The structure of the  $P_m$  state shown in Figure 1.8 has been proposed several times [57]. The existence of the covalent tyrosine-histidine cross-link should be taken as the evidence that a tyrosine radical is formed during the catalytic cycle of cytochrome *c* oxidase, because the cross-linking of tyrosines is typical of a radical reaction catalyzed by peroxidase. The  $P_r$  state of the binuclear center may be considered as a “high-energy” state, the decomposition of which will drive proton translocation [58]. The subsequent transformation of the  $P_r$  state into F state is coupled to translocation of the proton across the membrane, net uptake of another proton into the binuclear center and an accompanying generation of transmembrane electric potential [59].

**F STATE TO O STATE.** The F state results from the third electron transfer from cytochrome *c* together with the acquisition of two protons which convert the tyrosine radical to phenolate state. There are overwhelming spectroscopic data that heme  $a_3$  is in the oxyferryl form in F state (see Figure 1.8). However, there appears to be multiple forms and ambiguity with respect to the redox status and the protonation state of groups in the immediate vicinity of the active site, depending on how F state is generated. A fourth and final electron transfer and proton acquisition yields the oxidized O state via H state. The H state intermediate possesses hydroxyl group at the heme  $a_3$ .

**O AND R STATE.** In O state, all the four redox centers are fully oxidized and the heme  $a_3$  iron is bound with the water ligand and  $Cu_B$  center is bound with hydroxyl group. The tyrosine is covalently cross-linked to the  $Cu_B$  ligand of His284. The oxidized binuclear complex is reduced to R state via the formation of the one-electron reduced E state. Protons are translocated from matrix during this process.

Wikström and co-workers worked to establish much of the conceptual and an experimental framework for studying the oxidase mechanism, including important experimental evidence that proton pump is coupled only to the  $P \rightarrow F$  and  $F \rightarrow O$  state transitions and that each of these one electron redox steps results in the pumping of two protons [8, 60]. This paradigm has been generally accepted for the past decade, though some problems have been pointed out and discussed [61]. Now the paradigm that pumping is coupled only to the last two steps in the catalytic cycle is being challenged by Michel [53] on the basis of a re-evaluation of several key experiments. The conclusion of the re-evaluation is that the  $F \rightarrow O$  state transition is coupled to the pumping of only one proton and not two. Although considerable progress has

been made concerning the proton input pathways in cytochrome *c* oxidase, very little is known about the mechanism and how the proton pump actually works.

#### PROTON-COUPLED ELECTRON TRANSFER

The cytochrome oxidase energetically couple the electron transfer reactions associated with reduction of oxygen to water and pumps proton across the membrane. Even though a vast amount of structural and functional information of cytochrome *c* oxidase are available from experimental and theoretical data [8, 14, 16, 50–52], actual step of coupling the the redox reactions to the proton translocation is poorly understood. How the protons are delivered exactly to the binuclear center, the proton translocation and exit pathways and the chemical intermediates involved in the catalytic cycle are still remain unclear.

### 1.3 OUTLINE OF THE THESIS

The aim of the thesis is to understand the proton pumping mechanism of cytochrome *c* oxidase. Although the structures of cytochrome *c* oxidase has been solved for several organism, the molecular mechanism of proton pumping remains unclear. In this thesis, the reaction mechanism of cytochrome *c* oxidase is analyzed by combining the density functional theory (DFT) and continuum electrostatic calculations.

The theory behind the electrostatic calculations, Poisson-Boltzmann equation, titration behavior calculations and DFT are described in chapter 2 and different charge methods are reported in chapter 3. Protonation probabilities and redox potentials of cytochrome *c* nitrite reductase which serves as a simpler electron transfer system to study cytochrome *c* oxidase are discussed in chapter 4.

To investigate the reaction mechanism and the role of the histidines bound with the  $\text{Cu}_B$  center and the histidine coordinating the heme  $a_3$ , the  $\text{p}K_a$  calculations were performed on the  $\text{Cu}_B$  and heme  $a_3$  using DFT in combination with continuum electrostatic models. The  $\text{p}K_a$  values are calculated using different basis sets and different solvation models. The Finite Difference Poisson Boltzmann (FDPB) method and the conductor-like polarizable continuum model (C-PCM) are used to determine the solvation free energies in aqueous solution. The influence of different charges, basis sets and solvation models on the  $\text{p}K_a$  calculations of the  $\text{Cu}_B$  have been studied (Chapter 5). The  $\text{p}K_a$  values of the  $\text{Cu}_B$  ligands and heme  $a_3$  center in aqueous solution are reported in Chapter 5.

In chapter 6 the average  $\text{p}K_a$  values of the  $\text{Cu}_B$  ligands in the presence of different heme  $a_3$  redox states in cytochrome *c* oxidase are reported.

This thesis provides insights into the role of  $\text{Cu}_B$  ligands in the reaction mechanism of cytochrome *c* oxidase. The most accurate methods are reported for the  $\text{p}K_a$  calculations of the  $\text{Cu}_B$  and heme  $a_3$  center. This thesis is thus a step towards a better understanding of the role of the  $\text{Cu}_B$  ligand in the reaction mechanism of cytochrome *c* oxidase.



## CHAPTER 2

---

# COMPUTATIONAL METHODS

---

## 2.1 ACID-BASE AND REDOX REACTIONS EQUILIBRIA

### 2.1.1 FUNDAMENTAL DESCRIPTION OF ACID-BASE AND REDOX EQUILIBRIA

Biological molecules, such as proteins and nucleic acids bear numerous functional groups such as carboxyl and amino groups that can undergo acid-base reactions. The dissociation of a proton from a monoprotic acid is generally given by:



The free energy change ( $\Delta G_a$ ) of this reaction can be related to the equilibrium constant ( $K_a$ ),

$$\Delta G_a = -RT \ln K_a \quad (2.2)$$

$$K_a = \frac{[\text{H}^+][\text{A}^-]}{[\text{HA}]} \quad (2.3)$$

The pH of the solution is defined as the negative decadic logarithm of the hydrogen ion concentration and the  $\text{p}K_a$  of an acid is defined as the negative decadic logarithm of the  $K_a$  values.

$$\text{pH} = -\log[\text{H}^+] \quad (2.4)$$

$$\text{p}K_a = -\log K_a \quad (2.5)$$

The Henderson-Hasselbalch equation combines Eq. (2.4) and (2.5) to

$$\text{pH} = \text{p}K_a + \log \frac{[\text{A}^-]}{[\text{HA}]} \quad (2.6)$$

The protonation probability is given by [62]:

$$p_{\text{prot}} = \frac{[\text{HA}]}{[\text{HA}] + [\text{A}^-]} = \frac{10^{\text{p}K_a - \text{pH}}}{1 + 10^{\text{p}K_a - \text{pH}}} \quad (2.7)$$

The  $pK_a$  value of an acid is the pH value at which the concentration of the protonated and deprotonated forms of the acid equal. Similar to the above equilibrium protonation reaction, the equilibrium between the redox reaction is,



The reduction equilibrium constant  $K_{ET}$  for this reaction is,

$$K_{ET} = \frac{[A_{red}^-]}{[A_{ox}][e^-]} \quad (2.9)$$

Eq. (2.9) is analogous Eq. (2.6). The standard redox potential of the above redox reaction is given by:

$$E_0 = \frac{RT}{F} \ln K_{ET} \quad (2.10)$$

The redox potential of the solution is given by,

$$E = -\frac{RT}{F} \ln[e^-] \quad (2.11)$$

The Nernst equation combines the standard redox potential  $E_0$  and the redox potential of the solution  $E$

$$E = E_0 + \frac{RT}{F} \ln \frac{[A]}{[A^-]} \quad (2.12)$$

where  $F$  is the Faraday constant,  $R$  is the gas constant and  $T$  is the temperature.

The probability of finding group A in the reduced state is therefore given by [62]:

$$p_{red} = \frac{[A^-]}{[A^-] + [A]} = \frac{\exp\left(\frac{RT}{F}(E^0 - E)\right)}{1 + \exp\left(\frac{RT}{F}(E^0 - E)\right)} \quad (2.13)$$

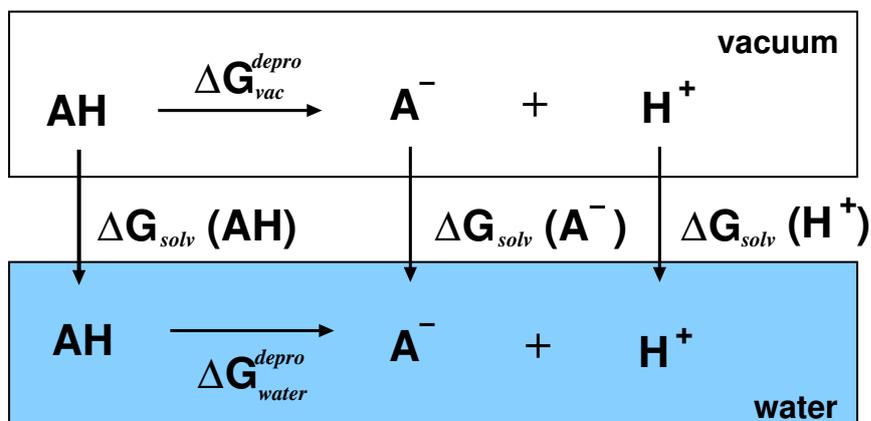
### 2.1.2 COMPUTATION OF ACID-BASE AND REDOX EQUILIBRIA

The  $pK_a$  is directly related to the free energy of the deprotonation reaction in aqueous solution  $\Delta G_{water}^{depro}$  by the following Eq. (2.14):

$$pK_a = \frac{1}{\ln 10 k_B T} \Delta G_{water}^{depro} \quad (2.14)$$

The  $\Delta G_{water}^{depro}$  can be expressed as a sum of two contributions: the solvation energy difference  $\Delta \Delta G_{solv}^{depro}$  between the associated and the dissociated system and the gas phase deprotonation energy  $\Delta G_{vac}^{depro}$ . These terms can be obtained from the thermodynamic cycle (see Figure 2.1).

$$pK_a = \frac{1}{\ln 10 k_B T} (\Delta G_{vac}^{depro} + \Delta \Delta G_{solv}^{depro}) \quad (2.15)$$



**Figure 2.1. Thermodynamic cycle to calculate absolute  $pK_a$  values.** The free energy of dissociation is calculated in vacuum and the reactant and products are then transferred from vacuum to water. The free energy of dissociation of a proton from an acid in water ( $\Delta G_{solv}^{depro}$ ) is calculated indirectly.

The solvation energy difference  $\Delta\Delta G_{solv}^{depro}$  is obtained from Eq. (2.16)

$$\Delta\Delta G_{solv}^{depro} = \Delta G_{solv}(A^-) + \Delta G_{solv}(H^+) - \Delta G_{solv}(AH) \quad (2.16)$$

The solvation energy of the protonated and deprotonated states,  $\Delta G_{solv}(A^-)$  and  $\Delta G_{solv}(AH)$  can be calculated by solving the Poisson-Boltzmann equation Eq. (2.19). The solvation energy of a proton is measured experimentally from the potential of the standard hydrogen electrode. The solvation energy of  $-264.6$  kcal/mol [63] is used for the proton in the present study.

The gas phase protonation energy  $\Delta G_{vac}^{depro}$  is given by Eq. (2.17)

$$\Delta G_{vac}^{depro} = \Delta H_{vac}^{depro} + \Delta H_{vib}^{depro} + H_{trans}(H^+) + \Delta(pV) - T[S(H^+)] \quad (2.17)$$

where  $\Delta H_{vac}^{depro}$  is the difference in the gas phase energy of the associated (protonated) and dissociated (deprotonated and hydrogen ion) system which can be obtained from quantum chemical calculations. The  $\Delta H_{vib}^{depro}$  is the change in the vibrational energy between the protonated and deprotonated states and this can be obtained from normal mode analysis. The  $H_{trans}(H^+) = \frac{3}{2}RT$  is the translation energy of a proton and  $\Delta(pV)$  is the energy change due to the volume change in the gas phase reaction which is estimated to be  $k_B T$  from the ideal gas approximation. The  $T[S(H^+)]$  is the entropic contribution of the gas phase free energy of proton and is set to  $7.8$  kcal/mol [62] as derived from the Sackur-Tetrode equation. The redox potential  $E_0^{redox}$  can also be computed by a similar approach.

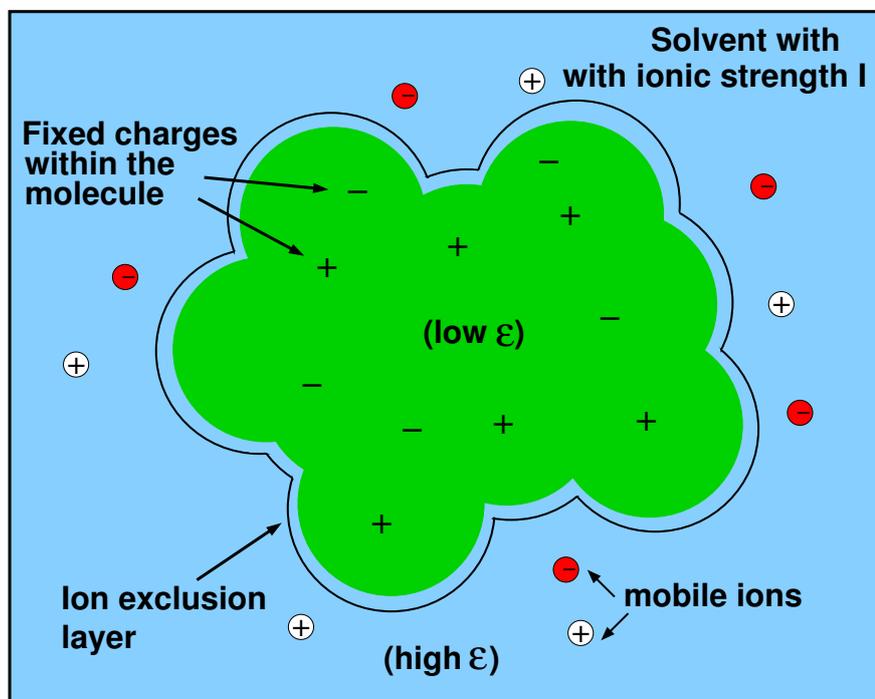
The redox potential  $E_0^{redox}$  can be calculated from Eq. (2.18)

$$E_0^{redox} = \frac{1}{F}(\Delta H_{vac}^{redox} + \Delta\Delta G_{solv}^{redox}) + \Delta SHE \quad (2.18)$$

where  $\Delta H_{vac}^{redox}$  is the difference in the gas phase energy between the oxidized and reduced states. The  $\Delta\Delta_{solv}^{redox}$  is the solvation energy difference between the oxidized and reduced states which can be calculated by solving the Poisson-Boltzmann equation.  $F$  is the Faraday constant (23.06 kcal/mol) and  $\Delta SHE$  is the standard potential of the hydrogen electrode (-4.43 V) [62].

## 2.2 CONTINUUM ELECTROSTATICS

Electrostatic effects are very important in many phenomena in biology and chemistry. In biology, electrostatics is a key component in determining the structure of proteins, nucleic acids and membranes. Electrostatics of proteins can be studied by different models. The quantum chemical calculations can give accurate results, but it is computationally expensive and time consuming method. An alternative approach involves continuum or macroscopic models, in which the solvent properties are described in terms of average values. By this model it is possible to describe solute molecules in atomic details and treating the solvent in terms of average properties [64]. Following approximations will be used in the continuum electrostatics model. The solvent water is assumed as a continuum liquid with high dielectric constant ( $\epsilon = 80$ ) (see Figure 2.2). The protein system will be assigned a low dielectric constant ( $\epsilon = 2 - 4$ ). The solvent which contains mobile ions are modeled as mobile charges. The protein system is represented as a low dielectric medium containing fixed point charges and the mobile ions are assumed to be infinitely small in size [62, 65, 66] (see Figure 2.2).



**Figure 2.2.** A molecule in a heterogeneous dielectric medium. A molecule of low dielectric with fixed charges in a high dielectric solvent with mobile charges (ion).

### 2.2.1 THE POISSON-BOLTZMANN EQUATION

The treatment of charges in vacuum is the simplest form of the classical electrostatics. Poisson equation can be applied to treat such electrostatic interactions.

The electrostatic potential  $\phi(\vec{r})$  at position  $(\vec{r})$ , arising from charge density  $\rho(\vec{r})$  in vacuum is given by,

$$\vec{\nabla}^2 \phi(\vec{r}) = -\frac{\rho(\vec{r})}{\varepsilon_0} \quad (2.19)$$

where  $\varepsilon_0$  is the dielectric constant in vacuum and  $\vec{r}$  is given in Cartesian coordinates. The  $\nabla$  is the Nabla-operator ( $\vec{\nabla} = (\partial/\partial x, \partial/\partial y, \partial/\partial z)$ ) and  $\vec{\nabla}^2 = \vec{\nabla} \cdot \vec{\nabla}$  is Laplace-operator. The Solutions of Eq. (2.19) yields the familiar coulombs law.

For inhomogeneous dielectric medium, the dielectricum varies with position  $(\vec{r})$ , and Poisson equation is modified to

$$\vec{\nabla} \cdot [\varepsilon(\vec{r}) \vec{\nabla} \phi(\vec{r})] = -\frac{\rho(\vec{r})}{\varepsilon_0} \quad (2.20)$$

The protein system is assumed to have mobile ions. The charge distribution of the mobile ions  $\rho_{ions}(\vec{r})$  can be described by Debye-Hückel theory. The theory describes the distribution of ions by combining electrostatics with statistical thermodynamics and uses the Boltzmann statistics based on the electrostatic interaction energy  $W_i$  between ion  $i$  and the solute [67]. The charge distribution resulting from the distribution of ions  $c_i(\vec{r})$  and other ions are

$$\rho_{ions}(\vec{r}) = \sum_{i=1}^K c_i(\vec{r}) \cdot Z_i \cdot e_0 = \sum_{i=1}^K c_i^{bulk} Z_i e_0 \cdot \exp\left(-\frac{W_i(\vec{r})}{RT}\right) \quad (2.21)$$

where  $c_i^{bulk}$  is the concentration of ion  $i$  in the bulk,  $K$  is the number of different type of ions,  $Z_i$  is the unitless formal charge of ion type  $i$ ,  $e_0$  is the elementary charge,  $R$  is the universal gas constant and  $T$  is the absolute temperature. The summation runs over all  $K$  types of ion  $i$ .

The interaction energy  $W_i$  is approximated by the potential energy  $E_{pot}$  of ion  $i$  in the electrostatic potential  $\phi(\vec{r})$  due to charge distributions of solute  $\rho_{solute}(\vec{r})$  and mobile ions  $\rho_{ions}(\vec{r})$ . This approximation  $W_i \approx E_{pot} = Z_i e \cdot \phi(\vec{r})$  yields

$$\rho_{ions}(\vec{r}) = \sum_{i=1}^K c_i^{bulk} Z_i e_0 \cdot \exp\left(-\frac{Z_i e_0 \cdot \phi(\vec{r})}{RT}\right) \quad (2.22)$$

With  $\rho(\vec{r}) = \rho_{solute}(\vec{r}) + \rho_{ions}(\vec{r})$ , Eq. (2.19) can be re-written to give the Poisson-Boltzmann equation (PBE)

$$\vec{\nabla} \cdot [\varepsilon(\vec{r}) \vec{\nabla} \phi(\vec{r})] = \frac{1}{\varepsilon_0} \left( \rho_{solute}(\vec{r}) + \sum_{i=1}^K c_i^{bulk} Z_i e_0 \cdot \exp\left(-\frac{Z_i e_0 \cdot \phi(\vec{r})}{RT}\right) \right) \quad (2.23)$$

Using the linear approximation  $\exp(-x) \approx 1 - x$  valid for small  $x \ll 1$ , for  $Z_i e_0 \cdot \phi(\vec{r})/RT \ll 1$

$$\sum_i^K c_i^{bulk} Z_i e_0 \cdot \exp\left(-\frac{Z_i e_0 \cdot \phi(\vec{r})}{RT}\right) = \sum_i^K c_i^{bulk} Z_i e_0 - \sum_i^K c_i^{bulk} Z_i e_0 \cdot \frac{Z_i e_0 \cdot \phi(\vec{r})}{RT} \quad (2.24)$$

As the total charge of the ions in solution is equal to zero, the first term on the right side of Eq. (2.24) is zero.

$$\vec{\nabla} \cdot [\varepsilon(\vec{r}) \vec{\nabla} \phi(\vec{r})] = \frac{1}{\varepsilon_0} \left( \rho_{solute}(\vec{r}) + \sum_i^K \left( \frac{c_i^{bulk} Z_i^2 e_0^2}{RT} \phi(\vec{r}) \right) \right) \quad (2.25)$$

The linear Poisson-Boltzmann equation (LPBE) can be rewritten by introducing the ionic strength and the Debye-Hückel inverse,  $\kappa$

$$I = \frac{1}{2} \sum_{i=1}^K c_i^{bulk} Z_i^2 \quad (2.26)$$

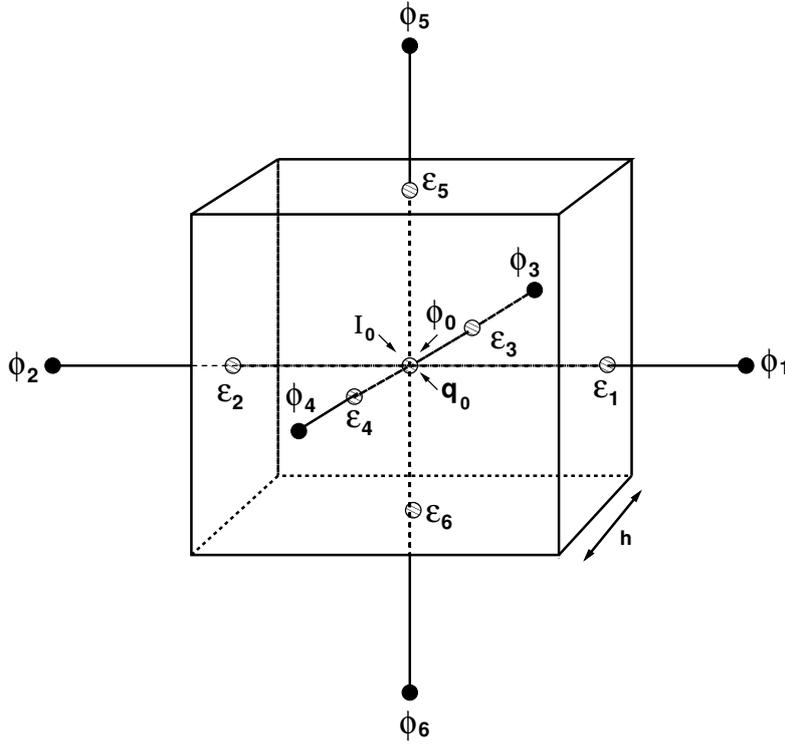
$$\kappa^2(\vec{r}) = \frac{8\pi e_0^2 I(\vec{r})}{RT} \quad (2.27)$$

$$\vec{\nabla} \cdot [\varepsilon(\vec{r}) \vec{\nabla} \phi(\vec{r})] = -4\pi \rho_{solute}(\vec{r}) + \kappa^2 \phi(\vec{r}) \quad (2.28)$$

Eq. (2.28) is solved numerically to yield  $\phi(\vec{r})$ .

The Poisson-Boltzmann equation is quite good for the mobile ions (monovalent and multivalent). The Linear Poisson-Boltzmann equation breakdowns for solvent with high concentration but these limitations will not affect results for the biological system since both ionic strengths and the electrostatic interactions are moderate in these systems.

NUMERICAL SOLUTION TO LINEAR POISSON-BOLTZMANN EQUATION. Analytical solutions are known only for simple geometries. For complex systems, numerical methods are applied to solve the Poisson-Boltzmann equation. A wide variety of problems have been studied using the finite difference method. However there are also other numerical methods such as boundary element method which tessellates the dielectric boundary between the interior and exterior of the molecule. A finite difference method is applied in the present work to solve the PBE. In finite difference method (see Figure 2.3), the protein (solute) is mapped into a cubic grid along with the solvent. The values of the electrostatic potential, charge density, grid constant  $h$ , dielectric constant and ionic strength are assigned to each grid point [68]. The atomic charges usually do not coincide with the grid points and so the charge is allocated to the eight surrounding grid points in such a way that the closer the charges to the grid point, the greater the proportion of its total charge that is allocated. The linear Poisson-Boltzmann



**Figure 2.3.** The cube used in the finite difference method for solving the Poisson-Boltzmann equation.

equation Eq. (2.28) is integrated over the volume  $V$  of the cubic grid:

$$\int \vec{\nabla} \cdot [\varepsilon(\vec{r}) \vec{\nabla} \phi(\vec{r})] d\vec{r} - \int \kappa_0^2(\vec{r}) \phi(\vec{r}) + 4\pi \int \rho_{solute}(\vec{r}) d\vec{r} = 0 \quad (2.29)$$

In the first step, the integral is transformed into a surface integral using Gauss's theorem: The volume integral is given by:

$$\int \vec{\nabla} \cdot [\varepsilon(\vec{r}) \vec{\nabla} \phi(\vec{r})] dA - h^3 \kappa_0^2 \phi_0 + 4\pi q_0 = 0 \quad (2.30)$$

The surface integral is calculated separately for all six sides of the cubic grid element.

$$\int \vec{\nabla} \cdot [\varepsilon(\vec{r}) \vec{\nabla} \phi(\vec{r})] dA = \sum_{i=1}^6 \frac{h^2 \varepsilon_i (\phi_i - \phi_0)}{h} \quad (2.31)$$

$$\sum_{i=1}^6 h \varepsilon_i (\phi_i - \phi_0) - h^3 \kappa_0^2 \phi_0 + 4\pi q_0 = 0 \quad (2.32)$$

the finite difference expression for  $\phi_0$ :

$$\phi_0 = \frac{(\sum_{i=1}^6 h \varepsilon_i \phi_i) + 4\pi q_0}{\sum_{i=1}^6 h \varepsilon_i + h^3 \kappa_0^2} \quad (2.33)$$

The electrostatic potential is iteratively calculated for each grid point by Eq. (2.33) until the convergence criterion is met.

## 2.2.2 CALCULATION OF PROTONATION AND REDOX EQUILIBRIA IN PROTEINS

One of the more frequent chemical reaction occurring in a protein-water solution is the uptake and release of protons by amino acids. The amino acids in a protein that can absorb or release protons are referred as titratable group. The Asp, Glu, Cys, Tyr, His, Lys, Arg, the C-terminal and N-terminal are basic titratable groups in proteins. Similarly the protein can contain redox active groups. The redox active groups also considered to be a titratable group which can undergo oxidation and reduction.

THE BEHAVIOR OF TITRATABLE SITES IN PROTEINS. The titration behavior of a single titratable group in aqueous solution will differ when the same titratable group is present in the protein [62, 69]. This section focuses on the behavior of protonatable groups though similar considerations apply to the behavior of redox groups.

The difference in the titration behavior in the protein is due to the following factors: change in dielectric environment, moving the neutral and charged form of the group from water to its position in the protein (Born energy) and the interaction with the neutral and charged form of the residue with the permanent dipoles of the protein (background interaction energy). The protonation probability of the respective group can alter the protonation state of the other titratable group (see Figure 2.4).

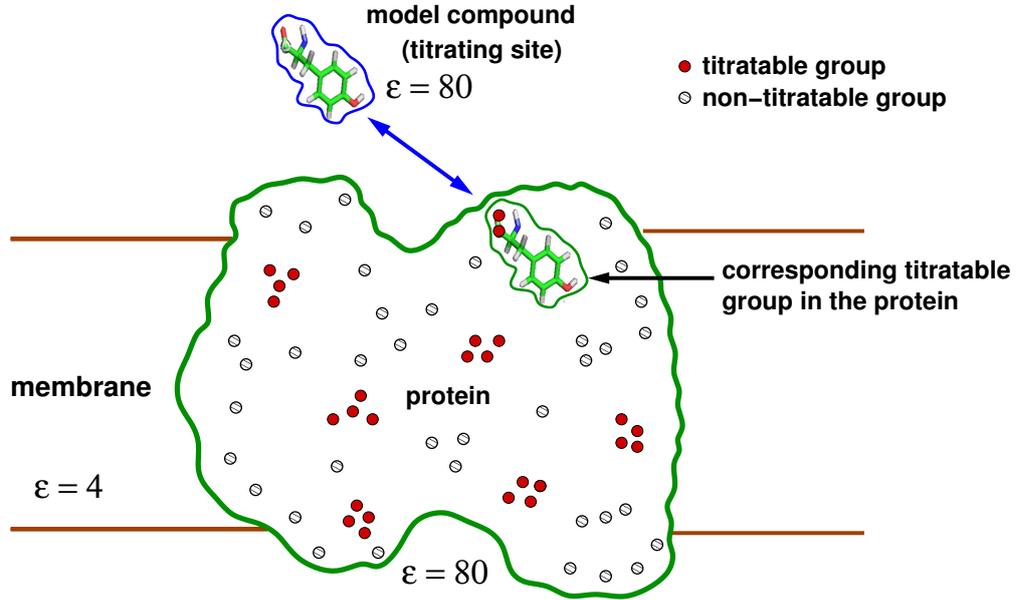
### ELECTROSTATIC MODELS

CALCULATION OF INTRINSIC  $pK_a$  VALUE. The  $pK_a$  value of a titratable group within a protein can considerably shifted compared to the  $pK_a$  value of the same titratable group in the aqueous solution [65]. The  $pK_a$  values of model compounds in aqueous solution are usually determined experimentally or the corresponding energies are calculated from the quantum chemical calculations. A hypothetical quantity, intrinsic  $pK_a$  introduced by Tanford and Kirkwood is used to describe the effect of background charges and the dielectric environment of a protein on the behavior of a titratable sites. The intrinsic  $pK_a$  value is the  $pK_a$  value that the particular titratable group would have, if all other titratable groups are in their reference protonation form i.e., the neutral form. The  $pK_a^{intr}$  of any titratable group can be obtained using a model  $pK_a$  and the energies for transferring the titratable group from aqueous solution to the protein environment [70, 71].

$$pK_{a,i}^{intr} = pK_{a,i}^{model} + \frac{1}{RT \ln 10} (\Delta\Delta G_i^{Born} + \Delta\Delta G_i^{back}) \quad (2.34)$$

$pK_a^{model}$ , is the  $pK_a$  value of titratable group in the aqueous solution and these  $pK_a$  values are known for all standard titratable groups in proteins.

Transferring the titratable group  $i$  from aqueous solution to protein, in which all other titratable groups are in their uncharged protonation form causes an energy shift. This energy shift is due to the contribution of two terms: Born energy term ( $\Delta\Delta G_i^{Born}$ ) and background energy term ( $\Delta\Delta G_i^{back}$ ).



**Figure 2.4. Titratable groups and model compound.** The protein is embedded in a membrane and is surrounded by solvent. The protein and the membrane are rigid low dielectric with permanent background charges. The model compound consists of a single titratable group in solution. The  $pK_a$  differences between the protein and the model compound are due to the following energies: Born solvation energy and background energy.

In a dielectric medium, any charge distribution induces a reaction field by polarizing the environment. The reaction field is always opposed to the original field and can be assumed to be proportional to the inducing charge  $q_0$  with  $\phi_{rf} = c \cdot q_0$  [67].

$$G_{Born} = \int_0^{q_0} \phi_{rf} dq = \int_0^{q_0} c \cdot q_0 dq = \frac{1}{2} c \cdot q_0^2 = \frac{1}{2} \phi_{rf} q_0 \quad (2.35)$$

The Born energy term  $\Delta\Delta G_i^{Born}$  is due to the interaction of the partial charges  $Q_k$  of the titratable group  $i$  with its reaction field:

$$\Delta\Delta G_i^{Born} = \frac{1}{2} \sum_{k=1}^{N_i} Q_{k,i}^{(p)} [\phi_{protein}(\vec{r}_k)] - \frac{1}{2} \sum_{k=1}^{N_i} Q_{k,i} [\phi_{model}(\vec{r}_k)] \quad (2.36)$$

where  $\phi_{protein}$  and  $\phi_{model}$  are potential at position ( $\vec{r}$ ) of partial charge  $Q$  of the atom  $k$  in  $N_i$  atoms of the titratable group  $i$ . The  $\phi_{protein}$  and  $\phi_{model}$  are calculated by solving the PBE by finite difference method in the protein and model compounds respectively and the superscript ( $p$ ) and ( $u$ ) indicate the protonated and unprotonated charge states. The background energy term  $\Delta\Delta G_i^{back}$  is due to the interaction of charges  $Q_k$  of titratable group  $i$  with the background charges of non-titratable groups and charges of the uncharged forms of all other titratable

groups in the protein.

$$\Delta\Delta G_i^{back} = \sum_{k=1}^{N_p} q_k [\phi_{protein}^{(p)}(\vec{r}_k) - \phi_{protein}^{(u)}(\vec{r}_k)] - \sum_{k=1}^{N_m} q_k [\phi_{model}^{(p)}(\vec{r}_k) - \phi_{model}^{(u)}(\vec{r}_k)] \quad (2.37)$$

where the  $q_k$  are the non-titrating group partial charges located at  $(\vec{r}_k)$ . The first summation in Eq. (2.37) runs over the  $N_p$  charges of protein that belong to atoms of non-titrating groups. The second summation in Eq. (2.37) runs over the  $N_m$  charges of atoms of the model.

Electrostatic potentials  $\phi_{protein}(\vec{r}_k)$  and  $\phi_{model}(\vec{r}_k)$  at position  $\vec{r}_k$  of permanent charge  $q_n$  are derived from solutions of the PBE.

The additivity of the solutions of the PBE makes it possible to separate the protonation energy of a titratable group in protein into several independent contributions.

INTERACTION ENERGIES. In proteins with multiple titration sites, the protonation states of each groups can influence each other. In addition to the  $pK_{a,i}^{intr}$  of each site in the protein, the interaction between titrating groups are required. The interaction energy  $W_{ij}$  between two titratable groups  $i$  and  $j$  is:

$$W_{ij} = \sum_{k=1}^{N_i} [Q_{k,i}^{(p)} - Q_{k,i}^{(u)}] [\phi_{protein,j}^{(p)}(\vec{r}_k) - \phi_{protein,j}^{(u)}(\vec{r}_k)] \quad (2.38)$$

where  $\phi_{protein,j}$  is the potential field due to the charges on site  $j$  and  $Q_{k,i}$  is the charge on the  $k$ th atom of site  $i$ .

COMPUTATION OF PROTONATION STATE ENERGIES. The protonation state of a protein with  $N$  titratable sites can be described by a  $N$ -dimensional protonation state vector  $\vec{x} = (x_1, x_2, \dots, x_N)$  where  $x_i$  is 1 or 0 depending on whether the site  $i$  is protonated or deprotonated respectively [62].

The energy of any protonation state can be calculated if the  $pK_a^{intr}$  and  $W_{ij}$  are known for all sites and pairs of site, respectively.

$$G^{(n)} = \sum_{i=1}^N \left( (x_i^{(n)} - x_i^{(0)}) RT \ln 10 \cdot (pH - pK_{a,i}^{intr}) \right) + \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \left( W_{ij} (x_i^{(n)} - x_i^{(0)}) (x_j^{(n)} - x_j^{(0)}) \right) \quad (2.39)$$

where  $x_i^{(n)}$  is 1 or 0 depending on whether group  $i$  is protonated or deprotonated in state  $n$ ,  $x_i^{(0)}$  is 1 or 0 depending on whether group  $i$  is protonated or deprotonated in the reference protonation state, and the same is applicable for group  $j$ . For redox active groups, an additional

term is needed.

$$\begin{aligned}
G^{(n)} = & \sum_{i=1}^N \left( (x_i^{(n)} - x_i^{(0)}) RT \ln 10 \cdot (pH - pK_{a,i}^{intr}) \right) \\
& - \sum_{i=1}^K \left( (x_i^{(n)} - x_i^{(0)}) F(E - E_v^{0,intr}) \right) \\
& + \frac{1}{2} \sum_{i=1}^{N+K} \sum_{j=1}^{N+K} \left( W_{ij} (x_i^{(n)} - x_i^{(0)}) (x_j^{(n)} - x_j^{(0)}) \right)
\end{aligned} \tag{2.40}$$

where,  $E_v^{0,intr}$  is the intrinsic standard redox potential of the redox-active group  $v$ , for redox-active groups  $x_v^n$  is either 1 or 0 depending whether group  $v$  is oxidised or reduced. The sum in the last term runs over all titratable and redox-active groups.

COMPUTATION OF PROTONATION PROBABILITIES BY QUANTUM CHEMICAL METHODS. The protonation and reduction probabilities are calculated from the grand canonical partition function. The energy of each possible state can be calculated from the microscopic  $pK_a$  and redox potentials. The macroscopic or apparent  $pK_a$  values can be obtained from the roots of the grand canonical partition function.

MULTIPLE SITES TITRATION. The titration behavior of individual residues is often complicated due to the interaction between the titratable residues in a protein.

A protein that has  $N$  proton binding sites and can exist in  $M$  conformation can adopt  $2^N M$  different microstates. Each microstate is characterized by a vector  $x$  which specifies the protonation state ( $x_i = 1$  if group  $i$  protonated, 0 if it is deprotonated), and a number  $m$  which specifies the conformation. The partition function of such a protein in terms of microstate standard free energy,  $G_{x,m}^o$  is then given by

$$Z = \sum_m^M \sum_x^{2^N} e^{-\beta G_{x,m}^o} e^{\beta n_x \mu_{H^+}} \tag{2.41}$$

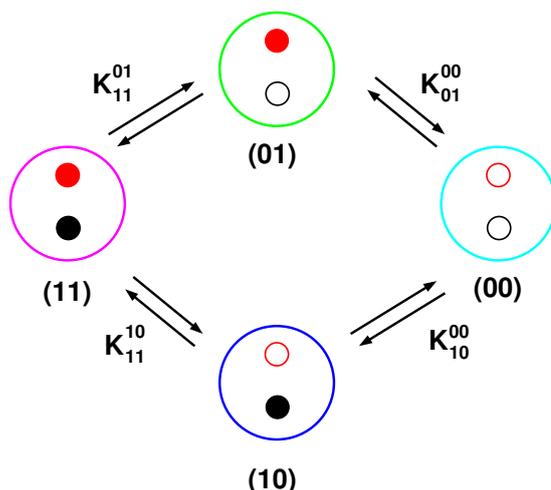
where  $n_x = \sum_{i=1}^N x_i$  is the number of protons bound to the molecule in state  $x$  and  $\beta = 1/RT$ . The partition function is a polynomial of the variable  $e^{\beta \mu_{H^+}}$  and Eq. (2.41) is also referred as binding polynomial [72]. The sum of the Boltzmann factors  $e^{\beta G_{x,m}^o}$  of the state  $x$  each having  $n_x$  protons bound in all  $M$  possible conformations give the  $n_x$ th coefficient of the polynomial:

$$\begin{aligned}
Z = & \sum_m^M \sum_x^{2^N} \delta(1) e^{-\beta G_{x,m}^o} e^{\beta n_x \mu_{H^+}} + \sum_m^M \sum_x^{2^N} \delta(2) e^{-\beta G_{x,m}^o} e^{\beta n_x \mu_{H^+}} \\
& + \dots + \sum_m^M \sum_x^{2^N} \delta(N) e^{-\beta G_{x,m}^o} e^{\beta n_x \mu_{H^+}}
\end{aligned} \tag{2.42}$$

with  $\delta(k) = 1$  if the state  $\{x, m\}$  has  $k$  protons bound and otherwise 0. The probability that a particular site  $i$  is protonated in the molecule i.e., the titration curve of an individual site is given by

$$\langle x_i \rangle = \frac{1}{Z} \sum_m \sum_x x_i e^{-\beta G_{x,m}^0} e^{\beta n_x \mu_{H^+}} \quad (2.43)$$

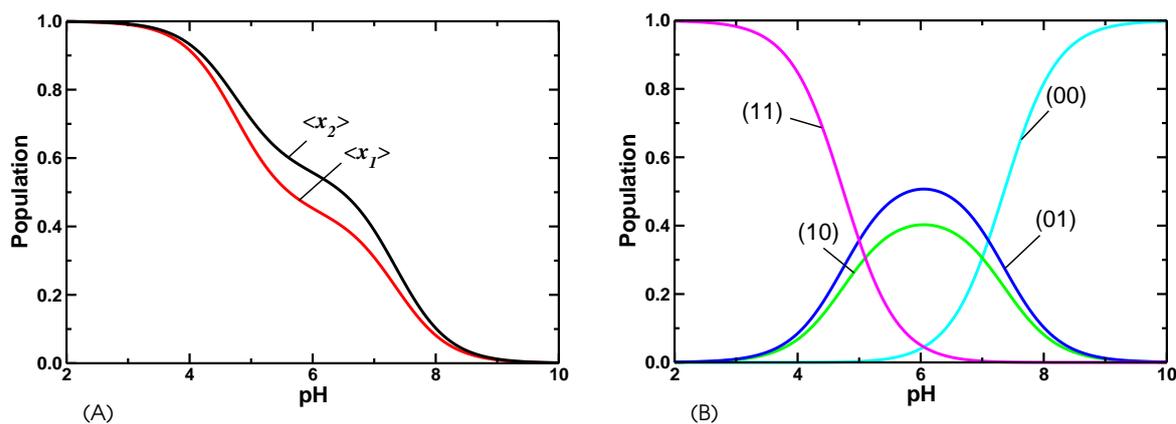
The microscopic equilibrium constants can be assigned to all equilibria between different microstates of the protein. These microscopic equilibrium constants can be defined in analogous to the equilibrium constant of a monoprotic acid. For example, a simple system with only two titratable sites has the following microscopic states (see Figure 2.5 and 2.6): both protonated (11), only one site deprotonated (01 and 10) and both the sites deprotonated (00) [73, 74]. The following microscopic equilibrium constants  $K_r^p$  where the  $r$  denotes the reactant state vector and  $p$  the product state vector, can thus defined as



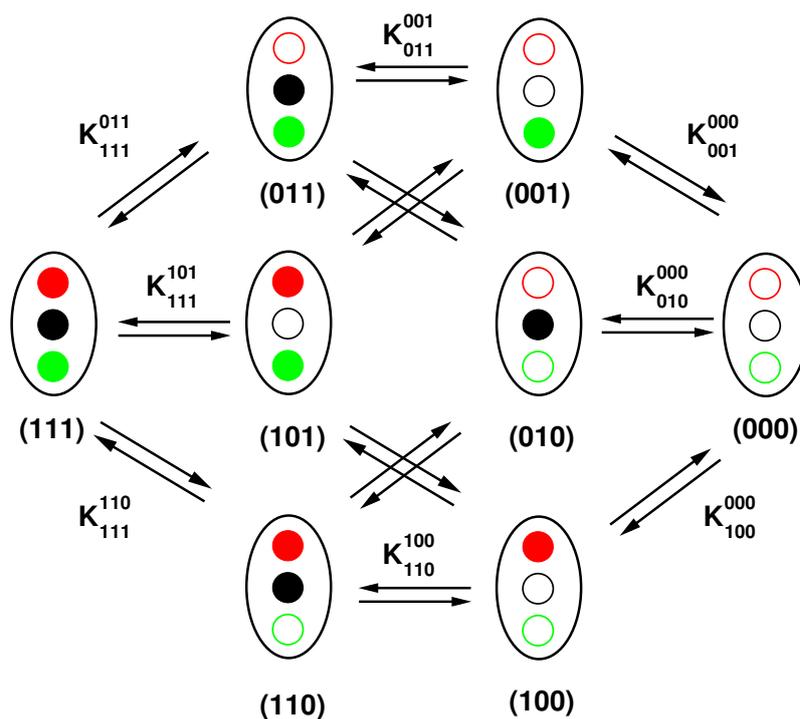
**Figure 2.5. Titration behavior of a diprotic acid.** Titration behavior of a diprotic acid with an interaction energy  $W = 2.0$  pK units between two sites, and microscopic  $pK_a$  values  $pK_{10}^{00} = 7.0$  and  $pK_{01}^{00} = 7.1$ . The four different protonation states of the system and their interconversion. Filled circles represent a protonated site, open circles are deprotonated sites. The corresponding protonation state vectors are given below the states.

$$\begin{aligned} K_{11}^{01} &= \frac{[(01)][H^+]}{[(11)]}; K_{11}^{10} = \frac{[(10)][H^+]}{[(11)]} \\ K_{01}^{00} &= \frac{[(00)][H^+]}{[(01)]}; K_{10}^{00} = \frac{[(00)][H^+]}{[(10)]} \end{aligned} \quad (2.44)$$

The microscopic  $pK_r^p$  values are defined as the negative decimal logarithm of the microscopic equilibrium constants  $K_r^p$  ( $pK_r^p = -\log K_r^p$ ). The free energy difference of the microstates at standard conditions



**Figure 2.6. Titration behavior of a diprotic acid.** (A). Probability of finding the system in either one of the four different microscopic states. Color-coding of the different microstates correspond to Figure 2.5. The curves are marked by corresponding protonation state vector. (B). The protonation probabilities of the two real sites are shown.



**Figure 2.7. The system with three proton binding sites and their interconversion.** The system with three proton binding sites with all possible microscopic states are shown. The three proton binding sites leads to eight microscopic states and twelve proton equilibrium reactions. Filled circles represent a protonated site and open circles are deprotonated sites. The  $C_{UB}$  center studied in this work is a good example for the three proton binding sites.

$$pK_r^p = \frac{1}{RT \ln 10} (G_p^o - G_r^o) \quad (2.45)$$

The macroscopic  $p\bar{K}_k$  values characterize the equilibria between two macroscopic protonation states of the protein i.e., between the two macrostates with  $k$  and  $k - 1$  protons bound. The  $k$ th macroscopic  $p\bar{K}_k$  value is given by

$$p\bar{K}_k = -\log \left( \frac{\sum_m^M \sum_x^{2^N} \delta(k-1) e^{-\beta G_{x,m}^o}}{\sum_m^M \sum_x^{2^N} \delta(k) e^{-\beta G_{x,m}^o}} \right) \quad (2.46)$$

A macroscopic  $p\bar{K}_k$  value is the thermodynamic average over all microscopic equilibria involved in the release of the  $k$ th proton. The product over the first to the  $k$ th macroscopic  $\bar{K}_k$  values equals to the  $k$ th polynomial coefficient  $\sum_m^M \sum_x^{2^N} \delta(k) e^{-\beta G_{x,m}^o}$  of the partition function in Eq. (2.42). The macroscopic  $p\bar{K}_k$  values can consequently also be used to formulate the partition function. Moreover, the total titration curve of a polyprotic acid is often interpreted in terms of macroscopic  $p\bar{K}_k$  values. The  $pK_a$  values for three proton binding sites are also calculated in similar way. The system with three proton binding sites are shown in Figure 2.7 [75].

#### CALCULATION OF TITRATION CURVES

The protonation state of a protein with  $N$  titratable group is  $2^N$ . The titration curve of a single site in a protein is given by the thermodynamic average over all possible protonation states at each pH-value

$$\langle x_i \rangle = \frac{\sum_{n=0}^{2^N} x_i^{(n)} \exp \left( -\frac{G^{(n)}(\text{pH})}{RT} \right)}{\sum_{n=0}^{2^N} \exp \left( -\frac{G^{(n)}(\text{pH})}{RT} \right)} \quad (2.47)$$

To obtain the protonation state energy  $G^n$ , the Eq. (2.47) should be evaluated for  $2^N$  times. Since cytochrome *c* oxidase and other model proteins are huge, the thermodynamic average  $\langle x_i \rangle$  grows exponentially with the number of sites. The average behavior of such system can be obtained by Metropolis Monte Carlo approach which is used to determine the protonation of many interacting sites as a function of pH.

**THE METROPOLIS MONTE CARLO APPROACH.** The Metropolis Monte Carlo method is used to obtain the titration curves of amino acids within the protein [76]. Initially a random protonation state  $n$  is chosen, the energy of the corresponding protonation state is calculated. Each Monte Carlo (MC) step chooses the protonation state randomly and the energy of the current protonation state is calculated and the change in energy is computed, i.e., the energy difference between current protonation state and the previous protonation state  $\Delta G = G^{old} - G^{current}$ .

The current state is accepted with a probability  $p$  according to Metropolis criterion

$$p = \begin{cases} 1 & \text{if } \Delta G \leq 0, \\ \exp\left(\frac{-\Delta G}{RT}\right) & \text{if } \Delta G > 0 \end{cases} \quad (2.48)$$

The above process is repeated for each MC step. The collection of protonation states start after the equilibration step. The steps involved in MC are shown in Figure 2.8. The Metropolis Monte Carlo approach also effectively sample the protonation states of the strongly coupled sites. The strongly coupled groups have the following states: the singly protonated state (1,0) and (0,1) the low energy states and state (0,0) and (1,1) high energy states. Therefore two transitions are needed to change from (1,0) to (0,1) to avoid the high- energy intermediate states (0,0) and (1,1). To solve this sampling problem, double and triple moves are introduced, which simultaneously change the protonation states of two or three sites in MC step.

**CORRELATION CURVES.** The coupling between the protonation form of titratable group  $i$  and group  $j$  is given by:

$$c = \langle x_{i,j} \rangle - \langle x_i \rangle \cdot \langle x_j \rangle \quad (2.49)$$

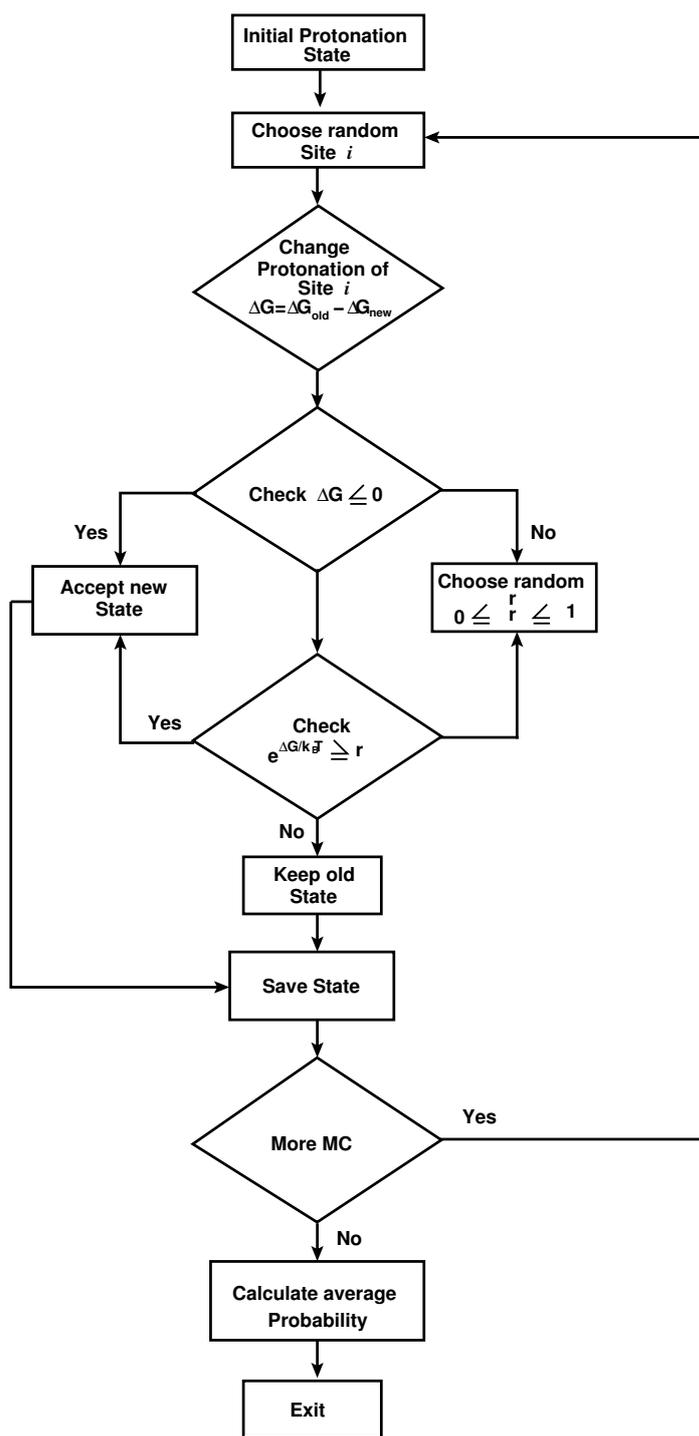
where the  $\langle x_i \rangle$  and  $\langle x_j \rangle$  are the probability of the titratable group  $i$  and group  $j$  which are protonated respectively and  $\langle x_{i,j} \rangle$  is the probability for both are protonated in the same time. The correlation between two sites can be obtained from above Eq. (2.49).

#### TANFORD-ROXBY $pK_a$ VALUES

The Tanford-Roxby approximation assumes that the average protonation of a titratable residue depends on the average charges of all other titratable groups. The average  $pK_a$  value of the residue  $i$  inside the protein [77, 78]  $\langle pK_{a,i}^{intr} \rangle_{prot}$  is obtained from Eq (2.50).

$$pK_{a,i} = pK_{a,i}^{intr} + \sum_{j=1}^N (\langle x_j \rangle - x_j^{(0)}) W_{ij} \quad (2.50)$$

where the intrinsic  $pK_a$  value ( $pK_{a,i}^{intr}$ ) is the  $pK_a$  value that the particular titratable group would have, if all other titratable groups are in their reference form. This term includes the solvation energy and the interaction with non-titrating residues and the protein backbone. The term  $W_{ij}$  represents the interaction energy between the titratable groups  $i$  and  $j$  in their charged form;  $\langle x_j \rangle$  represents the protonation probability of the group  $j$  which is obtained by a thermodynamic average over all possible protonation states and  $x_j^{(0)}$  is the reference protonation form of site  $j$ . The  $\langle x_j \rangle$  are obtained from MC calculations. These average  $pK_a$  values do not represent an equilibrium situation, but they are a good approximation of the real  $pK_a$  value of the site.



**Figure 2.8. The Metropolis Monte Carlo (MC) approach.** Initial random protonation state  $i$  is chosen and the protonation is changed. The energy difference between the new state and old state is calculated. The new state is accepted if the energy is smaller than zero. If the energy is greater than zero, the Boltzmann factor of the change in energy is checked. If the change in energy is larger than a number randomly chosen between 0 and 1 then the state is accepted. The steps are repeated until convergence of the sampled properties is achieved.

## 2.3 QUANTUM CHEMISTRY

The application of quantum-mechanical principles to chemical problems has gained momentum in past few decades. The quantum mechanical methods can be of great use in the simulation of proteins. They can be used to obtain parameters for molecular mechanics type potential energy functions. In the present work, the quantum chemical (QC) methods are used to calculate the free energies of the reactions, vibrational energies, solvation energies and partial atomic charges. This chapter described some of the basic principles, methods of quantum chemistry that lead to understand the electronic structure behavior.

### SCHRÖDINGER EQUATION

The principles of density functional theory are conveniently expounded by making reference to conventional wave function theory. The time independent, non-relativistic Schrödinger equation for a many electron system can be written as:

$$H\Psi = E\Psi \quad (2.51)$$

$H$  is the Hamiltonian operator for a system of nuclei and electrons described by position vectors  $\vec{R}_A$  and  $\vec{r}_i$ , respectively.  $\Psi$  is the wave function of the many particle system and  $E$  is the energy of the eigenfunction. In SI units the Schrödinger equation of the hydrogen atom can be written as:

$$-\frac{\hbar^2}{2m}\nabla_i^2\Psi_i(\vec{r}) - \frac{e^2}{4\pi\epsilon_0 r}\Psi_i(\vec{r}) = E_i\Psi_i(\vec{r}) \quad (2.52)$$

In Eq. (2.52), first term in the left hand side (LHS) is the kinetic energy of the electron system and the second term is the potential energy term and  $E_i$  is the energy of the system, where  $e$  and  $m$  are the electron charge and mass respectively,  $\hbar$  is Plank's constant divided by  $2\pi$ ,  $r$  is the distance between electron and proton and  $\epsilon_0$  is the permittivity of free space. In the SI system of units the above equation is rather unwieldy to deal with because of the tiny numbers associated with quantities as  $m$ ,  $e^2$  and  $\hbar^2$ . The inconvenience can be removed by rewriting the equation in terms of 'natural' atomic units.

The Schrödinger equation can be expressed in atomic units where  $m_e = |e| = \hbar = 4\pi\epsilon_0 = 1$  and Eq. (2.52) becomes Eq. (2.53)

$$\left(-\frac{1}{2}\nabla_i^2 - \frac{1}{r}\right)\Psi_i(\vec{r}) = E_i\Psi_i(\vec{r}) \quad (2.53)$$

The Hamiltonian for  $N$  electrons and  $M$  nuclei can be expressed as

$$H = - \sum_{i=1}^N -\frac{1}{2} \nabla_i^2 - \sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} + \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}} \quad (2.54)$$

In Eq. (2.54) the distance between the  $i$ -th electron and the  $A$ -th nucleus is  $\vec{r}_{iA}=|\vec{r}_i-\vec{R}_{iA}|$ , the distance between the  $i$ -th and  $j$ -th electron is  $\vec{r}_{ij}=|\vec{r}_i-\vec{r}_j|$  and the distance between the  $A$ -th and  $B$ -th nucleus is  $\vec{R}_{ij}=|\vec{R}_i-\vec{R}_j|$ .  $M_A$  is the ratio of the mass of nucleus  $A$  to the mass of an electron and  $Z_A$  is the atomic number of nucleus  $A$ . The Laplacian operators  $\nabla_i^2$  and  $\nabla_A^2$  involve differentiation with respect to the coordinates of the  $i$ -th electron and  $A$ -th nucleus, respectively. The first term in the above equation is the operator for the kinetic energy of the electrons; the second is the operator for the kinetic energy of the nucleus; the third term represents the Coulomb attraction between electrons and nuclei; the fourth and fifth term represents the repulsion between electrons and between nuclei, respectively.

The *Born-Oppenheimer approximation* [79, 80] considers electrons as moving particles in the field of fixed nuclei. Due to significant differences in mass between electrons ( $m_{el} \sim 10^{-31}$  kg) and the nucleus ( $m_{nuc} \sim 10^{-27}$  kg), the electronic degrees of freedom can be considered to respond instantaneously to any change in the nuclear configuration i.e., their wave functions always correspond to a stationary state. Hence, the interaction between different stationary states. Application of the Born-Oppenheimer approximation enables one to formulate the electronic Hamiltonian. The time-independent electronic Schrödinger equation describing the motion of electrons in the electrostatic field of the stationary nuclei comprises only term one, three and four of Eq. (2.54), where the electronic Hamiltonian and the electronic wave functions depend only parametrically on the coordinates of the nuclei.

### 2.3.1 DENSITY FUNCTIONAL THEORY

The density functional theory (DFT) is a remarkable theory that allows one to replace the complicated  $N$ -electron wave function by a much simpler electron density  $\rho(\vec{r})$ . The basis for DFT is the proof by Hohenberg and Kohn [81] that the ground-state electronic energy is determined completely by the electron density  $\rho(\vec{r})$  [79, 80, 82, 83]. The Hohenberg-Kohn says that the molecular energy, the wave function and all other molecular properties can be computed from the ground state electron probability density. Even with the proof that every specific electron density gives a specific energy (for the ground state), the connection between the energy and the functional has not been discovered. The hypothetical functional can be divided into three parts: the kinetic energy  $T[\rho(\vec{r})]$ , the core-electron attraction  $E_{ne}[\rho(\vec{r})]$ , and the electron-electron repulsion  $E_{ee}[\rho(\vec{r})]$ . The electron-electron repulsion is usually divided into an exchange part  $K[\rho(\vec{r})]$  and a Coulomb part  $J[\rho(\vec{r})]$ . The exact kinetic energy ( $T$ ) for a ground state is given by

$$T = \sum_i^{\infty} n_i \langle \Psi_i | -\frac{1}{2} \nabla_i^2 | \Psi_i \rangle \quad (2.55)$$

where  $\Psi_i$  are natural spin orbitals and  $n_i$  their occupation number,  $0 \leq n_i \leq 1$ . For an interacting system, Eq. (2.55) contains an infinite number of terms, making an exact solution unattainable. Kohn and Sham [81] presented a formalism for treating the kinetic part of the energy.

In Hartree's model [79, 80], the electrons move in an effective potential  $v_H(\vec{r})$  created by the atomic cores and a mean field created by the other electrons is given as:

$$v_H(\vec{r}) = - \sum_a^{\text{nuclei}} \frac{Z_a}{|\vec{R}_a - \vec{r}|} + \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}' \quad (2.56)$$

where  $a$  runs over all the nuclei,  $Z_a$  is the nuclear charge on atom  $a$ ,  $\vec{R}_a$  denotes the position of the nuclei and  $\rho(\vec{r})$  is the electron density.

In this approximation, a one-particle Schrödinger equation can be defined as:

$$\left[ -\frac{1}{2} \nabla_i^2 + v_H(\vec{r}) \right] \Psi_i = E_i \Psi_i \quad (2.57)$$

where  $E_i$  is the energy of the system.

The mean density is given by

$$\rho(\vec{r}) = \sum_i^N |\Psi_i(\vec{r})|^2 \quad (2.58)$$

For the ground state, a summation over the  $N$  spin orbitals with lowest eigenvalues is performed.

Eq. (2.56-2.58) are solved self-consistently. From a trial guess of the electron density, a potential is obtained by Eq. (2.56). This potential is used in Eq. (2.57). The one-electron orbitals obtained are put in to Eq. (2.58). The procedure is repeated until a set of converged one-electron orbitals are obtained. For a system of independent non-interacting electrons, the common one-body potential,  $v_{KS}$  is assumed.

$$\left[ -\frac{1}{2} \nabla_i^2 + v_{KS}(\vec{r}) \right] \Psi_i = E_i \Psi_i \quad (2.59)$$

The local one-body potential  $v_{KS}$  gives, by definition, a non-interacting electron density that has the same form as Eq. (2.57).

On comparing Eq. (2.55) and Eq. (2.59), the expression for the kinetic energy of the non-interacting electrons ( $T_s[\rho]$ ), becomes

$$T_s[\rho] = \sum_{i=1}^N \langle \Psi_i | -\frac{1}{2} \nabla_i^2 | \Psi_i \rangle \quad (2.60)$$

where the sum  $i$  again runs over all orbitals. The lower index  $s$  in  $T_s[\rho]$ , denotes the single-electron equations. As the electrons do interact in reality, the expression is not exact, and the following inequality holds where  $T[\rho]$  is the exact kinetic energy

$$T_s[\rho] \leq T[\rho] \quad (2.61)$$

The remaining part of the kinetic energy defines the correlation contribution:

$$T_c[\rho] = T[\rho] - T_s[\rho] \geq 0 \quad (2.62)$$

where  $T_c$  is included in a exchange/correlation term  $E_{xc}$ . The Kohn-Sham equations can be solved analogously to the Hartree equations, with the difference that the potential in Eq. (2.57),  $v_H$ , is replaced by  $v_{eff}$ :

$$v_{eff}(\vec{r}) = v(\vec{r}) + \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}' + v_{xc}(\vec{r}) \quad (2.63)$$

where  $v_{xc}(\vec{r})$  denotes the local exchange/correlation potential and  $v(\vec{r})$  corresponds to the external potential.

Within the KS formalism, the energy of the ground state can be obtained from

$$E^{DFT}[\rho] = \sum_i E_i + E_{xc}[\rho] - \int v_{xc}(\vec{r})\rho(\vec{r})dv - \frac{1}{2} \int \frac{\rho(\vec{r})\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} \quad (2.64)$$

or more generally, divided into its components:

$$E^{DFT}[\rho(\vec{r})] = T_s[\rho(\vec{r})] + E_{ne}[\rho(\vec{r})] + J[\rho(\vec{r})] + E_{xc}[\rho(\vec{r})] \quad (2.65)$$

An exact energy expression has been obtained from Eq. (2.65). The first term is the kinetic energy for the non-interacting electrons and the second term is the core-electron attraction and the third term represents the Coulombic repulsion functional, the electrostatic electron-electron repulsion and the final term is the exchange/correlation term. Except the last term exchange/correlation energy, all other terms are solved exactly. The new challenge was to find a solution to  $E_{xc}$ .

## DIFFERENT DFT MODELS

Most DFT models differ in their expression for the exchange/correlation energy  $E_{xc}$ . It is always split into separate exchange and correlation terms,  $E_x$  and  $E_c$ . The exchange energy is “by definition” given as a sum of contributions from  $\alpha$  and  $\beta$  spin densities, as exchange energy only involves electrons of the same spin. The kinetic energy, the nuclear-electron attraction and Coulomb terms are trivially separable. The correlation energy contains contributions from the interactions between all electrons:

$$E^x(\rho) = E_x^\alpha(\rho_\alpha) + E_x^\beta(\rho_\beta) \quad (2.66)$$

$$E^c(\rho) = E_c^{\alpha\alpha}(\rho_\alpha) + E_c^{\beta\beta}(\rho_\beta) + E_c^{\alpha\beta}(\rho_\alpha, \rho_\beta) \quad (2.67)$$

Where  $\rho_\alpha$  and  $\rho_\beta$  are densities of the  $\alpha$  and  $\beta$  spin. The total density is the sum of the  $\alpha$  and  $\beta$  contributions,  $\rho = \rho_\alpha + \rho_\beta$  the largest contribution to  $E_{xc}$  comes from the exchange part  $E_x$ .

## THE LOCAL DENSITY APPROXIMATION

In the *Local Density Approximation* (LDA) it is assumed that the density can be treated locally as a uniform electron gas or the electron density is assumed to be slowly varying in space

$$E_{xc}^{LDA}[\rho] = \int \rho(\vec{r}) \varepsilon_{xc}^{unif}(\rho) \quad (2.68)$$

$\varepsilon_{xc}^{unif}$  gives the exchange/correlation energy per electron in a uniform electron gas. The analytical expression for the exchange energy can be obtained from the Dirac formula

$$E_{xc}^{LDA}[\rho] = -C_x \int \rho^{4/3}(\vec{r}) d\vec{r} \quad (2.69)$$

where  $C_x$  is given as:

$$C_x = \frac{3}{4} \left( \frac{3}{\pi} \right)^{1/3}$$

In the more general case, where  $\alpha$  and  $\beta$  densities are not equal, LDA has been replaced by the *Local Spin Density Approximation* (LSDA).

$$E_{xc}^{LSDA}[\rho] = -2^{1/3} C_x \int \left( \rho_\alpha^{4/3}(\vec{r}) + \rho_\beta^{4/3}(\vec{r}) \right) d\vec{r} \quad (2.70)$$

For the LDA correlation energy, no exact solution is known but it has been determined by Monte Carlo methods for a number of different densities. In order to use these results in DFT calculations, a suitable analytic interpolation formula is desirable. Vosko, Wilk and Nusari [84] fitted various data to obtain reliable analytic expressions for the LDA correlation energy and in general considered to be a very accurate fit. This exchange functional is used in the calculations.

#### THE GENERALIZED GRADIENT APPROXIMATION

The *Generalized Gradient Approximation* (GGA) [79, 80] is the improvement over the LSDA approach where the electron distribution is considered as a non-uniform electron gas. The idea behind this is not only considering the electron density, but also the gradient of the density, that is the change in the density at a given point.

The term GGA was probably first introduced in connection with the PW86 functional, developed by Perdew and Wang in 1986 [85]. Perdew and Wang proposed modifying the LSDA exchange expression to that shown in Eq. (2.71)

$$E_x^{PW86} = \varepsilon_x^{LDA} (1 + ax^2 + bx^4 + cx^6)^{1/15} \quad (2.71)$$

$$x = \frac{|\nabla\rho|}{\rho^{4/3}} \quad (2.72)$$

where  $x$  is a dimensionless gradient variable, and  $a$ ,  $b$  and  $c$  being suitable constant where the summation over equivalent expression for  $\alpha$  and  $\beta$  densities is implicitly assumed.

A general GGA functional has the form:

$$\varepsilon_x^{GGA} = \varepsilon_x^{GGA}(\rho_\alpha, \rho_\beta, \nabla\rho_\alpha, \nabla\rho_\beta) \quad (2.73)$$

Several GGA functionals, for both exchange and correlation, have been proposed in the literature [85–90].

Becke [87] proposed a widely used correlation (B or B88) to the LSDA exchange energy:

$$\varepsilon_x^{B88} = \varepsilon_x^{LDA} + \Delta\varepsilon_x^{B88} \quad (2.74)$$

$$\Delta\varepsilon_x^{B88} = -\beta\rho^{1/3} \frac{x^2}{1 + 6\beta x \sinh^{-1} x} \quad (2.75)$$

The  $\beta$  parameter is determined by fitting to known atomic data and  $x$  is defined in Eq. (2.72).

Perdew and Wang have proposed an exchange functional similar to B88 to be used in connection with the PW91 correlation functional given below.

$$\varepsilon_x^{PW91} = \varepsilon_x^{LDA} \left( \frac{1 + xa_1 \sinh^{-1}(xa_2) + (a_3 + a_4 e^{-bx^2})x^2}{1 + xa_1 \sinh^{-1}(xa_2) + a_5 x^2} \right) \quad (2.76)$$

where  $a_i$  and  $b$  are suitable constants and  $x$  is defined in Eq. (2.72). where  $x$  is given in Eq. (2.72) This GGA functional is used for the calculations.

#### HYBRID METHODS

The Kohn-Sham extensions revealed that improved expressions for electron exchange and correlation can lead to better wave functions and electronic energy. Therefore major attempts were carried out in the past to find improved and better electron exchange and correlation expressions.

Wave function theory (Hartree-Fock) can in principle provide the exact exchange energy.

$$E_x^{HF} = -\frac{1}{2} \sum_i^n \sum_j^n \int \int \frac{\Psi_i(\vec{r}_1) \Psi_j(\vec{r}_1) \Psi_i(\vec{r}_2) \Psi_j(\vec{r}_2)}{|\vec{r}_1 - \vec{r}_2|} d\vec{r}_1 d\vec{r}_2 \quad (2.77)$$

It could be assumed that the ideal exchange/correlation energy would then be obtained from

$$E_{xc} = E_x^{HF} + E_c^{DFT} \quad (2.78)$$

In hybrid DFT methods in general, the exchange term of DFT is corrected by a contribution from exact exchange energy. In the above, the DFT exchange has been replaced by HF exchange.

**BECKE3** The most popular hybrid methods are based on *Becke's three functional B3* [91]. It contains three adjustable parameters  $a_0$ ,  $a_x$  and  $a_c$ :

$$E_{xc}^{B3} = a_0 E_x^{HF} + (1 - a_0) E_x^{LSDA} + a_x \Delta E_x^{B88} + E_c^{LSDA} + a_c \Delta E_c^{GGA} \quad (2.79)$$

Becke used PW91 for the correlation part [86, 88, 92]. The name of the functional, B3PW91, implies its use of a three-parameter scheme, as well as GGA exchange and correlation functionals B and PW91, respectively.

$$E_{xc}^{B3LYP} = a_0 E_x^{HF} + (1 - a_0) E_x^{LSDA} + a_x \Delta E_x^{B88} + E_c^{LSDA} + a_c \Delta E_c^{PW91} \quad (2.80)$$

THE BECKE 3 LEE, YANG AND PARR (B3LYP) HYBRID-FUNCTIONAL This is most frequently used in the quantum chemical computations. The DFT functional B3LYP comprises three different units: B stands exchange functional developed by A. Becke. The ‘3’ stands for the three adjustable parameters (see Eq. (2.79)), Stephens et al.,[93] modified LYP [90] correlation which was developed by Lee, Yang and Parr. Because LYP is designed to compute the full correlation energy and not a correction to LSDA. The B3LYP model is defined by

$$E_{xc}^{B3} = a_0 E_x^{HF} + (1 - a_0) E_x^{LSDA} + a_x \Delta E_x^B + (1 - a_c) E_c^{LSDA} + a_c E_c^{LYP} \quad (2.81)$$

Of all modern functionals, B3LYP has proved to be most popular to date. The B3LYP methods are used for the geometry optimization and solvation energy calculations.

### 2.3.2 BASIS SETS

There are two types of basis functions commonly used in the electronic structure calculations: *Slater Type Orbital* (STO) and *Gaussian Type Orbitals* (GTO). Slater Type Orbital have the functional form

$$\chi^{STO} = N r^{n-1} e^{-\zeta r} Y_{lm}(\Theta, \phi) \quad (2.82)$$

$N$  is a normalization constant and  $Y_{lm}$  are the usual spherical harmonic functions. The exponential dependence on the distance between the nucleus and the electron mirrors the exact orbitals for the hydrogen atom. The Amsterdam Density Functional code, ADF uses the Slater-type orbitals as basis functions.

Gaussian type orbitals can be written in polar or Cartesian coordinates.

$$\chi^{GTO} = N x^l y^m z^n e^{-\alpha r^2} \quad (2.83)$$

where the sum of  $x^l$ ,  $y^m$  and  $z^n$  determines the type of the orbital (for example  $x^l + y^m + z^n = 1$  is a p-orbital).

The preference for GTO basis functions in HF and related methods are due to the computational advantages these functions offer. The advantages include a very efficient algorithm for analytically calculating the huge number of four-center-two-electron integrals occurring in the Coulomb and HF-exchange terms. The STOs seem to be the natural choice for basis functions. They are simple exponentials that mimic the exact eigen functions of the hydrogen atom. GTOs are combined to reproduce as accurately as possible as STOs. It is referred to as a ‘contracted’ basis function. ‘Single- $\zeta$  basis set’ or minimal basis set implies that there is one and only one basis function defined for each type of orbital core through valance. A basis set with two functions for each atomic orbital is called a ‘double- $\zeta$  basis’ and three is called a ‘triple- $\zeta$  basis’.

Valence orbitals can vary widely as a function of chemical bonding. To increase the flexibility in the valence orbitals, 'split-valence' basis sets were developed. In such basis sets, core orbitals are represented by a single (contracted) basis function, while valence orbitals are split into arbitrarily many functions. The most widely used split-valence basis sets are 3-21G, 4-31G, 6-31G and 6-311G.

Polarization functions are used to give more flexibility to adequately describe the wave function, the problem with the calculation is that  $s$  and  $p$  functions centered on the atoms do not provide sufficient mathematical flexibility. Single (\*) and double(\*\*) polarization functions were introduced by Pople and co-workers [94–97]. Thus 6-31G\* implies a set of  $d$  functions added to polarize the  $p$  functions in  $d$ -orbital and in 6-31G\*\* the second star implies  $p$  functions on H and He.

The highest energy molecular orbitals of anions and loose supermolecular complexes are tend to be much more spatially diffuse than normal molecular orbitals. When a basis set does not have the flexibility necessary to allow a weakly bound electron to localize far from the remaining density, diffusion functions are used with the basis set. The presence of diffusion functions is indicated by a '+' in the basis set name. Thus 6-31+G\* indicates that heavy atoms have been augmented with an additional one  $s$  and one set of  $p$  functions having small exponents. The 6-31++G\*\* second plus indicates the presence of diffuse  $s$  functions on H.

### 2.3.3 SOLVATION FREE ENERGIES

The solvation energies can be obtained from electrostatic calculations by solving Poisson-Boltzmann equation as well as from QC calculations. The theory behind the solvation free energy and calculations are explained in the following section.

The interaction of a solute with the surrounding solvent is the free energy of solvation  $\Delta G_{solv}^{\circ}$ . This quantity is some times called the free energy of transfer which refers change in the free energy for a molecule leaving the gas phase and entering into the aqueous phase. Different physical effects contribute to the overall solvation process. The most important components are electrostatic interactions, cavitation, changes in dispersion and changes in bulk solvent structure. When a solute is immersed in a solvent, its charge distribution interacts with that of the solvent. The idea of modeling the electrostatic portion of solvation by placing the solute in a cavity immersed in a dielectric continuum dates back to Born model [98] for ionic solvation and to the work by Krikwood [99] and Onsager [100] on solvation effects in polar molecules.

Both microscopic and macroscopic theoretical methods are available for the estimation of solvation free energies. In determining the relative or absolute  $pK_a$  values,  $\Delta G_{solv}$  can be obtained from the experiments or by ab initio methods and the  $\Delta G_{solv}$  can be calculated by using Monte Carlo simulations, simulation methods combined with a reaction field approach, integral equation techniques and continuum dielectric methods [101].

#### CLASSICAL ELECTROSTATIC CONTINUUM DIELECTRIC METHOD

The solvation energies can be obtained from electrostatic calculations by solving Poisson-Boltzmann equation by finite difference method [101–104]. The solvation free energy is the

free energy difference for charging the solute in vacuum and in solution. Solute-solvent interactions is treated within the frame work of classical electrostatics. The polar solvation energy is given by the following equation where  $q_i^{depro}$  and  $q_i^{pro}$  denotes charges of the deprotonated and protonated state of the solute and  $\phi_i^{Rf,depro}$  and  $\phi_i^{Rf,pro}$  are the reaction field potential for the deprotonated and protonated state respectively. A variety of methods exist to compute atomic partial charges which are described in detail on chapter 3.

$$\Delta G_{solv} = \frac{1}{2} \sum_{i=1}^N q_i^{depro} \phi_i^{Rf,depro} - \frac{1}{2} \sum_{i=1}^N q_i^{pro} \phi_i^{Rf,pro} \quad (2.84)$$

The solute is represented by a set of atomic charges and Bondi radii and the solvent as a continuous dielectric medium (see section 2.2.1). The solvation energies are based on electrostatic calculations which can be done by a two-step procedure. In this method, first step is the QM calculations in vacuum are performed to compute the electron density distribution and wave function. From the electron density the point charges can be obtained by procedures described in section 3.1.2. In the second step the electrostatic free energies of solvation are evaluated from the point charge distribution by solving the Poisson-Boltzmann equation (see section 2.2.1). This approach is taken by most classical PBE solver as classically implying the charge distribution is not allowed to change.

The technical challenges associated with this method are like any numerical method. It is more successful when the density of discrete points is very high but more computationally demanding. The potential can be very sensitive to the grid points very near the cavity surface where the dielectric constants are changing instantaneously. Some times, the location of the point charges describing the solute in general do not coincide with the grid points and is called 'grid artifact'.

#### POLARIZABLE CONTINUUM MODEL (PCM)

The Polarizable Continuum Model (PCM) by Tomasi and coworkers [105–113] is one of the most frequently used continuum solvation methods and has seen numerous variations over the years [63, 105, 114–118]. In continuum model the solute is represented by a charge distribution  $\rho(r)$  in a cavity surface embedded in an infinite polarizable dielectric medium with permittivity  $\epsilon$ . The  $\rho_0(r)$  is change in the charge distribution when the charge distribution induces a reaction potential in the dielectric which can again influence on the charge distribution. The term  $\rho(r)$  is the whole charge distribution including electrons and nuclei. The classical electrostatics can be applied to calculate the free energy of solvation.

In PCM the polarized dielectric is considered as a perturbation to the solute Hamiltonian  $H_0$ . The surface charge distribution produces an electrostatic potential,  $V_\sigma(\vec{r})$  which can be added to the solute Hamiltonian  $H_0$  to take into account the polarization produced by the solvent in the continuum model:

$$H = H_0 + V_\sigma \quad (2.85)$$

Several other continuum models have been proposed and some of which use cavities of molecular shape, in particular models based on conductor-like description of the solvent.

The PCM model calculates the molecular free energy in solution as the sum over three terms:

$$\Delta G_{sol} = G_{es} + G_{ds} + G_{cav} \quad (2.86)$$

These components represent the electrostatic (*es*), the dispersion-repulsion (*dr*) contributions to the free energy and the cavitation energy (*cav*).

The cavitation energy arises from the work needed to form the cavity where the solute is accommodated and from the van der Waals interactions with the solvent molecules of the first solvent shells. The cavitation energy is expressed as a sum over the spheres forming the cavity

$$\Delta G_{cav} = \sum_i^{spheres} \frac{A_i}{4\pi R_i^2} G_i^{HS} \quad (2.87)$$

where  $R_i$  is the radius of the  $i$ th sphere and  $G_i^{HS}$  is the cavitation energy for a sphere of  $R_i$  in a fluid of hard spheres according to Pierotti's expression [119].  $A_i$  is the area of the portion of the sphere  $i$  that is actually exposed to the solvent [109].

The PCM suffers by the following limitations. Due to the fact that the wave function extends outside the cavity, the sum of the charges on the surface is not equal and opposite to the charge of the solute. This causes the 'outlying charge error'. The contribution of the charge on this cavity surface element is found separately using the Gauss' theorem.

#### CONDUCTOR-LIKE PCM (C-PCM)

In C-PCM method, the "outlying charge" error is not corrected according to Gauss' law as the outlying charge effects are included in the model.

In the C-PCM approach the solute is placed in a cavity formed by the envelope of the spheres centered on the atoms or atomic groups [120]. The dielectric constant inside the cavity is same as in vacuum and outside it takes the value of the desired solvent (e.g.,  $\epsilon = 32.6$  for methanol, 78.4 for water, etc.). Once the cavity has been defined, the surface is smoothly mapped by small regions called 'tesserae' (small tiles); each tessera is characterized by the position of its center, its area ( $a_i$ ) and the vector normal to the surface passing through its center.

The cavity is defined through interlocking van der Waals-spheres centered at atomic positions. The conductor-like reaction field is represented through point charges located on the surface of the molecular cavity (Apparent Surface Charge (ASC) model [121]).

For  $\epsilon = +\infty$ , the solvation charges are the solution of the following equation:

$$\vec{S}\vec{q} = -\vec{V} \quad (2.88)$$

where  $\vec{q}$  and  $\vec{V}$  are the solvation charges and the total electrostatic potential generated by solute nuclei and electrons on tesserae, respectively and matrix elements  $S$  [112] are

$$S_{ii} = 1.0694 \sqrt{\frac{4\pi}{a_i}} \quad (2.89)$$

$$S_{ij} = \frac{1}{|\vec{r}_i - \vec{r}_j|} \quad (2.90)$$

where  $a_i$  is the area of the tesserae. The quantity  $\vec{S}\vec{q}$  is the electrostatic potential due to the solvation charges. For real solvents with finite dielectric constant, Eq. (2.88) becomes

$$\vec{S}\vec{q} = -f(\epsilon)\vec{V} \quad (2.91)$$

The general form for the  $\epsilon$  dependent factor is  $f(\epsilon)=\epsilon-1/\epsilon-X$ : in the literature, different values for  $X$  are recommended. The solute-solute interaction energy is

$$E_{int} = \vec{V}^\dagger \vec{q} = \sum_i V_i q_i \quad (2.92)$$

where  $V_i$  is the solute electrostatic potential in tessera  $i$ . Because the solvation charges depend on the solute electron density linearly, it can be shown [122] that the quantity is variationally minimized in the Hartree-Fock (HF) and Kohn-Sham (KS) calculations which corresponds to the free energy. If  $E_0 = E[\rho_0] + V_{NN}$  is the solute energy in vacuum, the free energy in solution is

$$G = E[\rho] + V_{NN} + \frac{1}{2} \vec{V}^\dagger \vec{q} = E[\rho] + V_{NN} - \frac{1}{2} f(\epsilon) \vec{V}^\dagger \vec{S}^{-1} \vec{V} \quad (2.93)$$

where  $V_{NN}$  is the nuclear repulsion energy,  $\rho_0$  is the electronic density for the molecules in vacuum, and  $\rho$  is the density perturbed by the solvent. The Eq. (2.93) is obtained from the Eq. (2.91). This expressions apply to all variational methods for which the “electron density” can be defined.

# CHARGE DISTRIBUTION IN PEPTIDE CONFORMATIONS

---

Atomic charges are very helpful in understanding chemical reactivity and physical properties of solids and liquids [79, 80, 123]. The calculation of effective atomic charges plays an important role in the application of molecular mechanics, molecular dynamics and Monte Carlo calculations. However, a reliable method to calculate appropriate charges for the molecular systems is still lacking. Despite all fundamental problems, there is a need for reliable procedures for the calculation of atomic charges. Although the Mulliken population analysis is popular, it suffers from severe limitations and the charges cannot be compared for smaller systems with different level of theory.

Electrostatic interactions play a central role in molecular recognition, substrate diffusion, catalytic rates and protein folding [124–127]. In particular, they influence the  $pK_a$  of amino acid side chains and other protonatable groups and cofactors [128–130]. The charge distributions have more impact on the interaction energies within the proteins [131]. The charged and polar groups are found ubiquitously in biological macromolecules. More than 20% of all amino acids in globular proteins are ionized under physiological conditions and polar groups are found in more than 25% side chains [132].

This chapter focuses on exploring the relationship between the charge distributions obtained by different charge methods and the sensitivity of charges to the conformation of the tripeptide Ala-Asn-Ala. The charge distribution of asparagine which is present as a titratable group in proteins is studied with different conformations. Sigfridsson and Ryde [133] compared different charge methods to derive partial atomic charges from quantum chemical electrostatic potentials. The different charge methods include Mulliken population analysis, Natural Population Analysis (NPA), Atoms In Molecule (AIM), CHELP (CHarges from ELeCtrostatic Potentials), CHELPG (CHarges from ELeCtrostatic Potentials using a Grid based method), Merz-Kollman (MK) and RESP (Restrained ElectroStatic Potential model) [133]. They observed that among the standard methods, the MK charges gave the best moments and potentials, but they show an appreciable dependence on the orientation of the molecule.

Extensive studies have been reported considering water as a model system for the comparison of several procedures for the calculation of effective atomic charges [134]. An overview of atomic charges and the different procedures to derive them has been given by Wieberg and Rablen [135]. Maciel and co-workers [136] derived charges from the electrostatic potentials and explored the dependence of the theory and the geometry optimization levels to study the

effect of these methods on dipole moments. They checked the ability of these charges by reproducing the experimental dipole moments for 89 molecules. Various methods like RESP and CHELP-Bow were used to drive partial atomic charges of the disaccharides from the quantum chemical electrostatic potential and moments [137].

The atomic charges for various conformations was studied by Richards and co-workers [138] who reported two related methods, both of them employ a constrained minimization of the difference between the quantum mechanical and classical MEP (Molecular Electrostatic Potential) with respect to the atomic charges. The first method involves determining the MEP and constraining the charges to reproduce the dipole at an alternative geometry. The second method involves determining the MEP for each conformation of interest and weighting the MEP for each conformation according to the appropriate Boltzmann factor.

The charge distribution of the tripeptide Ala-Asn-Ala with different conformations have been analyzed using the hybrid density functional method (B3LYP) with 6-31G\* basis set. The population analysis methods such as Mulliken, NPA and the electrostatic potential methods CHELPG, MK and RESP are used to analyze the charge distribution of the tripeptide Ala-Asn-Ala. The extensive studies on the charge methods can provide a better understanding of each method and the parameters which influence the partial atomic charges. The different charge methods discussed in this chapter are used to obtain the point charges for the electrostatic calculations and for the calculation of  $pK_a$  values of  $Cu_B$  ligands in aqueous solution.

## 3.1 THEORETICAL BACKGROUND

In this section the theory behind charge methods, the advantage and disadvantages of each charge methods are explained.

### 3.1.1 WAVE FUNCTION BASED METHODS

#### MULLIKEN POPULATION ANALYSIS

In Mulliken population analysis, the electrons are divided up amongst the atoms according to the degree to which different atomic orbital (AO) basis functions contribute to the overall wave function [79, 80, 139, 140]. The total number of electrons are obtained by expanding the wave function in its AO basis set,

$$N = \sum_j^{electrons} \int \psi_j(\vec{r}_j) \psi_j(\vec{r}_j) d\vec{r}_j \quad (3.1)$$

$$= \sum_j^{electrons} \sum_{r,s} \int c_{jr} \varphi_r(\vec{r}_j) c_{js} \varphi_s(\vec{r}_j) d\vec{r}_j \quad (3.2)$$

$$= \sum_j^{electrons} \left( \sum_r c_{jr}^2 + \sum_{r \neq s} c_{jr} c_{js} S_{rs} \right) \quad (3.3)$$

where  $\psi$  is the wave function of molecule  $j$ ,  $r$  and  $s$  are the index of AO basis function  $\varphi$ ,  $c_{jr}$  is the coefficient of the basis function  $r$  in molecular orbital  $j$ ,  $c_{js}$  is the coefficient of the basis function  $s$  in molecular orbital  $j$ .  $S_{rs}$  is the usual overlap matrix element [79]. From Eq. (3.3) it's clear that the total number of electrons are divided into two sums, one including only the squares of single AO basis functions and the other including products of two different AO basis functions. The first term in the parentheses in Eq. (3.3), is the sum of electrons associated with single basis function only and the second term, which represents the electrons 'shared' between basis functions. The total number of electrons per atom  $N_k$  is given by

$$N_k = \sum_j^{electrons} \left( \sum_{r \in k} c_{jr}^2 + \sum_{r, s \in k, r \neq s} c_{jr} c_{js} S_{rs} + \sum_{r \in k, s \notin k} c_{jr} c_{js} S_{rs} \right) \quad (3.4)$$

The orthonormality of basis functions of different angular momentum both residing on the same atom  $k$  causes many terms in the second sum of Eq. (3.4) to be zero. The Mulliken partial atomic charge is defined as

$$q_k = Z_k - N_k \quad (3.5)$$

where  $Z_k$  is the nuclear charge of atom  $k$  and  $N_k$  is computed according to the Eq. (3.4).

With minimal or small split-valence basis sets, the Mulliken charges tend to be reasonably intuitive. The use of non-orthogonal basis set in the Mulliken analysis, however, can lead to some undesirable results. In Mulliken analysis where the total number of electrons are divided over AO basis functions, it is possible for individual basis functions to have occupation numbers greater than 1 or less than 0. Such situations obviously can have no physical meaning. The electrons are also divided equally in the case atoms with different electronegativities. Mulliken partial charges prove to be very sensitive to basis-set size, so that comparisons of partial atomic charges from different levels of theory are in no way possible. Moreover, with very complete basis sets, the Mulliken charges have a tendency to become unphysical.

#### NATURAL POPULATION ANALYSIS

Natural Atomic Orbitals (NAOs) were defined by P. O. Löwdin [141] as the eigenfunctions of the first-order reduced density matrix [142]. They are intrinsic orbitals  $\{\theta_k\}$  that arise as eigenfunctions of the first-order reduced density operator  $\hat{\Gamma}$ ,

$$\hat{\Gamma}\theta_k = q_k\theta_k \quad (3.6)$$

which is formed by ‘reducing’ the wavefunction probability distribution to the single-particle level,

$$\hat{\Gamma} = N \int \psi(1, 2, \dots, N) \psi^*(1', 2 \dots N) d\tau_2 \dots d\tau_N \quad (3.7)$$

and whose eigenorbitals are hence ‘natural’ to the  $\psi$  itself. The subsystems of the  $N$ -particle system are best formulated in terms of reduced density operators. In particular, the squared amplitude  $|\langle \psi(1, 2, \dots, N) | \phi(1) \rangle|^2$  that an electron of  $\psi(1, 2, \dots, N)$  is ‘in’ orbital  $\phi$  (i.e., the population of  $\phi(1)$  in the wavefunction) is rigorously expressed, for any possible orbital  $\phi$ , as

$$q_\phi = \langle \phi | \hat{\Gamma} | \phi \rangle \quad (3.8)$$

The occupancies  $q_\phi$  are intrinsically non-negative and limited by the *Pauli exclusion principle* e.g., for the spatial orbital  $\phi(r)$ ,

$$0 \leq q_\phi \leq 2 \quad (3.9)$$

The sum of occupancies  $q_k$  over any complete orthonormal set  $\{\theta_k\}$  accounts for all  $N$  electrons,

$$\sum_k \langle \phi_k | \hat{\Gamma} | \phi_k \rangle = \sum_k q_k = N \quad (3.10)$$

Mulliken population commonly fail to satisfy the physical constrains (Eq. 3.9 and 3.10). The natural orbitals  $\{\theta_k\}$  of Eq. (3.6) are intrinsically best possible for considering maximum occupancy in the smallest possible number of orbitals.

The starting point for the Natural Population Analysis (NPA) is the determination of an optimal set of effective atomic orbitals for the molecular environment.

To find the maximum-occupancy orbitals for atom A, the local eigenvectors  $\tilde{\theta}_i^{(A)}$  of the one-center atomic block  $\hat{\Gamma}^{(A)}$  associated with basis AOs on atom A is obtained first.

$$\hat{\Gamma}^{(A)} \tilde{\theta}_i^{(A)} = \tilde{q}_i^{(A)} \tilde{\theta}_i^{(A)} \quad (3.11)$$

After the maximum-occupancy orbitals  $\{\tilde{\theta}_i^{(A)}\}$  have been obtained for each atom A, it is necessary to restore full orthogonality. The ‘pre-orthogonal’ orbitals  $\{\tilde{\theta}_i^{(A)}\}$  are then subjected to an occupancy-weighted symmetric orthogonalization to find final NAOs  $\{\theta_i^{(A)}\}$ . Once the final orthonormal NAOs  $\theta_i^{(A)}$  are obtained, the Eq. (3.8) are used to rigorously compute their ‘natural’ populations  $q_i^{(A)}$ ,

$$q_i^{(A)} = \langle \theta_i^{(A)} | \hat{\Gamma} | \theta_i^{(A)} \rangle \quad (3.12)$$

which may be summed to give the total number of electrons  $N_A$

$$N_A = \sum_i q_i^{(A)} \quad (3.13)$$

and ‘natural charge’  $Q_A$  on atom A with atomic number  $Z_A$  is given as

$$Q_A = Z_A - N_A \quad (3.14)$$

In accord with Eq. (3.9 and 3.10), the natural populations  $q_i^{(A)}$  necessarily satisfy the Pauli principle ( $0 \leq q_i^{(A)} \leq 2$ ) and sum correctly to the total number of electrons. The Eq. (3.12-3.14) summarize ‘natural population analysis’.

The “Natural Population Analysis” is an alternative to the conventional Mulliken population analysis [79, 80, 143, 144]. It seems to exhibit improved numerical stability and better in describing the electron distribution in compounds of high ionic character. The most appealing feature of the NPA scheme is that each atomic partial charge effectively converges to a stable value with increasing basis-set size.

### 3.1.2 POTENTIAL BASED METHODS

Widely distributed methods to generate atomic partial charges from QC computations are called potential-based methods such as CHELP, CHELPG, MK and RESP method. The charges are calculated by matching the electrostatic potential (ESP) from the instantaneous atomic partial charge distribution with the QC ESP computed on optimized geometries in vacuum. These methods have in common that they use values of the electrostatic potential on spatial positions that are distributed around the atoms of the molecule. All potential-based methods take into account only the potential points outside the vdW radii of the atoms to determine charges. As they cannot be determined undoubtedly, effective vdW radii are a matter of debate in the literature and are treated differently in the potential-based methods. Different schemes are currently in use to select the points outside the vdW radii of the atoms of the molecule.

#### MOLECULAR ELECTROSTATIC POTENTIAL

A truncated multipole expansion is a computationally convenient single-center formalism that allows one to quantitatively compute the degree to which a positive or negative test charge is attracted or repelled by the molecule that is being represented by the multipole expansion [79]. The MEP can be computed exactly on any given position  $\vec{r}$

$$V_{MEP}(\vec{r}) = \sum_k^{nuclei} \frac{Z_k}{|\vec{r} - \vec{r}_k|} - \int \Psi(\vec{r}') \frac{1}{|\vec{r} - \vec{r}'|} \Psi(\vec{r}') d\vec{r}' \quad (3.15)$$

Where  $Z_k$  is the nuclear charge and the sum run overs all the atom  $k$ . The MEP is an observable, although in practice it is rather difficult to design appropriate experiments to measure it. Computationally, the molecular electrostatic potential is usually evaluated as

$$V_{MEP}(\vec{r}) = \sum_k^{nuclei} \frac{Z_k}{|\vec{r} - \vec{r}_k|} - \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}' \quad (3.16)$$

where  $\rho$  is the electron density. The partial atomic charges are typically derived from MEP. The common notation MEP is replaced with ESP for, 'electrostatic potential'. The ESP is the most obvious property to obtain the partial atomic charges that will be useful to model molecular-molecular interactions at short and long range. All ESP charge-fitting schemes involve determining atomic partial charges  $q_k$ ,

$$V_{ESP}(\vec{r}) = \sum_k^{nuclei} \frac{q_k}{|\vec{r} - \vec{r}_k|} \quad (3.17)$$

Although there are many positive aspects of the ESP derived charges, often they are found to be very sensitive to small changes to the molecular structure or in the sampling of the ESP.

#### CHELP METHOD

In CHELP (CHarges from ELectrostatic Potentials) method [79, 145] the electrostatic potential is determined at a selected number of points around the molecule chosen in spherical shells, 1 Å apart, of fourteen nearly symmetrically placed points around each atom. Point which fall within the vdW radius of any of the atoms are discarded due to the large distortions caused by close proximity to the nuclei. Typically 100-300 points are selected in the region extending to 3 Å from the vdWs surface of the molecule is considered.

From Eq. (3.16) the ESP is calculated quantum chemically and a least square procedure is used to fit the charge  $q_j$  to each atomic center  $j$  in the molecule. The calculated classical ESP is given by Eq. (3.17). The  $\chi_{esp}^2$  to be minimized by least-squares procedure is defined as in Eq. (3.18).

$$\chi_{esp}^2 = \sum_i (V_i - E_i)^2 \quad (3.18)$$

At the minimum, for all  $j$

$$\frac{\partial(\chi_{esp}^2)}{\partial q_j} = 0 \quad (3.19)$$

where

$$\frac{\partial(\chi_{esp}^2)}{\partial q_j} = -2 \sum_i \frac{V_i - E_i}{r_{ij}} = 0 \quad (3.20)$$

and thus a system of equations can be formed and solved in matrix form. The best least square fit of charges to the electrostatic potential is obtained by finding the minimum of  $y$ ,

$$y(q_1, q_2, \dots, q_n) = \sum_{i=1}^m [V_i - E_i(q_1, q_2, \dots, q_n)]^2 \quad (3.21)$$

where  $m$  is the total number of points to fit,  $V_i$  is from QC wave functions as given in Eq. (3.16) and  $E_i$  is the electrostatic potential in the monopole approximation and given in Eq. (3.17).

The minimum of  $y$  can be obtained by finding the stationary points of the Lagrangian function  $z$ :

$$z(q_1, q_2, \dots, q_n) = y(q_1, q_2, \dots, q_n) + \lambda g(q_1, q_2, \dots, q_n) \quad (3.22)$$

where  $g$  is the constraint imposed on the fit. For example, following constraints can be used: constraining the total charge, molecular dipole, the charges of a group of atoms to a particular value and two or more atoms having equivalent charges.

The linear least square problem with imposed constraints can be solved using the Lagrangian multiplier method. The function corresponding to total charge constraint is:

$$g(q_1, q_2, \dots, q_n) = \left( \sum q_i \right) - q_{tot} = 0 \quad (3.23)$$

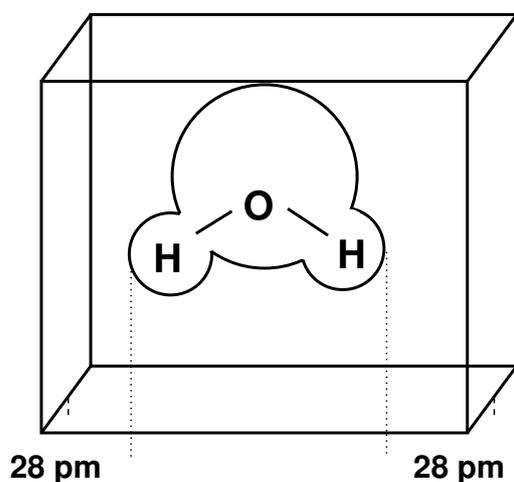
The extremum points of  $z$  are then obtained by solving  $\partial z / \partial \lambda = 0$  and  $\partial z / \partial q_k = 0$ . A set of  $n+1$  equations are obtained where  $n$  is the number of atoms. The solution of corresponding matrix equation yields the charges.

The CHELP method has rotational invariance problems due to the sparse point sampling and dependency upon the molecular coordinate system for choosing test points.

#### CHELPG METHOD

The principle difference between the CHELPG (CHarges from ELectrostatic Potentials using a Grid based method) and CHELP is that CHELPG [79, 146] procedure employs a point selection

algorithm based upon regularly spaced points. The algorithm involves defining a cube of points spaced 0.3-0.8 Å apart containing the molecule and including 2.8 Å of head space on all sides. Figure 3.1 describes the situation for water as a simple example. All points outside the 2.8 Å radius are also eliminated from the fitting procedure. The remaining points form a relatively homogeneous layer around the molecule. The electrostatic potential at each sample points is calculated analytically from the wave function and geometry data. These data are then used as input for the Lagrange multiplier least-squares routine, which has been constrained to fit the exact molecular charge, but not to the molecular dipole moment. CHELPG typically utilize 5000-6000 points. The CHELPG charges are obtained for all the five conformations by varying the grid points per square angstrom. In the present study the grid points in the CHELPG method is varied from 1 to 6 (points per square angstrom) in ESP fit.



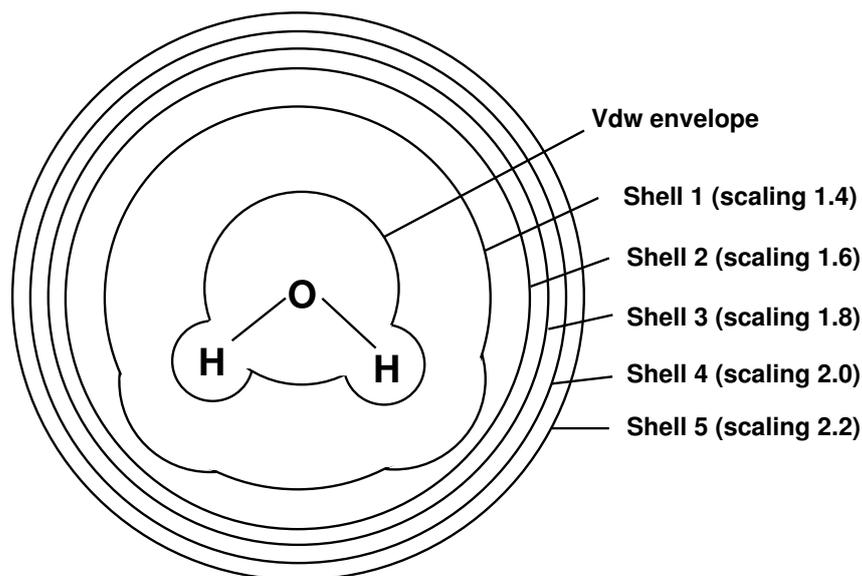
**Figure 3.1. CHarges from ELectrostatic Potentials using a Grid based method.** The dimensions of the cube are chosen such that the molecule is located at the center of the cube, adding 28.0 pm head space between the molecule and the end of the box in all three dimensions.

One of the weak points of CHELPG is the treatment of larger systems, in which some of the innermost atoms are located far away from the points at which the ESP is evaluated. In such situation, variations of the innermost atomic charges will not lead to significant changes in the ESP outside the molecule and fitting of these atomic charges will therefore not result in meaningful results. Representative atomic charges for flexible molecules should therefore be derived as average values over several conformers.

#### MK METHOD

The Merz Kollman (MK) method [79, 147, 148] uses the Connolly surface algorithm [149] to calculate a shell or a number of shells of points at a specified distance from the molecule and then calculate the electrostatic potential at these points. The use of successive shells of various multiples of the vdWs radii, ranging from 1.4 to 2.0 times the radius (smaller radii lead to positive potentials in the lone-pair region). The shells of 1.4, 1.6, 1.8, and 2.0

times the vdWs radius are used to avoid artifacts from being too close to the atoms. The density of points in these surfaces depended on the size of the molecule and typically 200-300 points/molecules are generated by varying 1-5 points per square angstrom. The Merz Kollman method is described for water in Figure 3.2. The water molecules is covered with four shells with 1.4, 1.6, 1.8, 2.0 and 2.2 times the vdWs radius. In the present study the MK charges are obtained for all the five conformations by varying the grid points and the shells. The grid points varied from 1 to 6 and the shell is varied from 2 to 9 to the maximum limit with 1.4, 1.6, 1.8, 2.0, 2.2, 2.4, 2.6, 2.8 and 3.0 times the vdWs radius.



**Figure 3.2. Merz-Singh-Kollman method.** The ESP is calculated at a number of grid points located on several shells around the molecule. The shells are constructed as an overlay of van der Waals spheres around each atom. The shells of 1.4, 1.6, 1.8, 2.0 and 2.2 times the vdWs radius are shown.

The major strength of the electrostatic potential charges is that they optimally reproduce the intermolecular interaction properties of the molecules with a simple two-body additive potential, if suitably accurate level of quantum chemical calculation is used to derive the ESP around the molecule. The major weakness of these charges is they are not easily transferable between common functional groups in related molecules [133].

#### RESTRAINED ELECTROSTATIC POTENTIAL METHOD (RESP)

In RESP method [150], the restraints are introduced in the form of penalty function into the fitting procedure considerably to overcome the limitations of other ESP charge methods.

With the charge fitting procedure  $\chi_{esp}^2$  (see section 3.1.2), an addition penalty function  $\chi_{rstr}$  is added. So  $\chi^2$  is to be minimized now

$$\chi^2 = \chi_{esp}^2 + \chi_{rstr}^2 \quad (3.24)$$

and the least squares minimum is defined as

$$\frac{\partial(\chi^2)}{\partial q_n} = \frac{\partial(\chi_{esp}^2)}{\partial q_n} + \frac{\partial(\chi_{rstr}^2)}{\partial q_n} \quad (3.25)$$

The simple harmonic Eq. (3.26) or hyperbolic Eq. (3.27) function is used as a penalty function.

$$\chi_{rstr}^2 = a \sum_m (q_{0n} - q_n)^2 \quad (3.26)$$

$$\chi_{rstr}^2 = a \sum_m ((q_n^2 + b^2)^{1/2} - b) \quad (3.27)$$

where  $a$  is the scaling factor determining the strength of the restraint and  $q_0$  is the target for the restraint. Mulliken charges are used as a target charges [150]. Generally the RESP charges are used in molecular mechanics form any organic and bio-organic systems.

In the present study, the RESP charges are obtained by using both Mulliken and NPA charges as initial charges for the harmonic restraint. The RESP charges are obtained from both electrostatic potentials calculated from CHELPG and MK methods.

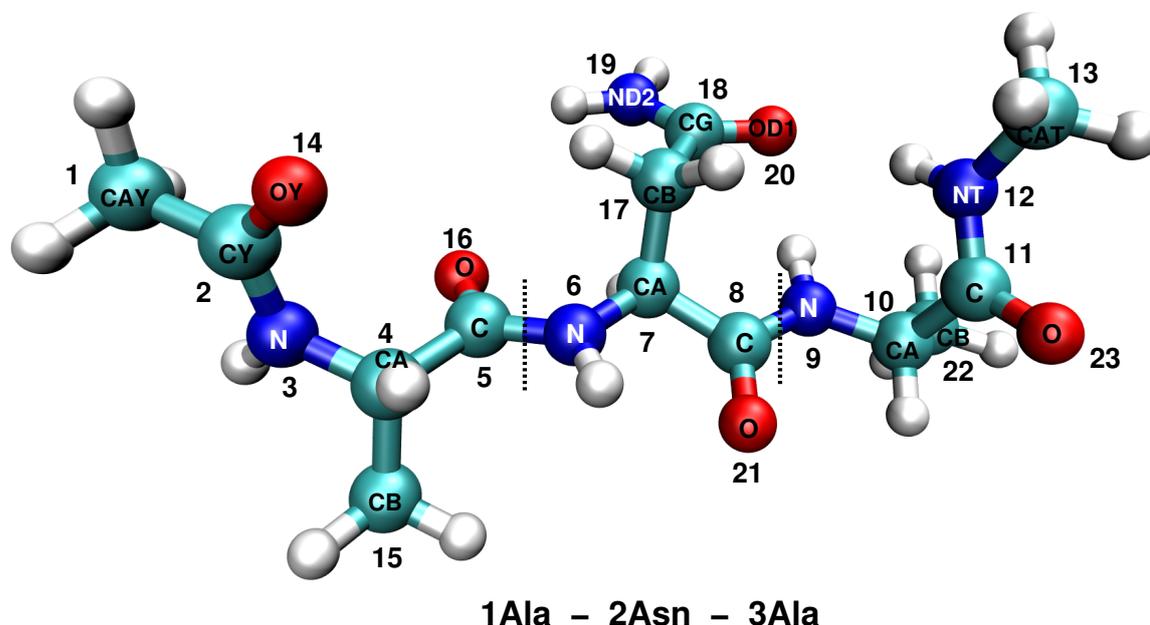
#### MULTIPOLE DERIVED CHARGES

Another set of charges used in the work is multipole derived charges. The multipole derived charges (MDC) are used to obtain the solvation energies based on FDPB method combination with DFT level of theory (see section 5.3.1). The MDC analysis is based on the ideas used for the *Dipole Preserving Charge analysis* [151] but formulated by using more accurate atomic dipoles. There are three stages involved in the method: First the molecular electron density is written as a sum of atomic densities. Second, from these atomic densities a set of atomic multipoles are defined, which is used to get the electrostatic potential outside the charge distribution. Third, these atomic multipoles are reconstructed by using a scheme that distributes charges over all atoms to reproduce these multipoles exactly. It is claimed that this method can overcome the drawbacks of the ESP methods because this method uses the atomic multipole expansion over molecular multipole expansion. A detailed description of atomic multipole expansion method is given in Ref [152]. The MDC used in the present work are charges derived from monopole (MDC<sub>m</sub>), dipole (MDC<sub>d</sub>) and quadrupole (MDC<sub>q</sub>) moments.

## 3.2 STRUCTURE PREPARATION AND COMPUTATIONAL METHODS

The initial structure of the tripeptide Ala-Asn-Ala was builded using CHARMM program [153] package. The C-terminal alanine and the N-terminal alanine are capped with acetate and methylate groups respectively. Five different conformations are generated by changing the

dihedral angles of atoms 6, 7, 17 and 18 from  $120^\circ$  to  $180^\circ$  by step of  $15^\circ$ . The tripeptide considered for the charge calculation is shown in Figure 3.3

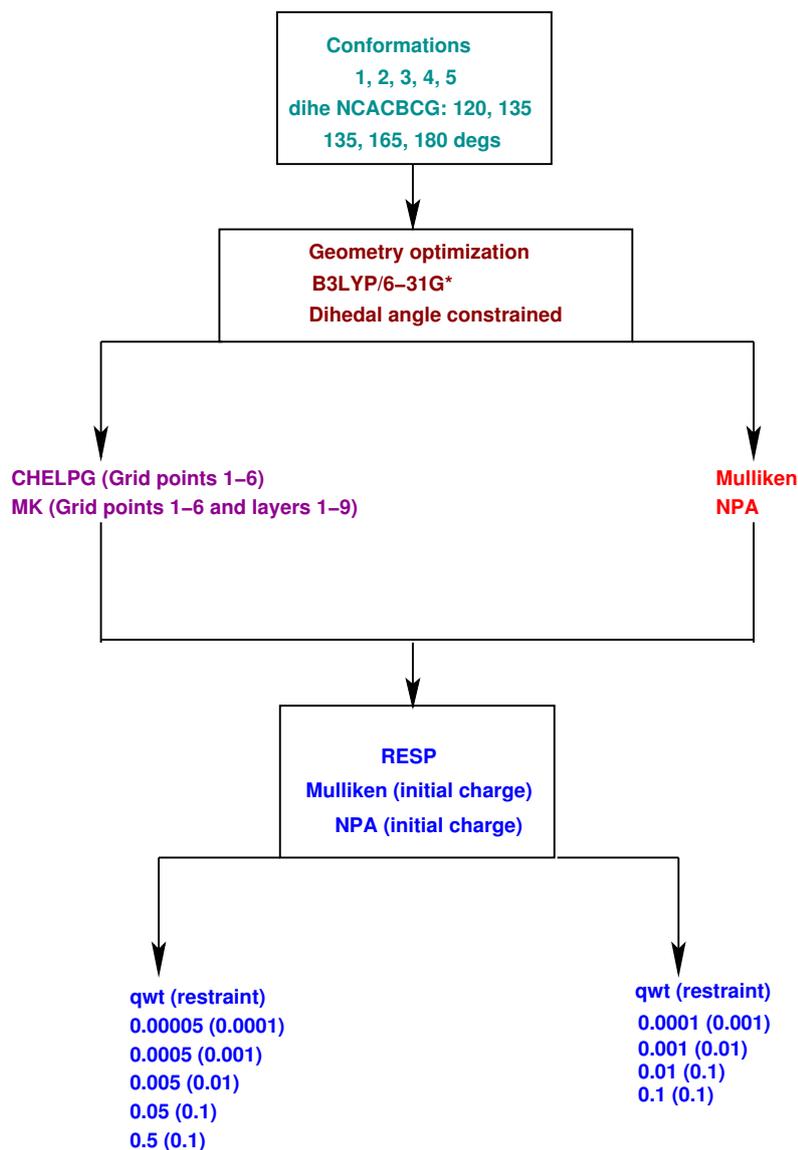


**Figure 3.3. The tripeptide Ala-Asn-Ala considered for the charge calculation is shown.** The tripeptide Ala-Asn-Ala is shown with the PDB atom type.

Each conformation was optimized by fixing the dihedral angle. All the conformations were optimized by hybrid density functional theory B3LYP with 6-31G\* basis set. The geometry optimizations were performed using GAUSSIAN 03 program [154]. The CHELPG and MK charges were also obtained using the same program package. The RESP charges were obtained using RESP program [155].

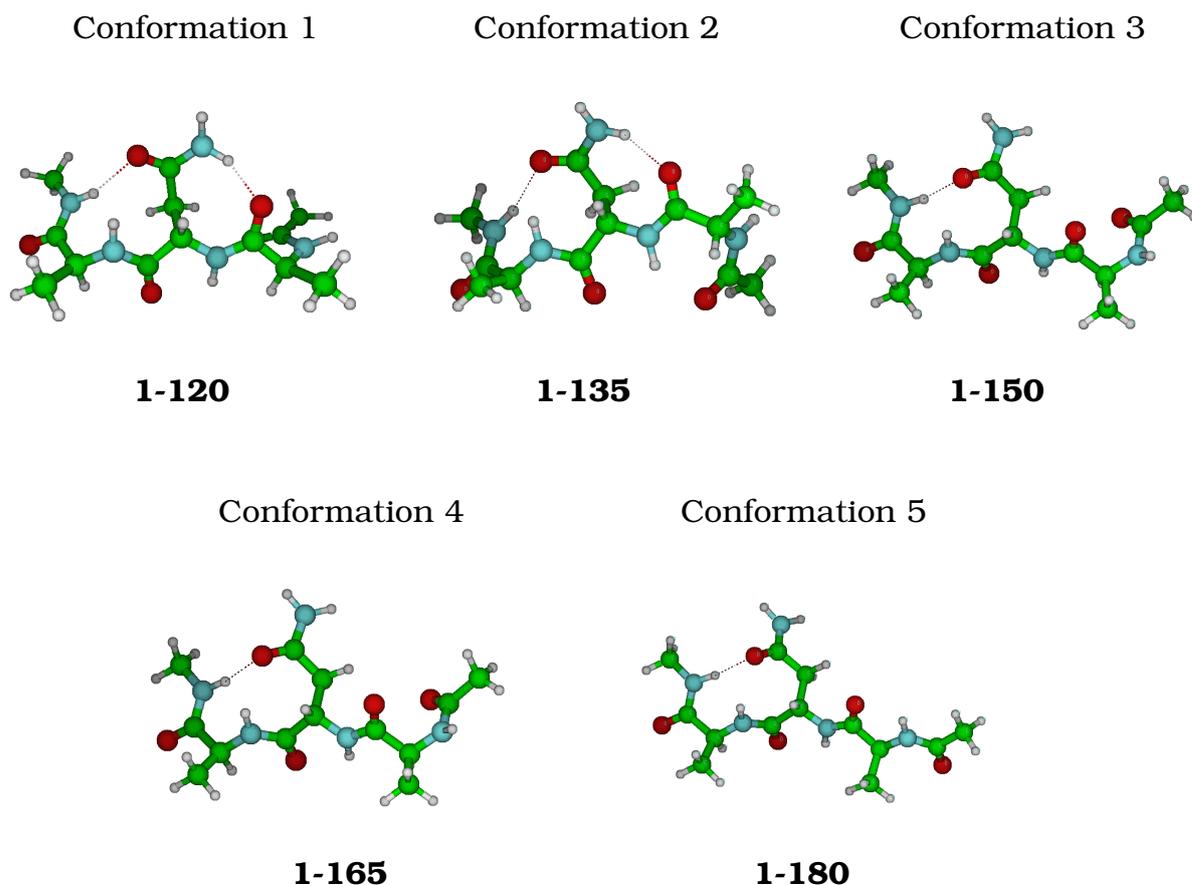
The complete set of calculations performed on tripeptide Ala-Asn-Ala are shown in Figure 3.4. Population analysis were performed for all the conformations. The choice of the population analysis method includes the traditional Mulliken population analysis [139, 140] and NPA [143, 144], both of these methods are based on the molecular wave function, but NPA uses the natural atomic basis function to overcome the drawback of the Mulliken population analysis. Various methods are available to fit the molecular electrostatic potential to the grid points with fitting procedures. The three methods considered in our studies are the CHELPG [145], MK [147, 148], and RESP [156] methods. The CHELPG method is a grid based method where points are selected in a regularly spaced cubic grid. To obtain charges independent of the conformations of the tripeptide, the sampling points were increased by varying the grid points in CHELPG and MK method and the charges were calculated for each conformation.

In CHELPG method, the grid points per square angstrom are varied from 1 to 6 in the ESP fit by invoking the GAUSSIAN 03 option **IOp(6/42=N)** where N is the points per square angstrom. MK method uses the Connolly surface algorithm [149] to calculate a shell or a number of shells



**Figure 3.4.** Scheme for calculations performed on the tripeptide Ala-Asn-Ala

of points at a specific distance from the molecule and the electrostatic potential is calculated at each point. The MK charges were calculated by varying the grid points and the number of shells for each conformations. The grid points varied from 1 to 6 using the IOP in Gaussian 03. The shells were varied from 2 to 9 to the maximum limit. This options is invoked by **IOP(6/41=N)** where N is number of shells in ESP charge fit. The RESP charges are calculated from both electrostatic potentials obtained by CHELPG and MK methods. The Mulliken and NPA charges were used as initial charges for the harmonic restraint. The B3LYP/6-31G\* optimized geometry of all the five different conformations of tripeptide Ala-Asn-Ala are shown in Figure 3.5.



**Figure 3.5.** The B3LYP/6-31G\* optimized geometry of all the five different conformations of tripeptide Ala-Asn-Ala and the nomenclature are shown. The B3LYP/6-31G\* optimized geometry of all the five different conformations of tripeptide Ala-Asn-Ala considered for the calculations. Five different conformations are generated by changing the dihedral angles of atoms 6, 7, 17 and 18 from 120° to 180° by step of 15°.

### 3.3 RESULTS AND DISCUSSION

The Mulliken and NPA charges calculated by B3LYP/6-31G\* basis for different conformations are given in Table 3.1 and 3.2. The partial atomic charges obtained from the ESP methods such as CHELPG and MK with varying grid points (1-6) can be better represented in graphical form as shown in Figures 3.6 and 3.7 respectively. The partial atomic charges obtained from MK method by increasing the shells (2-9) are shown in Figure 3.8. The RESP charges obtained from using Mulliken and NPA charges as initial charges for the harmonic restraint and the electrostatic potential obtained from CHELPG and MK with varying grid points and shells are given in electronic supporting information. All the partial atomic charges obtained of five different conformations 1-120, 2-135, 3-150, 4-165 and 5-180 are given in electronic supporting information.

### 3.3.1 MULLIKEN CHARGES

The Mulliken charges calculated by B3LYP/6-31G\* basis for different conformations are given in Table 3.1. Mulliken charges range from -0.73 to +0.60 (Table 3.1). There is a little variation in the partial atomic charge distribution as a function of the dihedral rotations. When the atomic charges of the heavy atoms are compared, Mulliken charges remain more negative in side chain of atom type 2Asn-ND2. All the oxygen atoms in the side chain show atomic charges in the range of -0.51 to -0.56. The atomic charge of nitrogen atoms in the backbone range from -0.58 to -0.61. Mulliken charges are used as initial target charge for the RESP method.

### 3.3.2 NPA CHARGES

The NPA charges calculated by B3LYP/6-31G\* basis for different conformations are given in Table 3.2. The NPA charges tend to be larger in magnitude compared to the Mulliken charges. NPA charges range from -0.83 to +0.71 (Table 3.2). The atom type 2Asn-ND2 has more negative charge similar to that of Mulliken charge. The variation of the atomic charge distribution as a function of the dihedral rotations is less compared to the Mulliken charges.

### 3.3.3 CHELPG CHARGES

The CHELPG method is an improved version of CHELP method for computing potential derived charges. It is shown that the charges derived from CHELPG method are considerably less dependent on molecular orientation than CHELP method. All atomic charges calculated by CHELPG method are given in the electronic supporting information. The atomic charges of the heavy atoms in the backbone and the side chains are shown in Figure 3.6 and 3.7 (A.1 to A.6) respectively. The ESP charges exhibit a much stronger conformational dependence than the Mulliken and NPA charges. All the atomic charges shown in Figure 3.6 and 3.7 (A.1 to A.6) were obtained by varying the grid points from 1 to 6, to get suitable point sampling where the charges do not show any dependency to the conformation. The density of points are increased to include more sample point in the ESP fit to find out the reasonable points per square angstrom. From Figure 3.6 and 3.7, it is clear that increasing the sampling points do not have any influence on the charges of different conformations.

### 3.3.4 MK CHARGES

All atomic charges calculated by MK method are given in the electronic supporting information. The atomic charges of the heavy atoms in the backbone and the side chains are given in Figure 3.6 and 3.7 (B.1 to B.6) respectively. Like CHELPG charges, the MK charges also exhibit a much stronger conformational dependency than the Mulliken and NPA charges. Compared to the CHELPG charges, the MK charges show more conformational dependency. In CHELPG method the charges were obtained from the electrostatic potentials using a grid based method, a point selection algorithm similar to the one employed by Cox and Williams [157] which dramatically reduces these orientation dependencies where as in MK method the charges were obtained from the electrostatic potentials using the Connolly surface. All the

| Residue     | Atom Type | 1-120   | 2-135   | 3-150   | 4-165   | 5-180   | $\bar{x}$ | $\sigma$ |
|-------------|-----------|---------|---------|---------|---------|---------|-----------|----------|
| Backbone:   |           |         |         |         |         |         |           |          |
| 1Ala        | CAY       | -0.5407 | -0.5330 | -0.5433 | -0.5461 | -0.5330 | -0.5392   | 0.0054   |
|             | CY        | 0.6087  | 0.6016  | 0.6079  | 0.6060  | 0.5945  | 0.6037    | 0.0052   |
|             | N         | -0.5832 | -0.5925 | -0.5793 | -0.5738 | -0.6050 | -0.5868   | 0.0110   |
|             | CA        | -0.0264 | -0.0526 | -0.0322 | -0.0374 | -0.0299 | -0.0357   | 0.0092   |
|             | C         | 0.6583  | 0.6305  | 0.6563  | 0.6450  | 0.6134  | 0.6407    | 0.0169   |
| 2Asn        | N         | -0.6161 | -0.6321 | -0.6170 | -0.6214 | -0.6324 | -0.6238   | 0.0071   |
|             | CA        | -0.0317 | -0.0297 | -0.0374 | -0.0248 | -0.0158 | -0.0279   | 0.0073   |
|             | C         | 0.5890  | 0.5929  | 0.6069  | 0.5991  | 0.5954  | 0.5967    | 0.0061   |
| 3Ala        | N         | -0.6132 | -0.6248 | -0.6262 | -0.6182 | -0.6150 | -0.6195   | 0.0052   |
|             | CA        | -0.0565 | -0.0563 | -0.0584 | -0.0553 | -0.0571 | -0.0567   | 0.0010   |
|             | C         | 0.6048  | 0.6028  | 0.6055  | 0.6043  | 0.6028  | 0.6040    | 0.0011   |
|             | NT        | -0.6336 | -0.6333 | -0.6299 | -0.6264 | -0.6304 | -0.6307   | 0.0026   |
|             | CAT       | -0.2954 | -0.2935 | -0.2934 | -0.3111 | -0.3080 | -0.3003   | 0.0077   |
| Side chain: |           |         |         |         |         |         |           |          |
| 1Ala        | OY        | -0.5182 | -0.5123 | -0.5162 | -0.5207 | -0.5220 | -0.5179   | 0.0034   |
|             | CB        | -0.4854 | -0.4610 | -0.4816 | -0.4764 | -0.4463 | -0.4701   | 0.0145   |
|             | O         | -0.5695 | -0.5610 | -0.5733 | -0.5559 | -0.5468 | -0.5613   | 0.0095   |
| 2Asn        | CB        | -0.3753 | -0.3664 | -0.3733 | -0.3854 | -0.3842 | -0.3769   | 0.0071   |
|             | CG        | 0.5562  | 0.5596  | 0.5717  | 0.6173  | 0.6187  | 0.5847    | 0.0277   |
|             | ND2       | -0.7396 | -0.7371 | -0.7252 | -0.7377 | -0.7382 | -0.7356   | 0.0052   |
|             | OD1       | -0.5309 | -0.5262 | -0.5274 | -0.5424 | -0.5391 | -0.5332   | 0.0064   |
| 3Ala        | O         | -0.5164 | -0.5077 | -0.5269 | -0.5330 | -0.5346 | -0.5237   | 0.0102   |
|             | CB        | -0.4437 | -0.4458 | -0.4506 | -0.4454 | -0.4430 | -0.4457   | 0.0027   |
|             | O         | -0.5295 | -0.5324 | -0.5331 | -0.5328 | -0.5299 | -0.5315   | 0.0015   |

**Table 3.1. Mulliken charges of the heavy atoms in the tripeptide backbone and side chains.** Partial atomic charges of the heavy atoms obtained at B3LYP/6-31G\* basis set, the average ( $\bar{x}$ ) and standard deviations ( $\sigma$ ) of atomic charges are given.

atomic charged shown in the Figures 3.6 and 3.7 (B.1 to B.6) were obtained by increasing the grid points from 1 to 6 and the number of shells are increased from 2 to 9 the maximum limit. Similar to CHELPG method, the MK method also show that increasing the sampling points per square angstrom do not have any influence on the charges of different conformations. Figure 3.8 show the MK charges obtained with different shells for different conformations. There is a slight deviation in the atomic charges when 2 shells are used. When the shells are

| Residue     | Atom Type | 1-120   | 2-135   | 3-150   | 4-165   | 5-180   | $\bar{x}$ | $\sigma$ |
|-------------|-----------|---------|---------|---------|---------|---------|-----------|----------|
| Backbone:   |           |         |         |         |         |         |           |          |
| 1Ala        | CAY       | -0.7635 | -0.7612 | -0.7638 | -0.7648 | -0.7601 | -0.7627   | 0.0017   |
|             | CY        | 0.7010  | 0.6969  | 0.7018  | 0.7028  | 0.6910  | 0.6987    | 0.0043   |
|             | N         | -0.6541 | -0.6529 | -0.6558 | -0.6567 | -0.6623 | -0.6564   | 0.0032   |
|             | CA        | -0.1311 | -0.1406 | -0.1331 | -0.1356 | -0.1411 | -0.1363   | 0.0040   |
|             | C         | 0.7071  | 0.6953  | 0.6950  | 0.6882  | 0.6969  | 0.6965    | 0.0061   |
| 2Asn        | N         | -0.6338 | -0.6434 | -0.6424 | -0.6505 | -0.6514 | -0.6443   | 0.0064   |
|             | CA        | -0.1352 | -0.1360 | -0.1312 | -0.1304 | -0.1295 | -0.1325   | 0.0026   |
|             | C         | 0.6973  | 0.6937  | 0.6942  | 0.6928  | 0.6894  | 0.6935    | 0.0025   |
| 3Ala        | N         | -0.6616 | -0.6648 | -0.6567 | -0.6519 | -0.6468 | -0.6564   | 0.0065   |
|             | CA        | -0.1448 | -0.1461 | -0.1473 | -0.1482 | -0.1484 | -0.1470   | 0.0014   |
|             | C         | 0.6780  | 0.6790  | 0.6793  | 0.6808  | 0.6801  | 0.6794    | 0.0010   |
|             | NT        | -0.6614 | -0.6617 | -0.6604 | -0.6568 | -0.6584 | -0.6597   | 0.0019   |
|             | CAT       | -0.4740 | -0.4737 | -0.4736 | -0.4769 | -0.4765 | -0.4749   | 0.0014   |
| Side chain: |           |         |         |         |         |         |           |          |
| 1Ala        | OY        | -0.6361 | -0.6440 | -0.6322 | -0.6347 | -0.6440 | -0.6382   | 0.0049   |
|             | CB        | -0.6851 | -0.6875 | -0.6844 | -0.6835 | -0.6801 | -0.6841   | 0.0024   |
|             | O         | -0.6810 | -0.6828 | -0.6746 | -0.6509 | -0.6560 | -0.6691   | 0.0131   |
| 2Asn        | CB        | -0.5536 | -0.5503 | -0.5559 | -0.5628 | -0.5622 | -0.5570   | 0.0049   |
|             | CG        | 0.6911  | 0.6931  | 0.6956  | 0.7105  | 0.7075  | 0.6996    | 0.0079   |
|             | ND2       | -0.8399 | -0.8411 | -0.8354 | -0.8355 | -0.8373 | -0.8378   | 0.0023   |
|             | OD1       | -0.6743 | -0.6722 | -0.6737 | -0.6808 | -0.6733 | -0.6749   | 0.0030   |
| 3Ala        | O         | -0.6253 | -0.6189 | -0.6388 | -0.6483 | -0.6487 | -0.6360   | 0.0121   |
|             | CB        | -0.6775 | -0.6772 | -0.6779 | -0.6770 | -0.6777 | -0.6775   | 0.0003   |
|             | O         | -0.6397 | -0.6423 | -0.6426 | -0.6463 | -0.6421 | -0.6426   | 0.0021   |

**Table 3.2. NPA charges of the heavy atoms in the tripeptide backbone and side chains.**

Partial atomic charges of the heavy atoms obtained at B3LYP/6-31G\* basis set, the average ( $\bar{x}$ ) and standard deviations ( $\sigma$ ) of atomic charges are given.

increased from 3 to 9, there is much change in the atomic charges, but increasing the shells do not help to find conformationally independent charges.

### 3.3.5 RESP CHARGES

RESP charges were obtained for different conformations using different combination of initial charges and the electrostatic potentials. The models used to obtain RESP charges are

described below. For the Model 1 to 4, both strong and weak restraint weights are used to assess the impact of the scaling factors on the charges and the quality of the fit. The RESP charges were calculated by a standard two-stage fitting procedure using the RESP program which is freely available [155]. We are interested in ascertaining the optimal parameters to obtain conformationally independent charges of the tripeptide Ala-Asn-Ala. The RESP charges obtained by using ESP from CHELPG and MK methods in combination with Mulliken charges and NPA charges as initial charges are given in Figure 3.9. The following scaling factors are used for strong harmonic restraint: 0.05(0.1), 0.5(0.1), 0.01(0.1) and 0.1(0.1) and the weak harmonic restraints as follows, 0.00005(0.0001), 0.0005(0.001), 0.005(0.01), 0.0001(0.001), 0.001(0.01). The values in the parenthesis show the restraint used for the second stage fit to obtain equal charges for the methyl groups.

**Model 1:** Harmonic restraint to the Mulliken charges are used with the combination of electrostatic potential from CHELPG method with varying grid points from 1 to 3 (see Figure 3.9).

**Model 2:** Harmonic restraint to the NPA charges are used with the combination of electrostatic potential from CHELPG method with varying grid points from 1 to 3 (see Figure 3.9).

**Model 3:** Harmonic restraint to the Mulliken charges are used with the combination of electrostatic potential from MK method with varying grid points from 1 to 6 (see Figure 3.9).

**Model 4:** Harmonic restraint to the NPA charges are used with the combination of electrostatic potential from MK method with varying grid points from 1 to 6 (see Figure 3.9).

**Model 5:** Harmonic restraint to the NPA charges are used with the combination of electrostatic potential from MK method with varying shells from 2 to 9.

#### MODEL 1

The ESP obtained from CHELPG method by varying grid points from 1-3 are used to calculate the RESP charges, whereas Mulliken charges are used as initial charges for harmonic restraint. The combination of the less grid points with strong harmonic restraint 0.5 or 0.1 in the first stage fit, the RESP charges approach the Mulliken charges. When weak restraints are used the RESP charges approach the ESP charges. The weak restraint beyond 0.00005 does not change the atomic charges much. Still the RESP charges depend on the conformation of the peptide because the ESP used to obtain RESP charges already depend on the conformations.

#### MODEL 2

In Model 2, the NPA charges are used as initial charges instead of Mulliken charges for harmonic restraint. The RESP charges show the same trend as model 1 i.e., when less grid points and strong restraint 0.5 or 0.1 are used 0.5 in the first stage fit, the RESP charges approach the NPA charges. Similar to model 1, restraint beyond 0.00005 does not change the atomic charges much. The restraint factor 0.0005 or 0.001 appeared to be optimal for both model 1 and model 2 to have a balance between the initial charges (Mulliken and NPA) and ESP charges (CHELPG charges).

### MODEL 3

In model 3, the ESP obtained from MK method by varying grid points from 1-6 are used to calculate RESP charges and the Mulliken charges are used as initial charges for harmonic restraint. As in model 1, when less grid points and strong restraint 0.5 or 0.1 are used in the first stage fit, the RESP charges approach the Mulliken charges. When the grid points are increased from 3 to 4, there is no change in the charge and increasing the sampling points do not have any effect on the charges. The calculations show that at least 3 point per square angstrom should be used to get a stable charges and the restraint factor of 0.0005 should be used to get a reasonable charges.

### MODEL 4

In Model 4, the NPA charges are used as initial charges with the combination of ESP from MK method. The trends are almost same. Analysis of the RESP charges also show that more than 3 points per square angstrom should be used to get stable charge with combination of restraint factor 0.0005.

### MODEL 5

In Model 5, the RESP charges are calculated by using the ESP obtained from MK method by varying the shells from 2 to 9. There is an irregular trend in the RESP charges with respect to the conformations. The analysis shows that at least more than 4 shells should be used to get a stable charges for each conformations.

All the models show that the RESP charges approach the initial charges (Mulliken and NPA) when strong restraint is used and the RESP charges approach the ESP charges (CHELPG and MK) when weak restraint is used. The restraint beyond 0.00005 does not make any change in the charges.

## 3.4 CONCLUSIONS

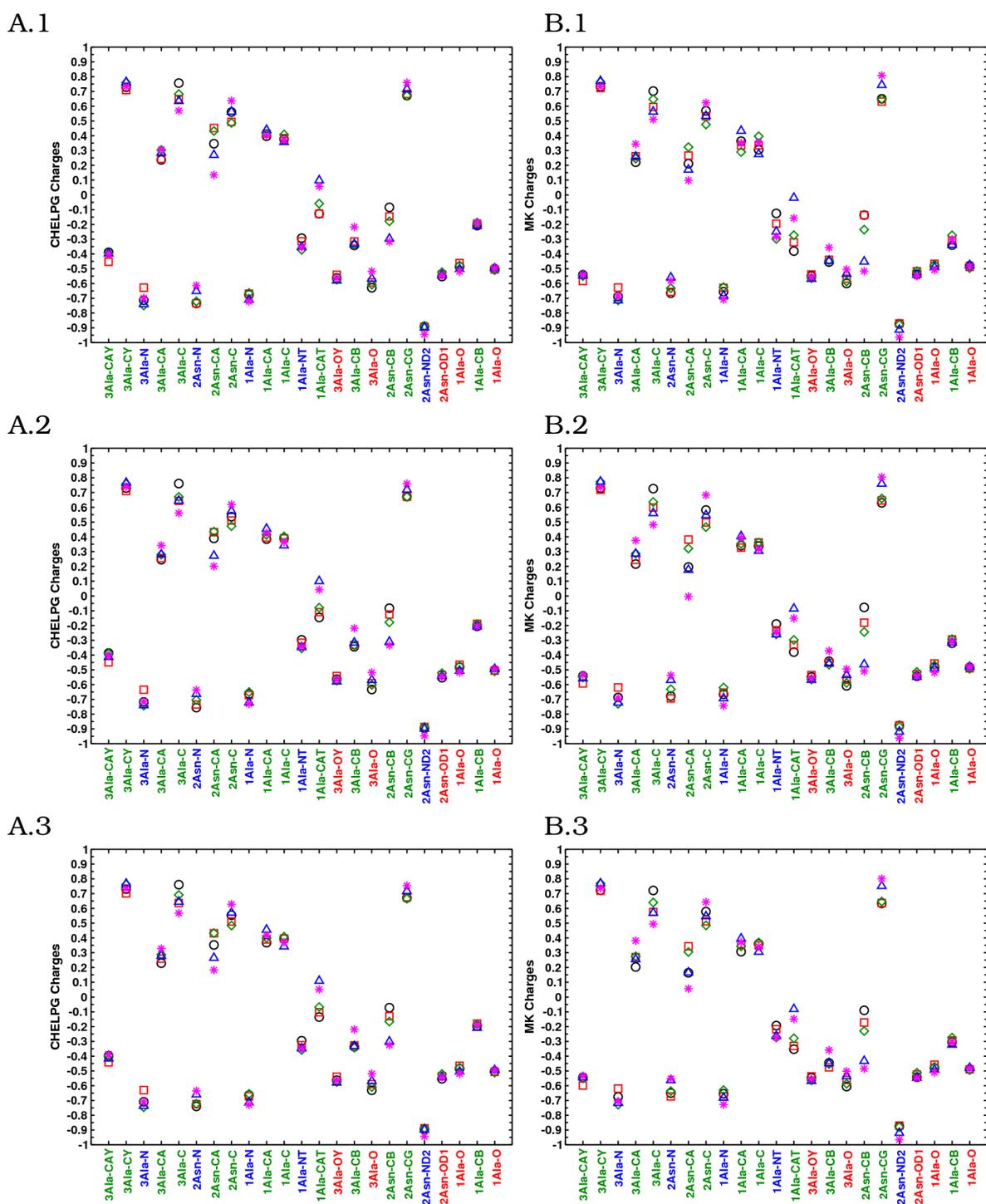
The Mulliken, NPA, CHELPG, MK and RESP atomic charges are derived for different conformation of the tripeptide Ala-Asn-Ala. All the partial atomic charges were obtained from hybrid density function theory B3LYP with 6-31G\* basis set. The sensitivity of charges with different conformations and the influence of wave-function based (Mulliken and NPA) and electrostatic potential based charge methods (CHELPG, MK and RESP) have been explored in detail. The factors influencing the charges are varied which includes increasing the points per square angstrom both in CHELPG and MK method and the shells are varied from 2-9 the maximum limit in MK method. The RESP charges were calculated by using the electrostatic potential from CHELPG and MK method using Mulliken and NPA charges as initial charges with harmonic function as penalty function. The RESP charges were also obtained by using different scaling factor from a strong to weak restraint.

The atomic charges obtained from both CHELPG and MK method are dependent on the conformations. The Mulliken and NPA charges show less conformational dependency compare to

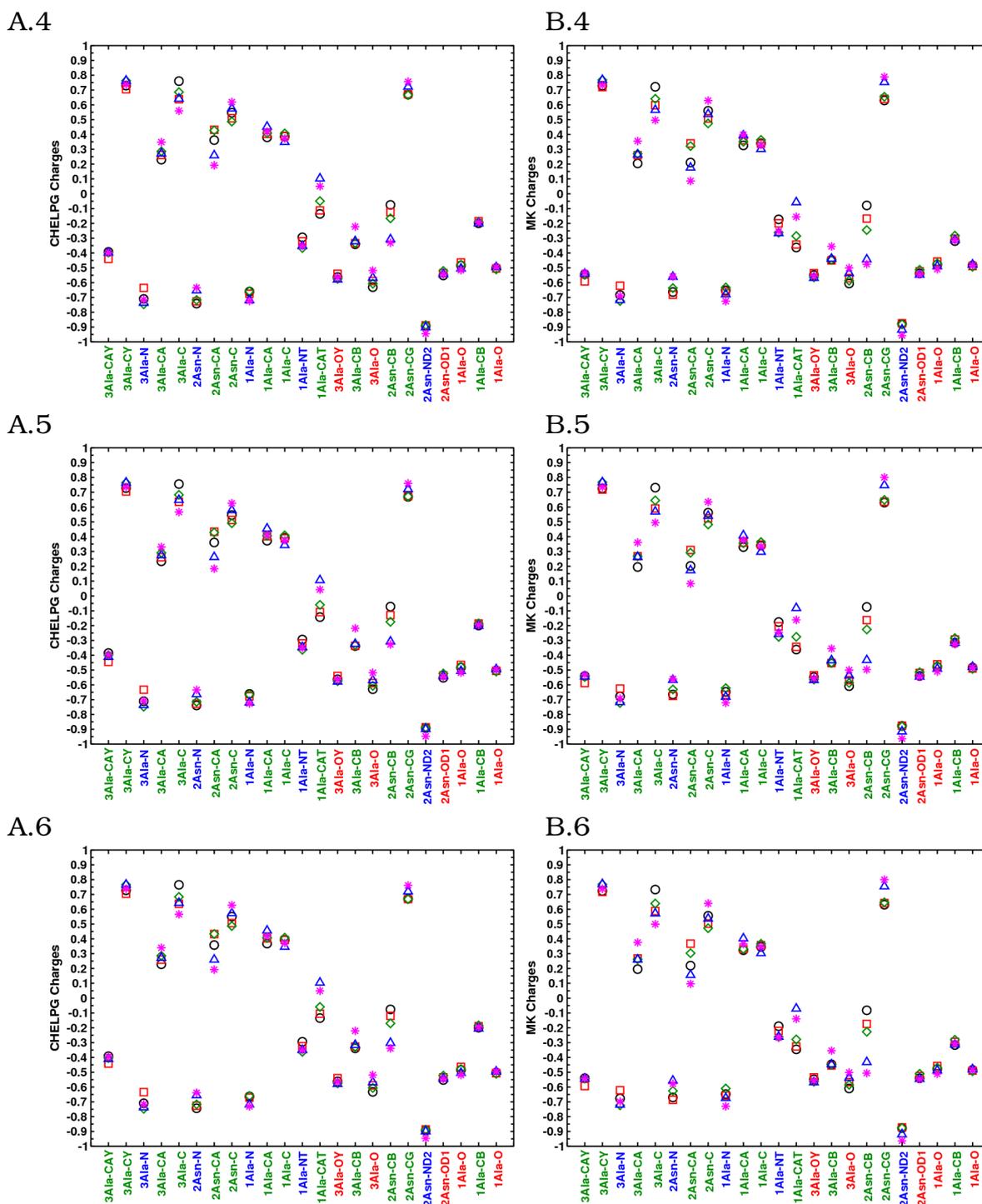
the CHELPG and MK charges. Difference in points selection for the ESP fits can dramatically affect the partial charges. The magnitudes of the calculated point charges were found to be somewhat dependent on the choice of points per square angstrom and the shells. The magnitude of the charges are more or less same for the both CHELPG and MK method when grid points of 3 to 4 per square angstrom were used. The results show that minimum 2 to 3 grid points per square angstrom should be used in both methods and increasing the grid points to maximum limit 6 does not change the magnitude of charge much. The MK charges show less dependent on conformation when minimum 3 to 4 shells were used and the charges remains almost same when shells are increased to maximum limit 9. The results show that minimum 3 to 4 shells should be used in MK method.

The RESP charges are mainly dependent on the strong or weak restraint used. When strong restraints are used the charges are closer to the Mulliken or NPA charges and if weak restraints are used the charges are close to the ESP charges. The magnitude of the charges exhibit much smaller fluctuations when restraint factor of 0.0005 was used in the first stage fit. The restraint weight of 0.0005 appears to be optimal for obtaining RESP charges for tripeptide Ala-Asn-Ala. The model 3 combination works fine to obtain RESP charges for tripeptide Ala-Asn-Ala. The charges predicted by the Mulliken population analysis and NPA schemes are larger than those predicted by the ESP schemes.

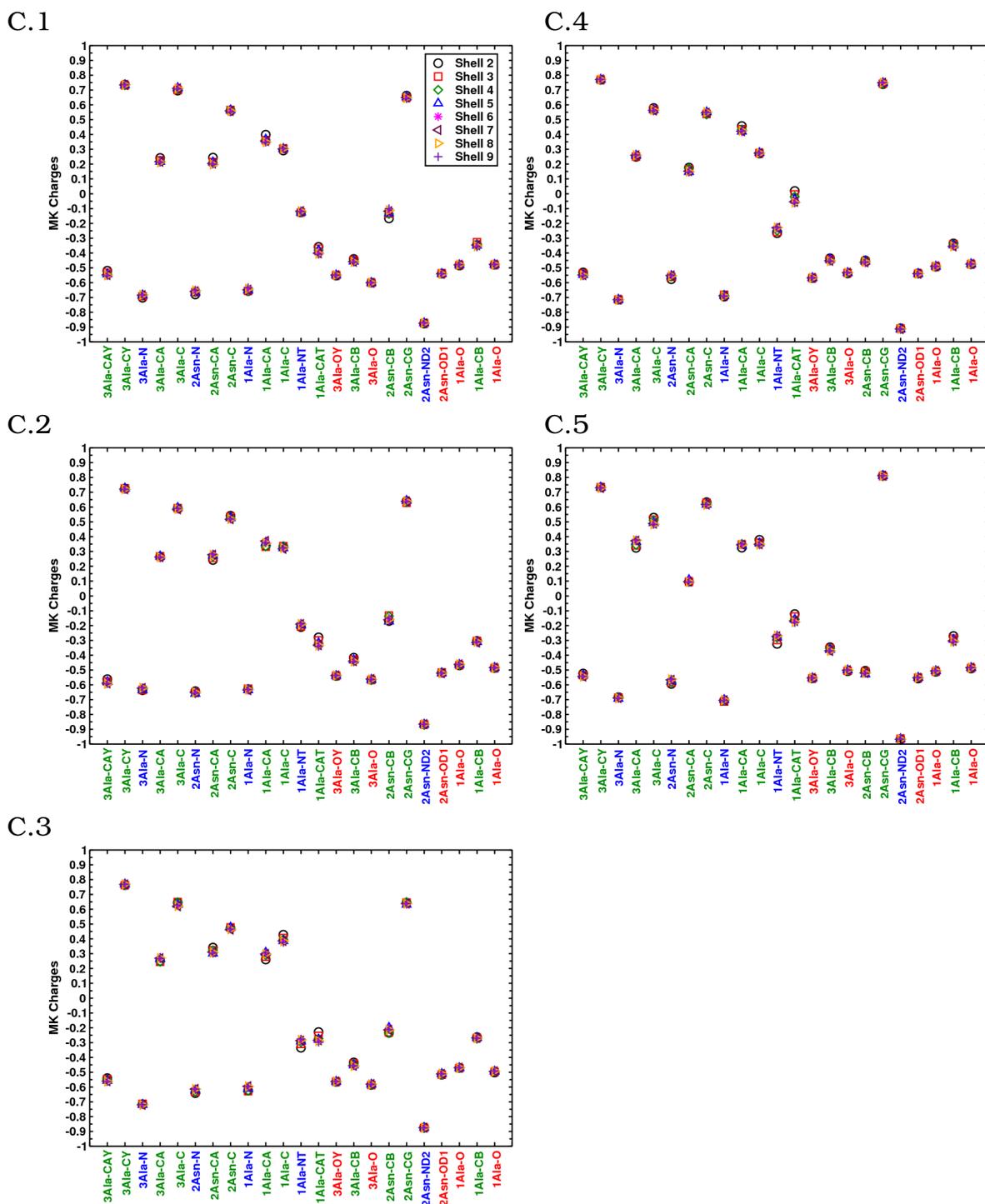
This conclusions should be carefully reconsidered for other systems with polar bonds in which electrons correlation effects may be more significant.



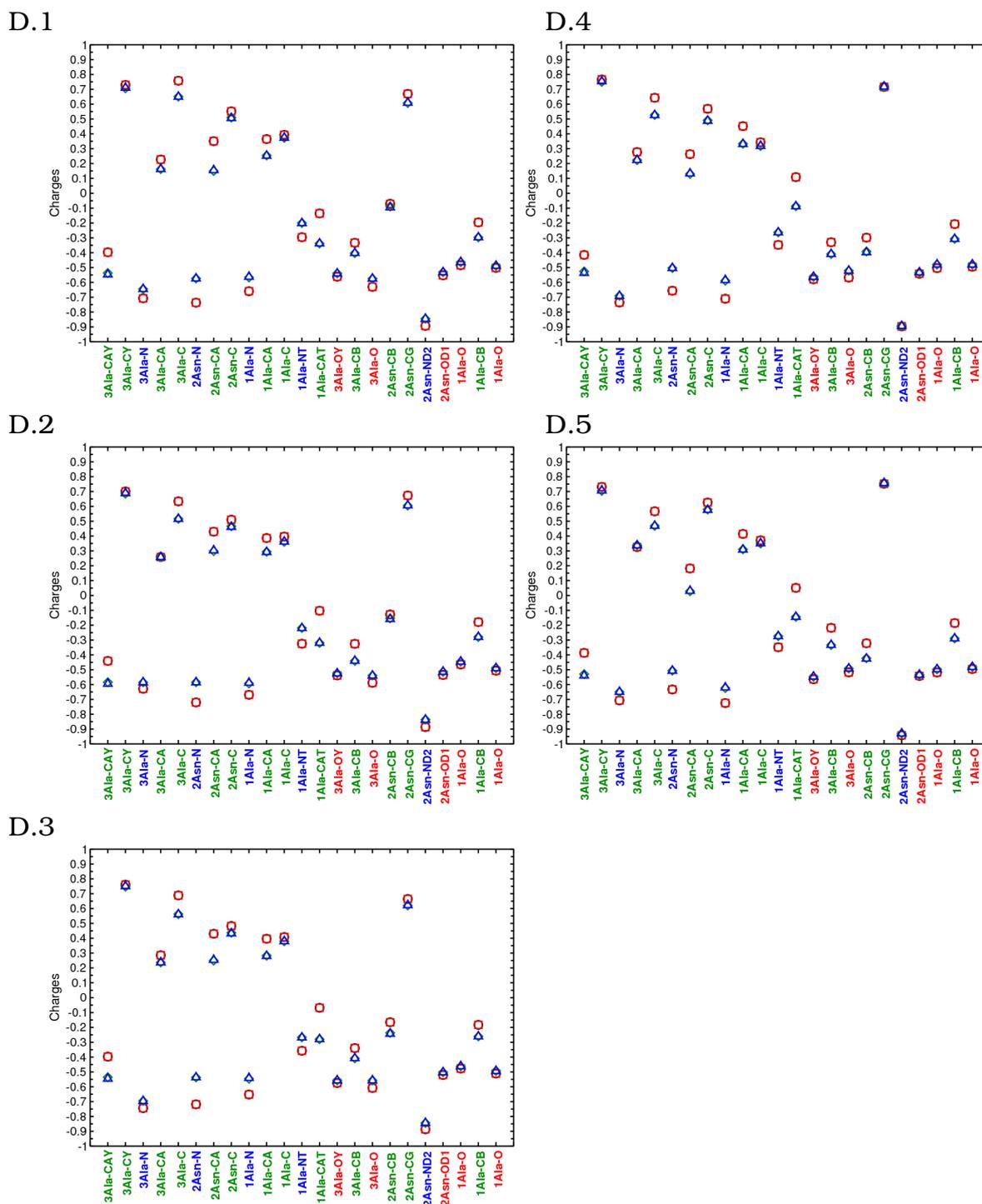
**Figure 3.6. Atomic charges for heavy atoms in the backbone and side chains for different conformations of tripeptide Ala-Asn-Ala obtained from CHELPG and MK methods are shown in the plots.** (A.1). CHELPG charges for heavy atoms in the backbone and side chain where 1 point per square angstrom is used in the CHELPG method. (A.2). CHELPG charges for heavy atoms in the backbone and side chain where 2 points per square angstrom are used in the CHELPG method. (A.3). CHELPG charges for heavy atoms in the backbone and side chain where 3 points per square angstrom are used in the CHELPG method. (B.1). MK charges for heavy atoms in the backbone and side chain where 1 point per square angstrom is used in the MK method. (B.2). MK charges for heavy atoms in the backbone and side chain where 2 points per square angstrom are used in the MK method. (B.3). MK charges for heavy atoms in the backbone and side chain where 3 points per square angstrom are used in the MK method. The black circle (○) corresponds to conformation 1 (1-20), red square (◻) corresponds to conformation 2 (2-135), green diamond (◇) corresponds to conformation 3 (3-150), blue triangle up (△) corresponds to conformation 4 (4-165), magenta star (☆) corresponds to conformation 5 (5-180)



**Figure 3.7. Atomic charges for heavy atoms in the backbone and side chains for different conformations of tripeptide Ala-Asn-Ala obtained from CHELPG and MK methods are shown in the plots.** (A.4). CHELPG charges for heavy atoms in the backbone and side chains where 4 points per square angstrom are used in the CHELPG method. (A.5). CHELPG charges for heavy atoms in the backbone and side chains where 5 points per square angstrom are used in the CHELPG method. (A.6). CHELPG charges for heavy atoms in the backbone and side chains where 6 points per square angstrom are used in the CHELPG method. (B.4). MK charges for heavy atoms in the backbone and side chain where 4 points per square angstrom are used in the MK method. (B.5). MK charges for heavy atoms in the backbone and side chains where 5 points per square angstrom are used in the MK method. (B.6). MK charges for heavy atoms in the backbone and side chains where 6 points per square angstrom are used in the MK method. The black circle (○) corresponds to conformation 1 (1-20), red square (◻) corresponds to conformation 2 (2-135), green diamond (◊) corresponds to conformation 3 (3-150), blue triangle up (△) corresponds to conformation 4 (4-165), magenta star (☆) corresponds to conformation 5 (5-180)



**Figure 3.8. Atomic charges for heavy atoms in the backbone and side chains for different conformations of tripeptide Ala-Asn-Ala obtained from CHELPG and MK methods is shown in the plots.** (C.1). MK charges for heavy atoms in the backbone and side chains of conformation 1 (1-120) where shells varied from 2-9 and 4 points per square angstrom is used in the MK method. (C.2). MK charges for heavy atoms in the backbone and side chains of conformation 2 (1-135) where shells varied from 2-9 and 4 points per square angstrom is used in the MK method. (C.3). MK charges for heavy atoms in the backbone and side chains of conformation 3 (1-150) where shells varied from 2-9 and 4 points per square angstrom is used in the MK method. (C.4). MK charges for heavy atoms in the backbone and side chains of conformation 4 (1-165) where shells varied from 2-9 and 4 points per square angstrom is used in the MK method. (C.5). MK charges for heavy atoms in the backbone and side chains of conformation 5 (1-180) where shells varied from 2-9 and 4 points per square angstrom is used in the MK method.



**Figure 3.9. Atomic charges for heavy atoms in the backbone and side chains for different conformations of tripeptide Ala-Asn-Ala obtained from RESP methods is shown in the plots.** (D.1). RESP charges for heavy atoms in the backbone and side chains of conformation 1 (1-120). (D.2). RESP charges for heavy atoms in the backbone and side chains of conformation 2 (1-135). (D.3). RESP charges for heavy atoms in the backbone and side chains of conformation 3 (1-150). (D.4). RESP charges for heavy atoms in the backbone and side chains of conformation 3 (1-165). (D.5). RESP charges for heavy atoms in the backbone and side chains of conformation 3 (1-180). The black circle ( $\circ$ ) corresponds to RESP charges where ESP from CHELPG method is used with combination of Mulliken charges as initial charges. The red square ( $\square$ ) corresponds to RESP charges where ESP from CHELPG method is used with combination of NPA charges as initial charges. The green diamond ( $\diamond$ ) corresponds to RESP charges where ESP charges MK method is used with combination of Mulliken charges as initial charges. The blue triangle up ( $\Delta$ ) corresponds to RESP charges where ESP from MK method with combination of NPA charges as initial charges.



## PROTONATION AND REDOX POTENTIALS OF CYTOCHROME *c* NITRITE REDUCTASE

---

Cytochrome *c* oxidase is a large membrane protein designed to utilize the energy of electron transfer and oxygen reduction to pump protons across the membrane. The molecular mechanism of the energy conversion process is not been well understood [8, 40–42]. Electrostatic calculations on other proteins with simpler, better resolved structures can help to understand the possible mechanism of electron transfer in cytochrome *c* oxidase.

Cytochrome *c* nitrite reductase can serve as such a simple system. Cytochrome *c* nitrite reductase catalyzes the six electron reduction of nitrite to ammonia. This second part of the respiratory pathway of nitrate ammonification is the key step in the biological nitrogen cycle [158–161].

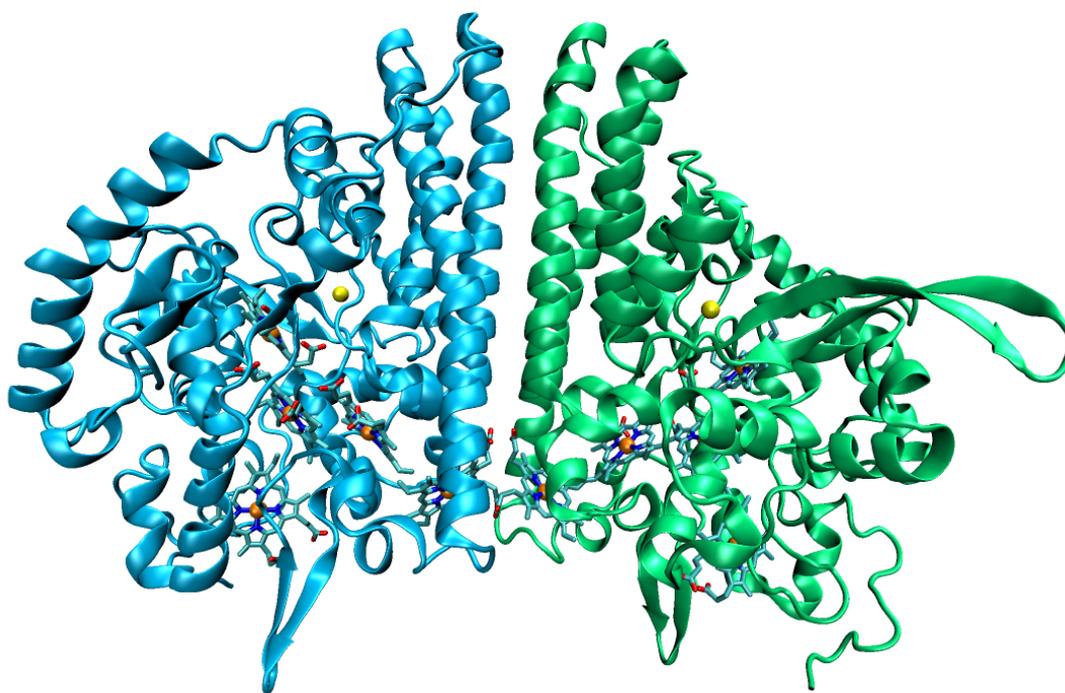
Cytochrome *c* nitrite reductase is a multiheme *c* enzyme and its active site is a protoporphyrin IX which is covalently linked to the protein backbone. A lysine was found to replace the usual histidine as a proximal ligand to the heme iron [162–164]. The five hemes in the monomer of cytochrome *c* nitrite reductase are in close contact with Fe-Fe distance of 9 to 12.8 Å. The cytochrome *c* nitrite reductase dimer is shown in Figure 4.1 and the arrangement of hemes in the dimer is shown in Figure 4.2. The Heme I forms the active site and the other hemes, Hemes II, III and IV are almost coplanar with the catalytic Heme I. Hemes II and V are farther apart and are not coplanar with Hemes I, III and IV. All the hemes except Heme I are bis-histidinyl-coordinated and linked to the peptide backbone by thioether bonds to the cysteine residues of a classical heme-binding motif for periplasmic proteins, Cys-X<sub>1</sub>-X<sub>2</sub>-Cys-His. Heme I however has the binding motif Cys-X<sub>1</sub>-X<sub>2</sub>-Cys-Lys, where the nitrogen atom of the lysine replacing that of histidine.

The cytochrome *c* nitrite reductase catalyzes the reduction of both nitrite and sulfide with high specific activity and gets electrons from the membranous quinol pool, thereby generating a proton motive force. Heme I is clearly the substrate-binding site and previous spectroscopic studies [165, 166] in a similar nitrite reductase had proved that one of the heme was high-spin and it was the active site of the enzyme. The cytochrome *c* nitrite reductase can be found in the periplasm and forms a stable, membrane associated complex with its electron donor NrfH, a member of the NapC/NirT family of the tetraheme cytochromes [167, 168].

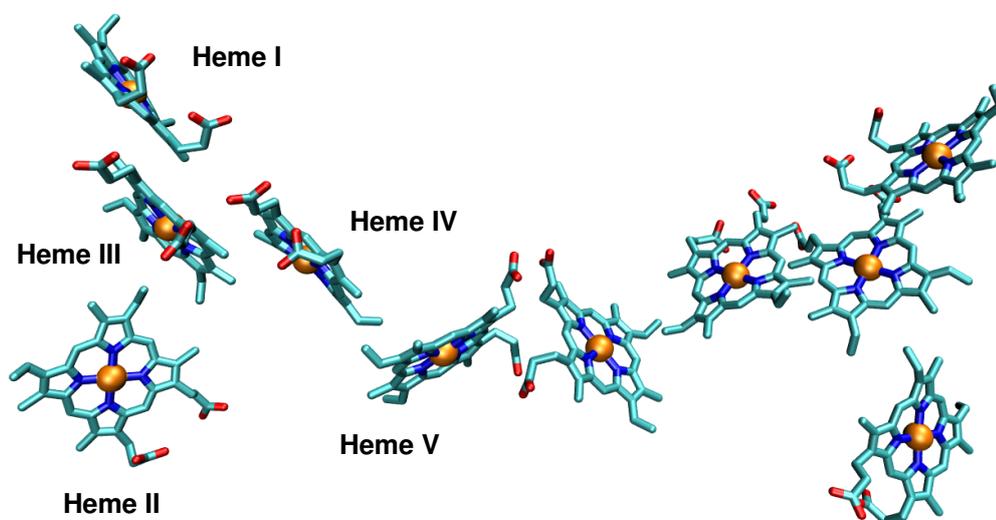
Cytochrome *c* nitrite reductase possesses more than 180 protonatable groups, of which ten of them are the heme propionates in each monomer. In total, cytochrome *c* nitrite reductase

contains 10 *c*-type hemes. In ten hemes, eight hemes are *c*-type heme i.e., two histidines coordinating to the iron and in other hemes the lysine is covalently linked as a proximal ligand to the heme iron of the *c*-type heme. Although all *c*-type hemes are of the same chemical nature, their midpoint potentials differ.

The cytochrome *c* nitrite reductase is a homodimeric enzyme with 10 *c*-type hemes which are arranged such that the nearest neighbors are in close proximity. It has been reported [164, 169] that the reduction potential of the 10 heme centers ranges from ca. -30 to -320 mV in cytochrome *c* nitrite reductase from *Escherichia coli*. The protein film voltammetric experiments revealed that the heme oxidation state has a profound, and often unanticipated effect on the interactions with substrate molecules, nitrite, hydroxyl amine and inhibitor cyanide [169]. The oxidation probabilities of the hemes and protonation probabilities of the heme propionates and redox potentials of the hemes in cytochrome *c* nitrite reductase are calculated to study the redox potential profiles. The electrostatic interactions between the protonatable and redox-groups are studied by solving the Poisson-Boltzmann equation using finite difference method. The electrostatic calculations are performed in cytochrome *c* nitrite reductase, a less complex protein before performing calculations on the complex protein like cytochrome *c* oxidase which contains four redox centers  $Cu_A$ , heme *a*, heme  $a_3$  and  $Cu_B$ . The aim of the present study is to obtain the redox mid-point potentials of the hemes to understand the redox potential profiles of hemes when the catalytic Heme I was blocked with the inhibitor cyanide.



**Figure 4.1. The nitrite reductase dimer and the heme arrangement.** A front view with the dimer axis orientated vertically, the five hemes in each monomer (stick model) and the  $Ca^{2+}$  (yellow) are shown.



**Figure 4.2.** The arrangement of hemes in nitrite reductase dimer is shown. The overall orientation of hemes corresponds to nitrite reductase dimer is shown. Hemes in the left monomer are labeled according to their attachment to the protein chain. The same labels are used throughout the study.

#### 4.1 HEMES AND THE CALCIUM BINDING SITE IN CYTOCHROME *c* NITRITE REDUCTASE

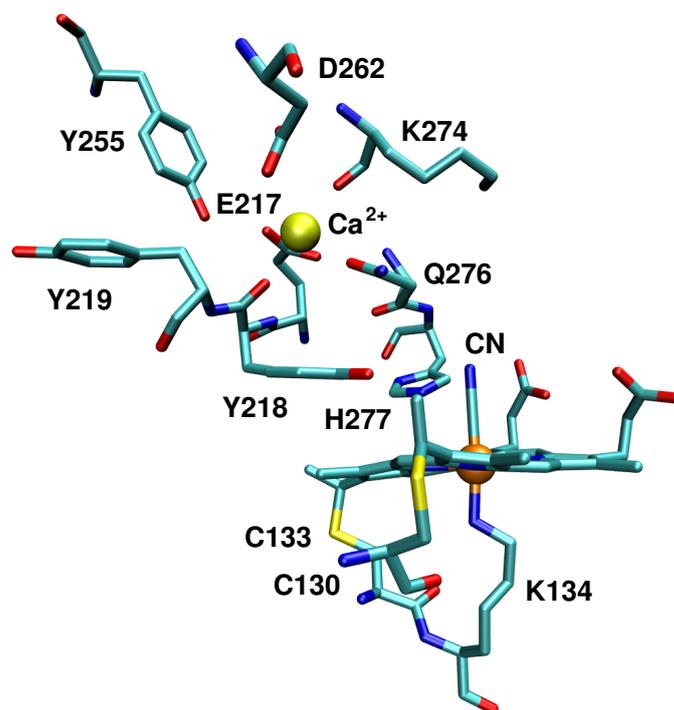
Hemes I, III and IV are close enough to allow direct  $\pi$ -electron interaction of the porphyrin rings. The propionate side chains of heme I form part of the active-site cavity, while those of Heme IV are exposed to the solvent and the Heme III propionates are hydrogen-bonded inside the protein. All the hemes show slight distortion from ideal planarity and it is more pronounced in Heme II and less in Heme I. It is suggested [162] that Heme II could function as the entry point for electrons (see Figure 4.2). The  $\text{Ca}^{2+}$  binding site appears to be an essential structural feature in the overall architecture of the enzyme and the region surrounding the calcium binding site is one of the most highly conserved parts of the whole sequence [163]. The calcium binding site of cytochrome *c* nitrite reductase with the important residues are shown in Figure 4.3. The residues Tyr281, Lys274, Gln276 and His277 act as ligands for the calcium binding site. It was proposed that the calcium binding site hold the key residues needed for catalysis [162]. The set of tyrosine residues near the calcium binding site and active site might play a role in the reaction mechanism of cytochrome *c* nitrite reductase by forming possible radical intermediates of the stepwise reduction of nitrite to ammonia [170].

## 4.2 PREPARATION OF THE CRYSTAL STRUCTURES OF CYTOCHROME *c* NITRITE REDUCTASE FOR CONTINUUM ELECTROSTATICS CALCULATIONS

The X-ray structure of cytochrome *c* nitrite reductase (PDB-code 1FS9, 1.6 Å resolution) from  $\epsilon$ -proteobacterium *Wolinella succinogenes* with azide ion as inhibitor was considered for the present study. To block the electron transfer from catalytic heme to inhibitor, the azide ion was replaced with cyanide by `InsightII` [171] program package. The structures were prepared for electrostatic calculations using `CHARMM` [153] molecular modeling package. Hydrogen atom positions were generated using the `HBUILD` algorithm implemented in the `CHARMM` program. All the heavy atoms were fixed and the energy minimizations were performed using 500 steepest decent (SD) steps, followed by 500 conjugate gradient (GC) steps. Partial atomic charges for standard amino acids were taken from `CHARMM` parameter set. For the redox-center (Heme I to V) and the non-redox centers ( $\text{Ca}^{2+}$ ) partial charges were obtained from quantum chemical calculations (see section 4.2.1). The protein structures were minimized with the crystallographic water molecules. In electrostatic calculations the crystallographic water molecules were removed since the construction of water hydrogens would arbitrarily assign a certain orientation to the water molecules that would affect electrostatic calculations. The intrinsic  $\text{p}K_a$  and the redox potentials were calculated by solving the Poisson-Boltzmann equation using the `QMPB` program. All aspartates, glutamates, histidines, lysines, argenines, tyrosines, the N-terminus and C-terminus of the protein and the heme propionates were considered as protonatable groups. The  $\text{p}K_a$  values of the model compounds were taken from the literature [70, 71].

### 4.2.1 DENSITY FUNCTIONAL CALCULATIONS

The partial charges of the hemes and the residues linked with the calcium ion were obtained by DFT [81, 172] calculations. The X-ray crystal structure of the heme I with inhibitor cyanide and *c*-type heme are shown in Figure 4.4. The histidines of *c*-type heme were modeled by methyl-imidazoles and the cysteines were modeled by methyl-thiols. In catalytic Heme I, cysteines were modeled by methyl-thiols and the lysine next to the  $\text{C}_\delta$  carbon was cut and hydrogens were added. The propionates of the hemes are replaced with methyl groups. The model compounds considered for heme I and heme (II-V) are shown in Figure 4.5. The density functional calculations were performed on both heme models to obtain partial atomic charges both in reduced and oxidized states. The DFT calculations were performed the Amsterdam density functional program (ADF 2004.01) [173]. In the local density approximation for the exchange and correlation, the Vosko-Wilk-Nusair [84] local spin density potential was used. For the generalized gradient approximation the Parr-Wang 91 exchange and correlation functionals [86] were used. Based on the spin states of the redox centers, the spin-unrestricted calculations were done with double- $\zeta$  STO basis sets. All the core densities were kept frozen and orthogonal to the valence orbitals. The  $\text{C}_\beta$  of the cysteines and histidines,  $\text{C}_\delta$  carbon of lysine and the C atom of methyl groups representing propionates were fixed during minimization. Both the hemes, heme-lys (Heme I) and heme *c* (Hemes II-V) were considered in the low-spin state. The reduced *c*-type heme has total charge of +1 and S=0 and the oxidized



**Figure 4.3. The active site of cytochrome *c* nitrite reductase with important residues.** The catalytic heme I bound with the inhibitor cyanide and the calcium binding site with important residues are shown.

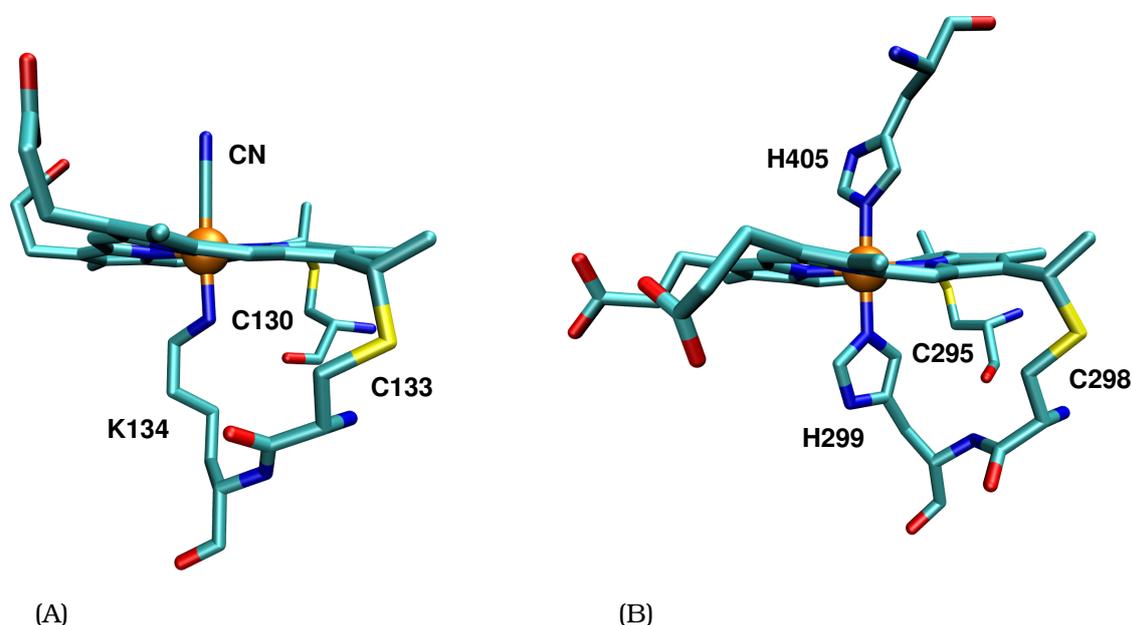
hemes *c*-type heme has total charge of 0 and  $S=1/2$ . The reduced heme-lys has total charge of 0 and  $S=0$  and the oxidized heme-lys has total charge of -1 and  $S=1/2$ .

#### CHARGE FITTING

Point charges were obtained by fitting the quantum electrostatic potential using a CHELPG algorithm [174]. The CHELPG algorithm was explained in detail in section 3.1.2. The ESP charges were calculated on a grid of uniform spacing of 0.2 Å with 3 Å as outer boundary cutoff around each atom of the molecule. The atoms were assigned the Bondi radii [175] values of 1.7 for carbon, 1.2 for hydrogen, 1.55 for nitrogen, 1.5 for oxygen, 1.3 for iron and 1.8 for sulfur. Total charge and the dipole moments were applied as constraints. The charges obtained for oxidized Fe(III) and reduced Fe(II) heme with lysine and cyanide and heme *c* are given in Table 4.1 and 4.2 respectively.

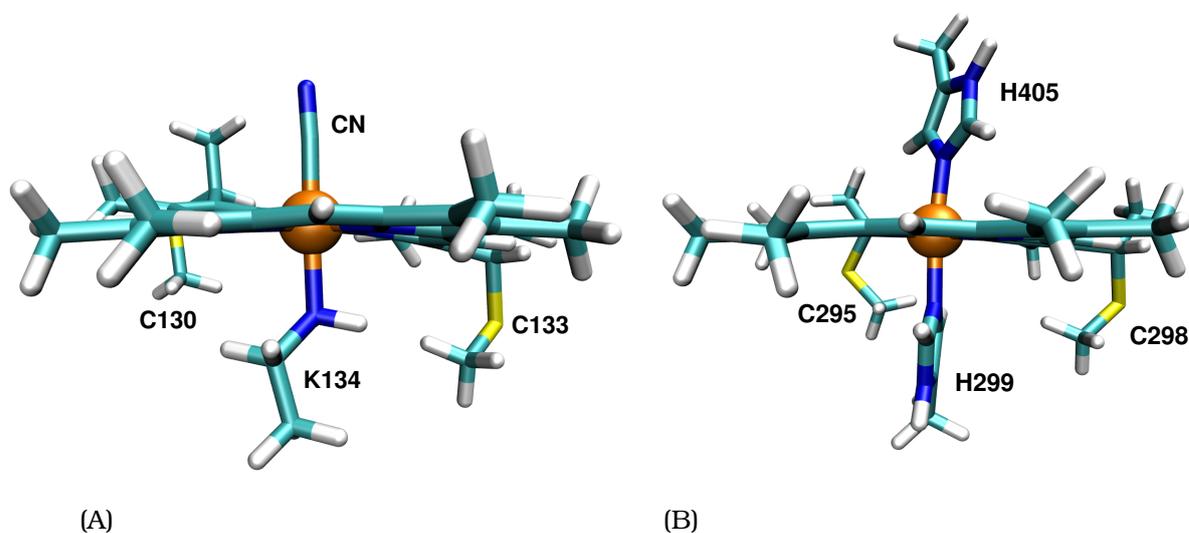
#### 4.2.2 CONTINUUM ELECTROSTATIC CALCULATIONS

Continuum electrostatic calculations were performed with the QMPB (Quantum Mechanical based Poisson-Boltzmann method) program which numerically solves the linear Poisson-Boltzmann equation by finite difference method. The QMPB program is developed in our group



**Figure 4.4. The X-ray crystal structure of catalytic heme (Heme I) and other *c*-type heme of cytochrome *c* nitrite reductase from *W. succinogenes*.** (A). The catalytic Heme I with lysine and cyanide are shown. (B). The other *c*-type heme with two histidines residues coordinating to the iron are shown. All the four hemes from II to V belongs to the *c*-type heme. Only Heme II is shown in this figure.

by Timm Essigke. The monomer of cytochrome *c* nitrite reductase was considered for the electrostatic calculations. The protein structure for electrostatic calculations were obtained by the minimization procedure described in section 4.2. The ionic strength of the protein was set to 0.1 M and the dielectric constant of protein and water were set to 4.0 and 80.0 respectively. The linear Poisson-Boltzmann equation was solved by using two successively finer grids with grid spacing of 1.0 and 0.25 Å and grid size of 121<sup>3</sup>. Initially the grid was centered on the protein model compound and the second grid was centered on the titratable group. The intrinsic  $pK_a$  values were obtained by Eq. (2.34). The Monte Carlo program GMCT developed by our group was used to calculate the titration curves of amino acids and hemes within the proteins. Aspartates, glutamates, lysine, histidines, cysteine, tyrosine and N- and C-termini were treated as titratable groups. The redox centers, Heme II to Heme IV were treated as redox-active sites. The redox potential of -220 mV [176] was used as model redox-potential for *c*-type hemes. The catalytic Heme I was fixed in either in the oxidized or in the reduced state. The mid-point potential of *c*-type hemes were calculated both in the reduced and in the oxidized state of catalytic Heme I. The protonation and oxidation probabilities were computed in the pH range of 0.0 to 12.0, in steps of 0.1 pH units and the redox potential in the range of -600 to 600 mV in steps of 10 mV. The mid-point potentials of Heme II to V were calculated at pH 7.



**Figure 4.5. The model compound of heme-lys (Heme I) and heme *c* (Hemes II-V).** (A). The model compound of heme-lys (Heme I) is shown. The lysine coordinating the iron is truncated next to the C $\delta$  carbon and two cysteines are modeled by methyl-thiols. The azide ion is replaced with cyanide. (B) The model compound of heme *c* is shown. The two histidines coordinating the iron are modeled by methyl imidazole and cysteines are modeled by methyl-thiols. The heme propionates were cut off and substituted by hydrogen atoms.

### 4.3 RESULTS AND DISCUSSION

All *c*-type hemes (Heme II-V) are considered as redox active groups. The oxidation probabilities of Hemes II-V from *W. succinogenes* at pH=7.0 are given in Figure 4.6. The oxidation probabilities of the Heme II-V at pH 0.0 to 12.0 are given in Figure 4.7. The protonation probability of the Heme I propionates and Heme II-V propionates at pH 0.0 to 12.0 are given in Figure 4.8 and 4.10 respectively.

Butt and co-workers performed protein film voltammetry experiments on cytochrome *c* nitrite reductase from *Escherichia coli* and studied the effects of two inhibitors, azide and cyanide anion on the activity of the enzyme [177]. They measured the redox-potential of hemes by treating the enzyme with nitrite inhibitor and obtained the redox potential,  $E_{1/2}^{\circ} = -103$  mV for Heme I and III, and Heme II was assigned  $E_{1/2}^{\circ} = \sim -37$  mV. The Heme IV and V were assigned  $E_{1/2}^{\circ} = \sim -320$  mV. The assignment of the mid-point potential to a particular heme is, however very controversial [169]. They also studied the affinity of cyanide on the reduced and oxidized states of cytochrome *c* nitrite reductase. The order of assignment of redox-potential is challenging because some protonatable groups in the protein can change the protonations and conformations upon reduction as well as in the presence of paramagnetic centers. The coupling between protonation and reduction reactions has been observed in many biological charge transfer reactions and are known as redox Bohr effect. It is caused by the electrostatic interactions between protonatable and redox-active groups [178]. The oxidation probabilities of Heme II-V and the protonation probabilities of Hemes (I-V) propionates are discussed in the following section.

| Residue | Atom | Charges |         | Residue | Atom | Charges |         |
|---------|------|---------|---------|---------|------|---------|---------|
|         |      | Oxa.    | Red.    |         |      | Oxa.    | Red.    |
| HEME I  | FE   | 0.8370  | 0.7190  | HEME I  | HAB1 | -0.0310 | -0.0750 |
| HEME I  | NA   | -0.0880 | -0.0540 | HEME I  | CBB  | -0.3110 | -0.3290 |
| HEME I  | NB   | -0.1860 | -0.0820 | HEME I  | HXB1 | 0.0970  | 0.0833  |
| HEME I  | NC   | -0.1630 | -0.1370 | HEME I  | HXB2 | 0.0970  | 0.0833  |
| HEME I  | ND   | -0.2950 | -0.2290 | HEME I  | HXB3 | 0.0970  | 0.0833  |
| HEME I  | C1A  | 0.0810  | 0.0300  | HEME I  | CMC  | -0.2560 | -0.2420 |
| HEME I  | C2A  | -0.2250 | -0.2080 | HEME I  | HMC1 | 0.0970  | 0.0717  |
| HEME I  | C3A  | 0.0920  | 0.0500  | HEME I  | HMC2 | 0.0970  | 0.0717  |
| HEME I  | C4A  | -0.0720 | -0.1040 | HEME I  | HMC3 | 0.0970  | 0.0717  |
| HEME I  | C1B  | -0.0690 | -0.1730 | HEME I  | CAC  | 0.4170  | 0.6130  |
| HEME I  | C2B  | 0.0490  | 0.0890  | HEME I  | HAC1 | -0.0490 | -0.1090 |
| HEME I  | C3B  | -0.1600 | -0.2210 | HEME I  | CBC  | -0.2360 | -0.2680 |
| HEME I  | C4B  | 0.0880  | 0.0480  | HEME I  | HYC1 | 0.0780  | 0.0650  |
| HEME I  | C1C  | 0.0770  | 0.0290  | HEME I  | HYC2 | 0.0780  | 0.0650  |
| HEME I  | C2C  | -0.0120 | 0.0470  | HEME I  | HYC3 | 0.0780  | 0.0650  |
| HEME I  | C3C  | -0.2310 | -0.3590 | HEME I  | CMD  | -0.4080 | -0.3570 |
| HEME I  | C4C  | 0.1300  | 0.1050  | HEME I  | HMD1 | 0.1407  | 0.1117  |
| HEME I  | C1D  | 0.0990  | 0.0400  | HEME I  | HMD2 | 0.1407  | 0.1117  |
| HEME I  | C2D  | 0.0570  | 0.0410  | HEME I  | HMD3 | 0.1407  | 0.1117  |
| HEME I  | C3D  | -0.2670 | -0.2610 | HEME I  | CAD  | 0.1790  | 0.1360  |
| HEME I  | C4D  | 0.1950  | 0.1340  | HEME I  | HAD1 | -0.0030 | -0.0060 |
| HEME I  | CHA  | -0.3660 | -0.3500 | HEME I  | HAD2 | -0.0030 | -0.0060 |
| HEME I  | HA   | 0.2540  | 0.2390  | CYS130  | SG   | -0.3510 | -0.4270 |
| HEME I  | CHB  | -0.1410 | -0.1290 | CYS130  | CB   | 0.3390  | 0.2930  |
| HEME I  | HB   | 0.1600  | 0.1570  | CYS130  | HB1  | -0.0840 | -0.0740 |
| HEME I  | CHC  | -0.3480 | -0.3850 | CYS130  | HB2  | -0.0840 | -0.0740 |
| HEME I  | HC   | 0.2420  | 0.2450  | CYS133  | SG   | -0.3780 | -0.4670 |
| HEME I  | CHD  | -0.3950 | -0.3910 | CYS133  | CB   | 0.3010  | 0.2560  |
| HEME I  | HD   | 0.2730  | 0.2740  | CYS133  | HB2  | -0.0695 | -0.0505 |
| HEME I  | CMA  | -0.3840 | -0.3420 | CYS133  | HB1  | -0.0695 | -0.0505 |
| HEME I  | HMA1 | 0.1380  | 0.1107  | LYS134  | CD   | 0.0490  | -0.0230 |
| HEME I  | HMA2 | 0.1380  | 0.1107  | LYS134  | HD1  | -0.0115 | -0.0090 |
| HEME I  | HMA3 | 0.1380  | 0.1107  | LYS134  | HD2  | -0.0115 | -0.0090 |
| HEME I  | CAA  | 0.1880  | 0.1280  | LYS134  | CE   | 0.0790  | 0.1440  |
| HEME I  | HAA1 | -0.0070 | -0.0075 | LYS134  | HE1  | -0.0020 | -0.0180 |
| HEME I  | HAA2 | -0.0070 | -0.0075 | LYS134  | HE2  | 0.0140  | -0.0090 |
| HEME I  | CMB  | -0.1080 | -0.1170 | LYS134  | NZ   | -0.4780 | -0.4760 |
| HEME I  | HMB1 | 0.0527  | 0.0397  | LYS134  | HZ1  | 0.2400  | 0.2280  |
| HEME I  | HMB2 | 0.0527  | 0.0397  | LYS134  | HZ2  | 0.2180  | 0.2030  |
| HEME I  | HMB3 | 0.0527  | 0.0397  | CN      | C    | -0.0230 | -0.0180 |
| HEME I  | CAB  | 0.3710  | 0.5310  | CN      | N    | -0.4560 | -0.5710 |

**Table 4.1. Charges of the oxidized Fe(III) and reduced Fe(II) heme *c* coordinated with lysine and cyanide.** The charges are obtained from chargefit program.

#### OXIDATION PROBABILITIES OF HEME II-V

The mid-point redox potentials obtained for Hemes II-V in the reduced and in the oxidized states are given in Table 4.3. There is no proper model compound for catalytic Heme I. So the redox state of the catalytic Heme I was fixed either in reduced or in oxidized state and the mid-point potential of the other hemes were calculated.

The experimentally determined midpoint potentials of Heme II, Heme III, Heme IV and Heme V are -37, -107, -323, -323 mV respectively [169]. Based on these redox potentials, Butt et al. proposed the following mechanism for the six electron reduction of nitrite to ammonia by cytochrome *c* nitrite reductase. Initially the activity of the enzyme is turned on by the reduction of Hemes I and III where the catalytic Heme I is involved in the nitrite coordination. Heme

| Residue | Atom | Charges |         | Residue | Atom | Charges |         |
|---------|------|---------|---------|---------|------|---------|---------|
|         |      | Oxa.    | Red.    |         |      | Oxa.    | Red.    |
| HEME C  | NA   | -0.3120 | -0.2580 | HEME C  | HMC1 | 0.0843  | 0.0597  |
| HEME C  | NB   | -0.2290 | -0.1240 | HEME C  | HMC2 | 0.0843  | 0.0597  |
| HEME C  | NC   | -0.1680 | -0.1120 | HEME C  | HMC3 | 0.0843  | 0.0597  |
| HEME C  | ND   | -0.2380 | -0.1770 | HEME C  | CAC  | 0.3320  | 0.5290  |
| HEME C  | C1A  | 0.1760  | 0.1220  | HEME C  | HAC1 | 0.0040  | -0.0570 |
| HEME C  | C2A  | -0.2520 | -0.234  | HEME C  | CBC  | -0.4060 | -0.4440 |
| HEME C  | C3A  | 0.0140  | -0.0100 | HEME C  | HYC1 | 0.1300  | 0.1197  |
| HEME C  | C4A  | 0.0950  | 0.0350  | HEME C  | HYC2 | 0.1300  | 0.1197  |
| HEME C  | C1B  | 0.0330  | -0.0460 | HEME C  | HYC3 | 0.1300  | 0.1197  |
| HEME C  | C2B  | 0.0690  | 0.0870  | HEME C  | CMD  | -0.3610 | -0.3210 |
| HEME C  | C3B  | -0.2550 | -0.3310 | HEME C  | HMD1 | 0.1360  | 0.1100  |
| HEME C  | C4B  | 0.3000  | 0.2820  | HEME C  | HMD2 | 0.1360  | 0.1100  |
| HEME C  | C1C  | 0.1310  | 0.1190  | HEME C  | HMD3 | 0.1360  | 0.1100  |
| HEME C  | C2C  | -0.0510 | -0.0560 | HEME C  | CAD  | 0.1020  | 0.0560  |
| HEME C  | C3C  | -0.0790 | -0.2010 | HEME C  | HAD1 | 0.0420  | 0.0365  |
| HEME C  | C4C  | -0.0600 | -0.0720 | HEME C  | HAD2 | 0.0420  | 0.0365  |
| HEME C  | C1D  | -0.0370 | -0.0910 | CYS     | SG   | -0.3510 | -0.4130 |
| HEME C  | C2D  | 0.0900  | 0.0820  | CYS     | CB   | 0.3640  | 0.3210  |
| HEME C  | C3D  | -0.2750 | -0.2540 | CYS     | HB1  | -0.0815 | -0.0640 |
| HEME C  | C4D  | 0.2130  | 0.1420  | CYS     | HB2  | -0.0815 | -0.0640 |
| HEME C  | CHA  | -0.3830 | -0.3650 | CYS     | SG   | -0.3390 | -0.4260 |
| HEME C  | HA   | 0.2310  | 0.2240  | CYS     | CB   | 0.3310  | 0.2870  |
| HEME C  | CHB  | -0.3140 | -0.2920 | CYS     | HB2  | -0.0710 | -0.0500 |
| HEME C  | HB   | 0.2430  | 0.2330  | CYS     | HB1  | -0.0710 | -0.0500 |
| HEME C  | CHC  | -0.4940 | -0.5320 | HIS     | CB   | 0.1530  | 0.1140  |
| HEME C  | HC   | 0.2740  | 0.2810  | HIS     | HB1  | 0.0215  | 0.0135  |
| HEME C  | CHD  | -0.2110 | -0.2060 | HIS     | HB2  | 0.0215  | 0.0135  |
| HEME C  | HD   | 0.2160  | 0.2150  | HIS     | ND1  | -0.3020 | -0.3470 |
| HEME C  | CMA  | -0.3410 | -0.2860 | HIS     | HD1  | 0.3710  | 0.3610  |
| HEME C  | HMA1 | 0.1413  | 0.1100  | HIS     | CG   | -0.0620 | -0.0230 |
| HEME C  | HMA2 | 0.1413  | 0.1100  | HIS     | NE2  | -0.1540 | -0.1260 |
| HEME C  | HMA3 | 0.1413  | 0.1100  | HIS     | CD2  | -0.1920 | -0.2540 |
| HEME C  | CAA  | 0.1960  | 0.1530  | HIS     | HD2  | 0.2110  | 0.2120  |
| HEME C  | HAA1 | 0.0085  | 0.0005  | HIS     | CE1  | -0.0260 | -0.0270 |
| HEME C  | HAA2 | 0.0085  | 0.0005  | HIS     | HE1  | 0.1620  | 0.1530  |
| HEME C  | CMB  | -0.4610 | -0.4190 | HIS     | CB   | 0.1830  | 0.1350  |
| HEME C  | HMB1 | 0.1580  | 0.1307  | HIS     | HB1  | 0.0145  | 0.0085  |
| HEME C  | HMB2 | 0.1580  | 0.1307  | HIS     | HB2  | 0.0145  | 0.0085  |
| HEME C  | HMB3 | 0.1580  | 0.1307  | HIS     | ND1  | -0.3560 | -0.4170 |
| HEME C  | CAB  | 0.3240  | 0.4560  | HIS     | HD1  | 0.3880  | 0.3830  |
| HEME C  | HAB1 | -0.0160 | -0.0520 | HIS     | CG   | -0.0520 | -0.0050 |
| HEME C  | CBB  | -0.2880 | -0.3290 | HIS     | NE2  | -0.2230 | -0.2120 |
| HEME C  | HXB1 | 0.1023  | 0.0977  | HIS     | CD2  | -0.1420 | -0.1760 |
| HEME C  | HXB2 | 0.1023  | 0.0977  | HIS     | HD2  | 0.1460  | 0.1400  |
| HEME C  | HXB3 | 0.1023  | 0.0977  | HIS     | CE1  | 0.0240  | 0.0260  |
| HEME C  | CMC  | -0.1850 | -0.1510 | HIS     | HE1  | 0.1530  | 0.1470  |
| HEME C  | FE   | 0.9620  | 0.7780  |         |      |         |         |

**Table 4.2. Charges of the oxidized Fe(III) and reduced (II) heme c.** The charges are obtained from chargefit program.

I, III and II in each monomer are reduced first and the dimer holds the six electrons required for the reduction of nitrite molecule bound in one active site to ammonium. The reduction of Hemes IV and V is a step in the process of transformation of nitrite to hydroxylamine [169].

In proteins, the protonation or oxidation probability of more than one titratable group usually depends on the pH or redox potential of the solution. Consequently, the interaction between these titratable groups also depends on the pH or redox potential of the solution. Because

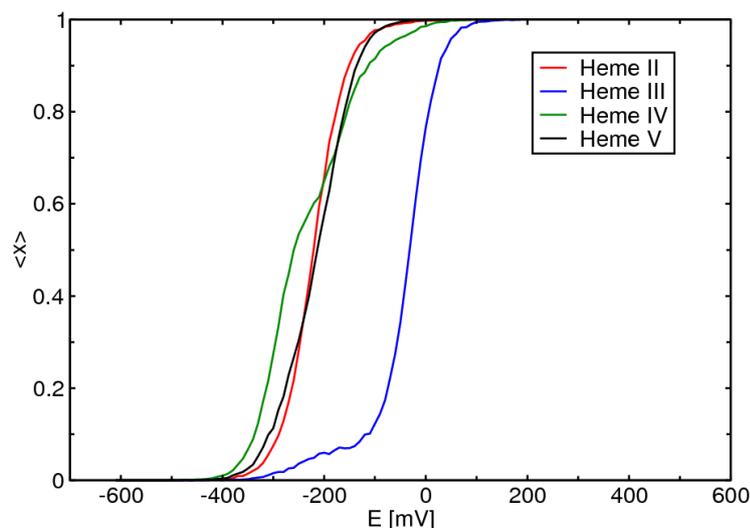
| Hemes    | Mid-point redox potentials |                   |                   |
|----------|----------------------------|-------------------|-------------------|
|          | in [mV]                    |                   |                   |
|          | Oxd. <sup>a</sup>          | Red. <sup>b</sup> | Exp. <sup>c</sup> |
| Heme II  | -210                       | -210              | -323 (Heme IV)    |
| Heme III | -30                        | -50               | -37 (Heme II)     |
| Heme IV  | -270                       | -270              | -107 (Heme III)   |
| Heme V   | -210                       | -210              | -323 (Heme V)     |

**Table 4.3. The calculated mid-point redox potentials of Hemes II to V at pH=7.** The mid-point redox potentials obtained for the Hemes II-V in the reduced and oxidized states of Heme I are given. The experimental mid-point redox potentials are arranged where the experimental values match closely with the calculated mid-point potentials. The order of assignment of redox-potential to a particular heme is very controversial [169].

<sup>a</sup> The mid-point redox potentials were calculated for Hemes II-V by fixing the catalytic Heme I in the oxidised state.

<sup>b</sup> The mid-point redox potentials were calculated for Hemes II-V by fixing the catalytic Heme I in the reduced state.

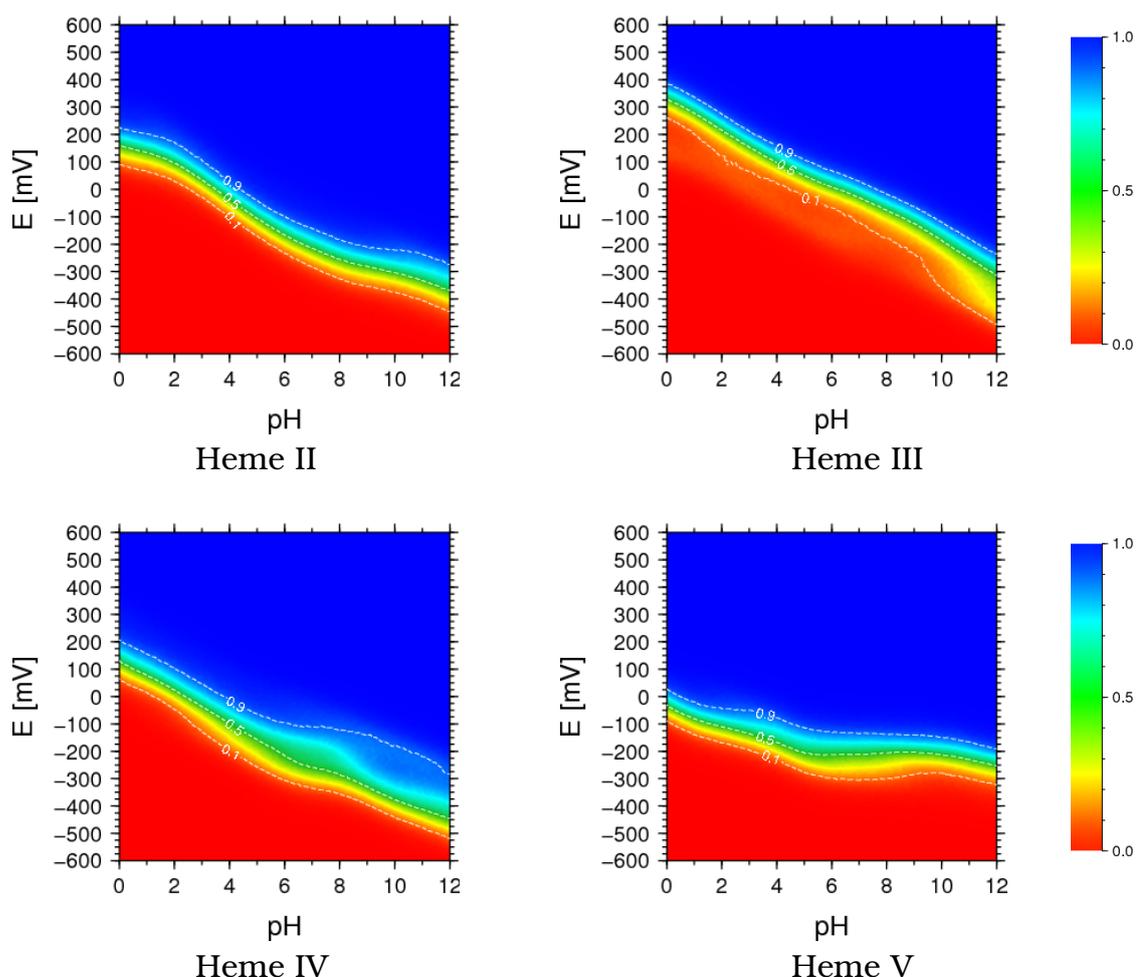
<sup>c</sup> The experimental mid-point redox potentials of Hemes II-V from Ref [169].



**Figure 4.6. Oxidation probabilities of different hemes in cytochrome *c* nitrite reductase from *W. succinogenes* at pH=7.** The hemes do not show a standard Nernst titration behavior because of the interaction with each other and with the titratable groups in their vicinity.

of these interactions, titration curves of titratable groups in proteins can deviate considerably from sigmoidal Nernst or Henderson-Hasselbalch titration curves of isolated titratable groups. Thus, it is not always possible to assign a unique mid-point potential or  $pK_a$  value to a specific titratable group. Therefore, the pH value at which the protonation probability of the protonatable group is 0.5 is often used to describe the titration behavior. Likewise, the

solution redox potential  $E_{1/2}^{\circ}$  at which the oxidation probability of the redox-active group is 0.5 is used to describe the titration behavior. The oxidation probabilities of different hemes in cytochrome *c* nitrite reductase at pH=7 are given in Figure 4.6. The titration curves deviates from the standard Nernst titration curves. This deviation is expected because of the interaction between the hemes and the interaction of hemes with other protonatable groups and heme propionates.



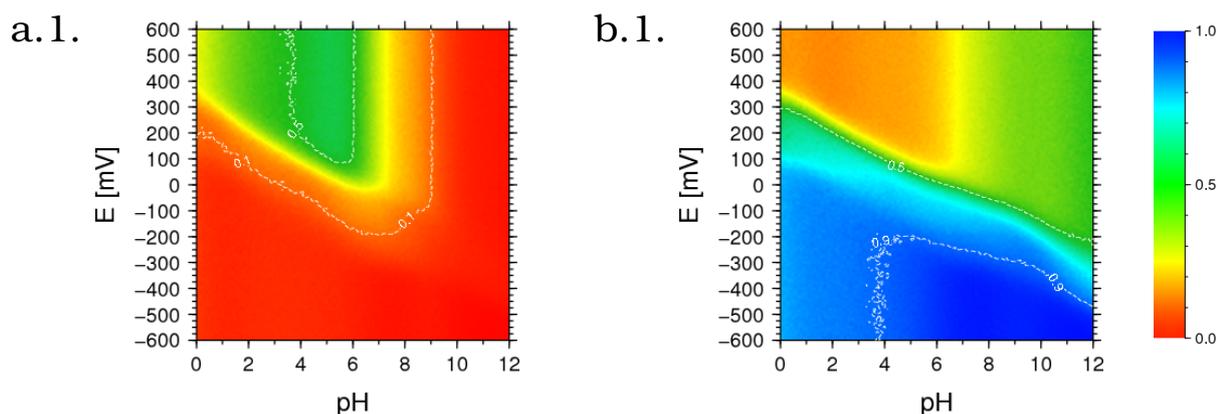
**Figure 4.7. Oxidation probabilities of the Hemes II-V.** Calculated oxidation probabilities of heme II-V of cytochrome *c* nitrite reductase from *W. succinogenes*. The probabilities are color coded as indicated by the scale next to the diagram. The green color indicates that the oxidation probability of the hemes is about 0.5.

The computed redox potentials for hemes range from  $-270$  to  $-30$  mV. However the mid-points cannot be compared with the experimental mid-point potential because the experimental mid-point potentials were obtained using nitrite as inhibitor. The redox potentials in the present studies were computed by blocking the catalytic Heme I by a cyanide where the blocking prevents the electron transfer from the catalytic heme to the inhibitor. Butt et al. [169] obtained redox potential of  $-107$  mV for Heme III. The redox potential of  $-30$  mV was obtained for Heme III from present calculations. On comparing the redox potentials of Hemes II-V, in

the oxidized and in the reduced state of Heme I, the redox potentials remain almost same except for Heme III where the difference is  $-20$  mV. Hemes II and V have same mid-point potential even though the distance between the hemes are more compare to heme IV and V. The Heme III has high redox potential compared to other hemes.

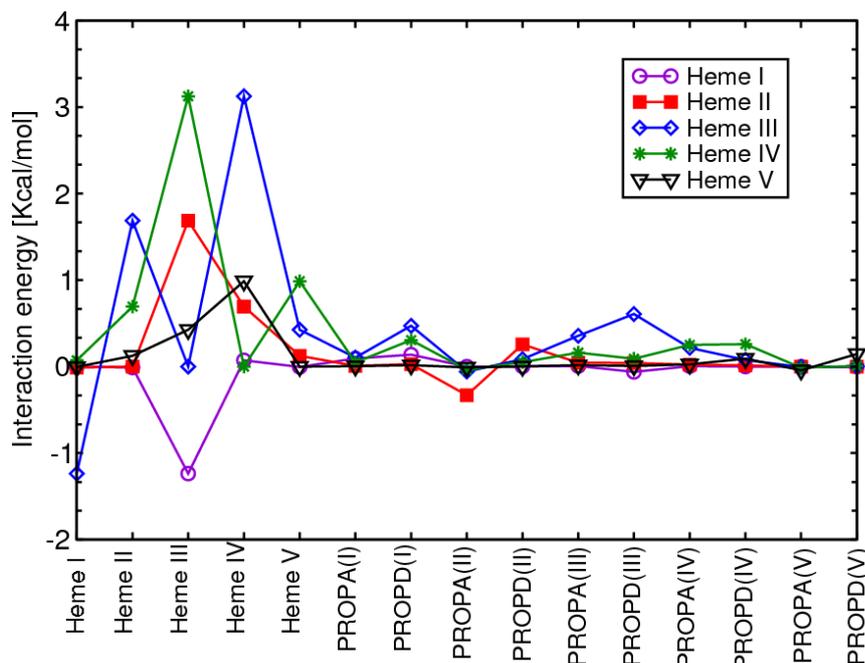
The oxidation probabilities of the Hemes II-V are shown in Figure 4.7. Within the protein, the heme mid-point redox potentials are affected by charges on the ionizable amino acids, polar groups and other hemes. The mid-point potential of the Hemes II-V decreases with the increase in pH. The oxidation probabilities of the Hemes II-V (see Figure 4.7) show that redox states of the hemes are strongly depend on the pH. The pH dependence is less pronounced in Hemes II and V which are not coplanar with other hemes and more pronounced in Heme III and IV.

#### PROTONATION PROBABILITIES OF HEMES I-V PROPIONATES



**Figure 4.8. Protonation probabilities of the hemes propionates A and D of Heme I.** (a.1)-(b.1). Calculated protonation probabilities of heme propionates A and D of Heme I. The probabilities are color coded as indicated by the scale next to the diagram. The green color indicates that the protonation probability of the residues is about 0.5.

The protonation probabilities of heme propionates (A and D) from Heme I and Heme II-V are shown in Figure 4.8 and 4.10 respectively. The interaction energies between the hemes and heme propionates are given in Figure 4.9. The protonation probabilities of Heme I propionates are strongly depend the redox states of other hemes. The protonation probabilities of Heme I propionate A (plot a.1) strongly depend on both the pH and redox potential compared to Heme I propionate D (plot b.1). Comparing the protonation probabilities of heme propionates A of Hemes II-V, Heme IV and Heme III propionates show strong dependence on both the pH and redox potential where as Heme II and V show dependence of the pH only. In heme propionates D of Hemes II-V, only Heme IV and V show strong dependence on pH and Heme II and III propionates are completely deprotonated and protonated respectively. The interaction between the heme charges depend on how deeply the hemes are buried in the protein as well as the distance between the hemes. The stabilization of the hemes by propionic acids can



**Figure 4.9. Interaction energies of Hemes (I-V) with heme propionates A and D in cytochrome *c* nitrite reductase.** The plot shows the interaction energies of Hemes (I-V) with heme propionates A and D in cytochrome *c* nitrite reductase.

also have influence on the redox potential of the hemes. The interaction energies between the hemes are higher compared to the heme propionates.

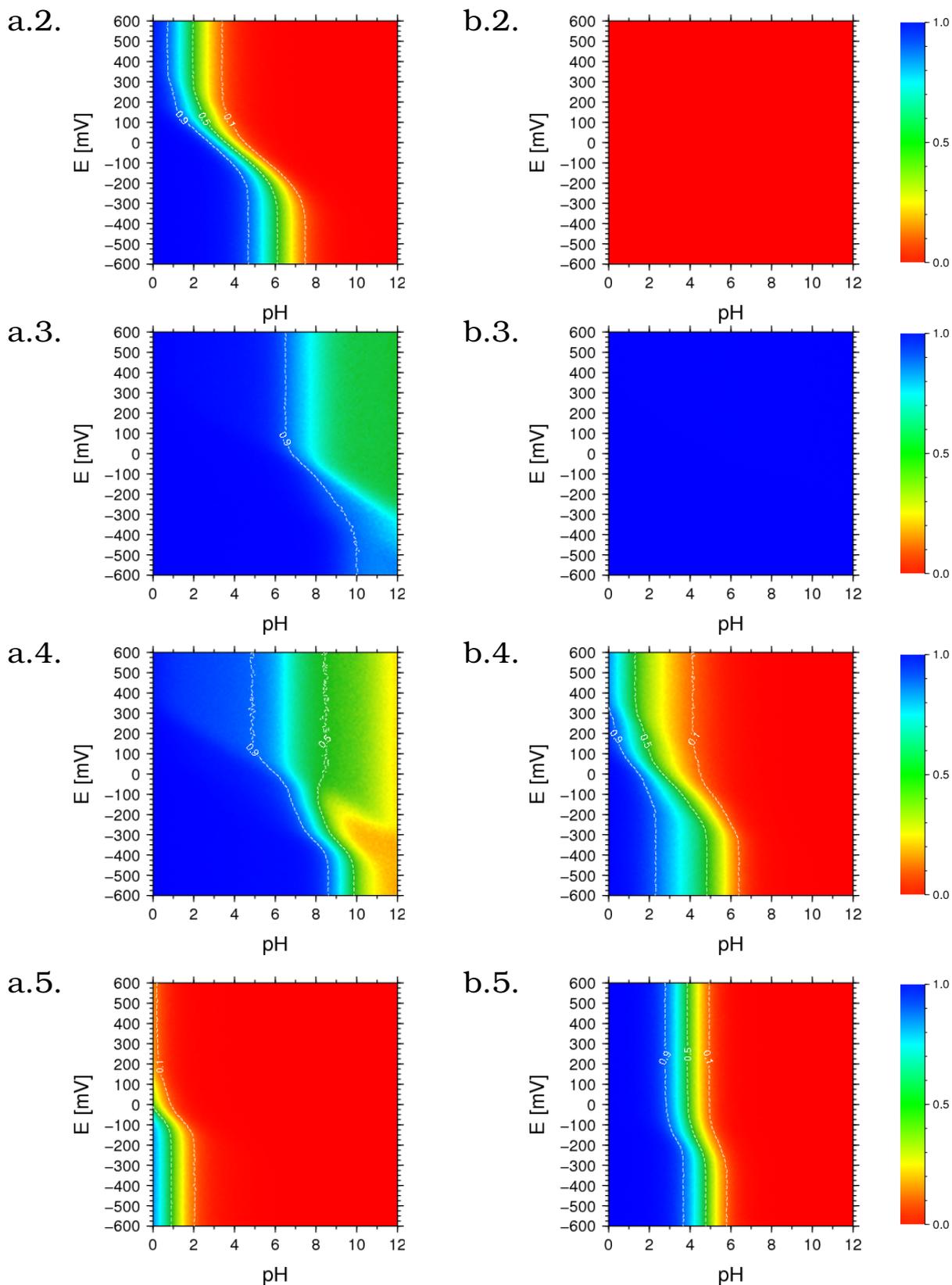
The interaction energies between Heme III and IV is around 3 kcal/mol due to the staking position of Heme III and IV with respect to Heme I and V. The analysis of the interaction between hemes show that they have influence on the heme mid-point potentials. The coupling of the  $pK_a$  shift of propionates with redox potentials of the hemes are supported from the structure because of the proximity of hemes to these propionates.

## 4.4 CONCLUSIONS

Electrostatic calculations were performed on a simpler, better resolved crystal structure of cytochrome *c* nitrite reductase before doing calculations on the complex system cytochrome *c* oxidase. The redox potentials were calculated and in the range between  $-270$  to  $-30$  mV. The redox potentials are almost same for the oxidized and reduced enzyme except for Heme III. The oxidation and protonation probabilities of the hemes and heme propionates were calculated. The redox potentials of the hemes are sensitive to their location and are also influenced by the protonation probabilities of the surrounding titratable groups and as well as the propionates in the protein. Heme I, III and IV are close enough to allow direct  $\pi$ -electron interactions of the porphyrin ring.

The mid-point potential of the Hemes II-V decreases with increase in pH. The oxidation probabilities of the Hemes II-V show less dependence on pH and Heme III and IV shows strong dependence. The oxidation probabilities show that the redox states of the hemes are strongly

depend on the pH. The protonation probabilities of Heme I propionates are strongly depend on the redox states of other hemes. On comparing the protonation probabilities of propionates A of Hemes II-V, propionates of Heme IV and Heme III show strong dependence on both the pH and redox potential and where as Heme II and V show dependence on the pH only. In heme propionates D of Hemes II-V, only Heme IV and V show strong dependence on pH and Heme II and III propionates are completely deprotonated and protonated respectively. The interaction between the hemes and heme propionates play a major role in tuning the redox potentials of the hemes.



**Figure 4.10. Protonation probabilities of the heme propionates A and D of Hemes II-V.** (a.2)-(a.5) Calculated protonation probabilities of heme propionates A of Hemes II-V. (b.2)-(b.5) Calculated protonation probabilities of heme propionates D of Hemes II-V. The probabilities are color coded as indicated by the scale next to the diagram. The green color indicates that the protonation probability of the residues is about 0.5.



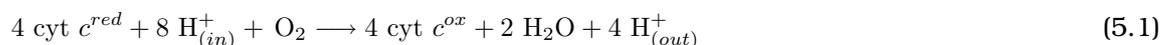
---

## $pK_a$ CALCULATIONS OF BINUCLEAR CENTER OF CYTOCHROME $c$ OXIDASE USING DFT CALCULATIONS

---

Knowledge of the  $pK_a$  values of ionizable groups are important for an understanding of many phenomena of chemistry of both the gas phase and the solution. The  $pK_a$  values are of particular interest in elucidating the reaction mechanisms, especially those involving proton transfers and interpreting the binding of substrates or inhibitors to enzymes. Experimental determinations of individual  $pK_a$  values are difficult in complex systems. One case in point concerns the direct measurement of  $pK_a$  values of titrating groups of catalytically important residue or substrate in the enzyme-substrate complexes [101].

Cytochrome  $c$  oxidase contains four redox centers  $Cu_A$ , heme  $a$ , heme  $a_3$  and  $Cu_B$ . It catalyzes the reduction of oxygen to water. In the course of its catalytic activity, electrons delivered by cytochrome  $c$  are transferred via  $Cu_A$  and heme  $a$  to the  $Fe_{a_3}$ - $Cu_B$  the binuclear center, where the reduction of oxygen takes place. For every molecular oxygen a total of eight protons are consumed from the matrix side of the mitochondrion: four protons are delivered to the binuclear center for water formation and another four protons are translocated across the membrane (see Eq. (5.1)).



A detailed picture of the underlying mechanism of how the protons are translocated to the binuclear center and how they reach the opposite side of the enzyme is a matter of debate [57, 179–181]. The analysis of the many available crystallographic structures of cytochrome  $c$  oxidase does not show any specific pathways connecting the D-channel to the binuclear center or leading protons to the exterior side of the membrane. Recent studies have proposed the direct involvement of the one of the  $Cu_B$  histidine ligands His291 in bovine heart or His334 in *R. sphaeroides*, a proton loading site in the pumping mechanism [1, 2, 8, 12].

Based on electrostatic calculations Stuchebrukhov et al. proposed that  $Cu_B$  bound imidazole ring of His334 play a key role in the proton pumping mechanism [1, 2]. The central feature of the proposed mechanism is that the  $pK_a$  values of the imidazole vary significantly depending on the redox state of the metals in the binuclear center [5, 6]. The proposed model states

that upon reduction of the binuclear center, His334 deprotonates at its  $N_{\delta 1}$  [5]. From DFT calculations on the models of  $Cu_B$  in aqueous solution, these authors found that the  $pK_a$  value of His334 is 8.6 for the oxidized  $Cu_B$  center and 13.2 for the reduced  $Cu_B$  center. Another DFT study [182] on the  $Cu_B$  center obtained a  $pK_a$  value of above 13.5 for His334 in the oxidized state making it unlikely that a redox-coupled protonation state change of His334 directly involved in the proton pumping mechanism. The discrepancy between these two studies lead to some debate in the field [4, 183]. However in the two studies [5, 182], the crystal coordinates were not fully relaxed leading to unrealistic bond lengths. Moreover in the later study [183], it was not considered that some residues of the cytochrome *c* oxidase can adopt a protonation that deviates from their usual protonation in aqueous solution. Siegbahn and co-workers, performed hybrid density functional theory calculations using B3LYP functional to study the energetics of the proton translocation in cytochrome *c* oxidase [184]. They also computed the redox potentials of the metal centers and the tyrosyl radicals, as well as  $pK_a$  values of important groups along the translocation paths [184, 185]. Ryde and co-workers systematically investigated how the axial ligand in heme proteins influences the geometry, electronic structure and spin states of the active site using density functional B3LYP method and medium-sized basis sets [186].

To investigate the reaction mechanism and as well as the role of the histidines bound with the  $Cu_B$  center and the histidine coordinating the heme  $a_3$ , we calculated the  $pK_a$  values of the Cu-bound imidazole in various  $Cu_B$  and  $Fea_3$  complexes using density functional theory in combination with classical continuum electrostatics models. In the previous studies only, the  $pK_a$  value of His334 was determined [5, 6, 182] and not all possible microscopic  $pK_a$  values were calculated in order to check for internal consistency. In the present study, we calculated all microscopic  $pK_a$  values of the  $Cu_B$  and heme  $a_3$  center both in the reduced and in the oxidized states [187]. The Finite Difference Poisson Boltzmann (FDPB) method and conductor-like polarizable continuum model (C-PCM) are used to determine the solvation free energies in aqueous solution. The dependence of the solvation free energies on the parameters are explored by comparing these two solvation models. The present study attempts to find reliable method to obtain proper  $pK_a$  values of the  $Cu_B$  center in aqueous solution.

## 5.1 $pK_a$ CALCULATIONS

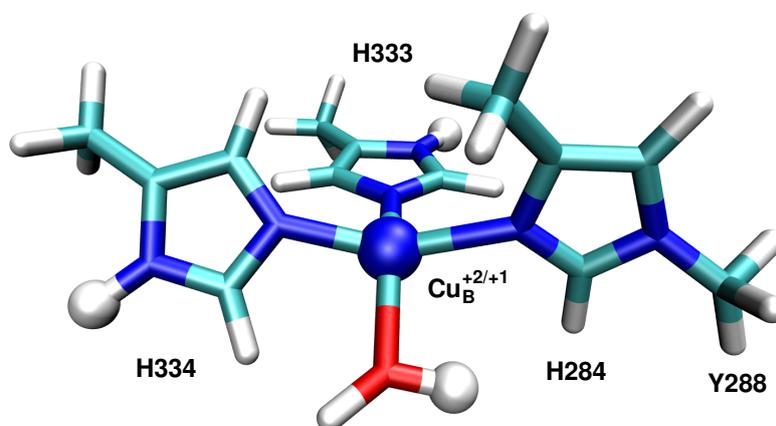
$pK_a$  calculations are used mainly in studies of enzyme mechanisms. Brief introduction of acid-base equilibria reactions are given in section 2.1.1 and 2.1.2 [62]. It is required to have reliable and accurate means of calculating relative and/or absolute  $pK_a$  values and proper understanding of the factors involved in  $pK_a$  calculations. The  $pK_a$  values are very sensitive to the charges used for the solvation free energy calculations. In this chapter, the  $pK_a$  calculations on  $Cu_B$  and heme  $a_3$  center in the aqueous solution are reported. The  $pK_a$  calculations in cytochrome *c* oxidase are reported and discussed in chapter 6. The accuracy of the  $pK_a$  calculations depend on different factors such as the solvation energies of the reactant and product, the solvation energy used for the proton and the model used for the  $pK_a$  calculations. The calculation of absolute  $pK_a$  values of species in solution is a delicate task, which necessitates the use of high-level QM methods capable of attaining chemical accuracy [63]. The treatment of solvation represents the least reliable aspects of  $pK_a$  calculations [182]. The heart of the  $pK_a$  estimation methods are the calculated atomic charges in case of FDPB

method. To gauge the dependence of  $pK_a$  values on the level of theory used to describe the solvation, FDPB method and C-PCM [121, 188] are used to determine the solvation free energies in bulk water.

## 5.2 COMPUTATIONAL METHODS

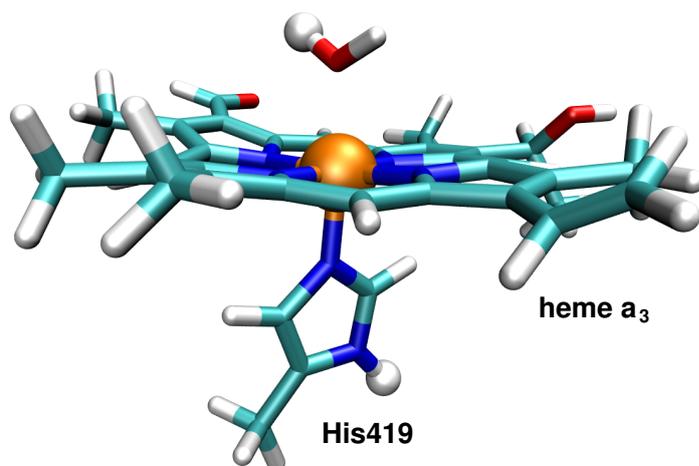
### 5.2.1 DENSITY FUNCTIONAL CALCULATIONS

The starting structures of the  $Cu_B$  and heme  $a_3$  centers were taken from the X-ray crystal structure of bovine heart cytochrome *c* oxidase obtained by Yoshikawa et al. at 2.3 Å resolution (2OCC in the protein data bank) [16]. The minimal model required to calculate the  $pK_a$  of a  $Cu_B$  center consists of the Cu ion, methylimidazole representing the coordinating histidines 284, 333 and 334, a methyl group representing Tyr288 (which is cross-linked to His284) and the bound  $H_2O$  molecule as  $Cu_B$  ligand. The methyl groups attached to the imidazole rings represent the  $C_\beta$  of histidines. A full description of Tyr288 is not included because the calculations are performed only to obtain the  $pK_a$  calculations of histidines and  $H_2O$  coordinating to  $Cu_B$  center. The model compound considered for the  $Cu_B$  center is given in Figure 5.1. The model for the heme  $a_3$  center consists of the heme  $a_3$  and an axial histidine which is modeled by methylimidazole. The sixth coordination is modeled with  $H_2O$  molecule as in the  $Cu_B$  center. The hydrophobic hydroxyethyl-farnesyl group is truncated next to the hydroxyl group. The model compound considered for the heme  $a_3$  center is given in Figure 5.2.



**Figure 5.1. The model compound of  $Cu_B$  center.** The histidines coordinating the copper ion are modeled by methyl imidazoles. The cross-linked tyrosine 288 is replaced by methyl group. The fourth coordination is the  $H_2O$  molecule. The proton binding sites are indicated by white spheres.

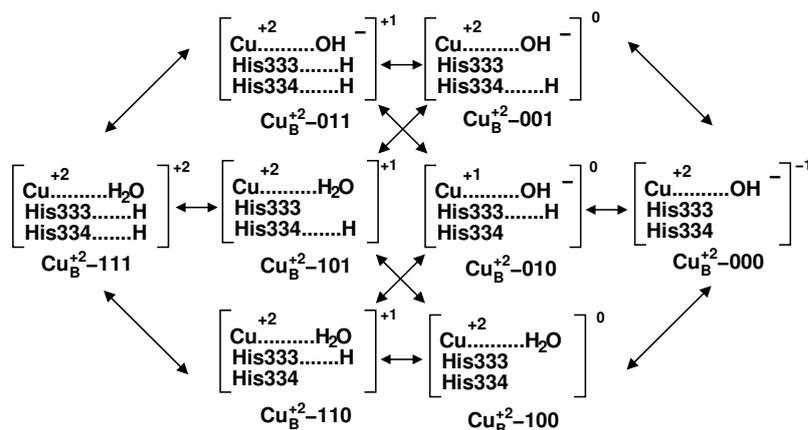
The  $Cu_B$  center has three proton binding sites which are the  $H_2O$  molecule,  $N_{\delta_1}$  of His333, and  $N_{\delta_1}$  of His334. The three proton binding sites lead to eight possible microscopic states and twelve proton equilibrium reactions [73]. In heme  $a_3$  center, there are two proton binding sites the  $H_2O$  molecule and the  $N_{\delta_1}$  of His419. The two proton binding sites lead to four possible microscopic states and four proton equilibrium reactions [73]. All possible protonation states



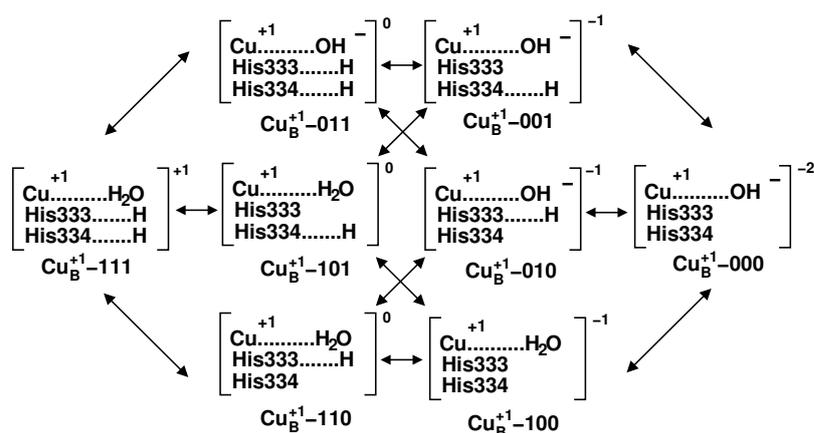
**Figure 5.2. The model compound of heme  $a_3$  center.** The histidine coordinating the iron are modeled by methyl imidazoles. The hydrophobic hydroxyethyl-farnesyl group was truncated next to the hydroxyl group. The heme propionates were cut off and substituted by hydrogen atoms. The  $H_2O$  molecule is considered in the six coordination position. The proton binding sites are indicated by white spheres.

of the  $Cu_B$  center and the possible proton equilibrium reactions in the oxidized and in the reduced states are given in Figure 5.3 and 5.4 respectively. All possible protonation states of the heme  $a_3$  center and the possible proton equilibrium reactions in the oxidized and in the reduced states are given in Figure 5.5 and 5.6 respectively. Each of these protonation states can be described by a protonation state vectors in which '1' indicates that the proton is bound and '0' indicates that the proton is not bound. The state vector of each protonation state of the complex is given at the bottom of the corresponding complex. For example the state vector '111' denotes that all the sites ( $H_2O$ , His333 and His334) are protonated and state vector '011' represents deprotonated  $H_2O$  and protonated His333 and His334.

All microscopic states were optimized with density functional theory (DFT) use of the Perdew-Wang exchange and correlation functional (PW91) and hybrid density functional theory (B3LYP). The geometry optimizations using PW91 were performed with ADF 2004.01 [173] program. In the local density approximation for the exchange and correlation, the Vosko-Wilk-Nusair [84] local spin density potential was used. For the generalized gradient approximation the Parr-Wang 91 exchange and correlation functionals [86] were used (see section 2.3.1). The numerical integration method was used with the general accuracy parameter 4.0. According to the redox states the spin-unrestricted calculations were done with high-level basis sets triple- $\zeta$  plus polarization, (TZP: triple- $\zeta$  plus polarization -ADF Basis IV) and triple- $\zeta$  plus polarization with diffusion function for iron and copper was used. All core densities were kept frozen and orthogonal to the valence orbitals. The frozen core includes the (1s, 2s, 2p) orbitals of Fe and Cu and the 1s orbitals of N, C, and O. The geometry optimizations using B3LYP method were performed using GAUSSIAN 03 [154] using triple- $\zeta$  valence plus polarization (TZVP), 6-31+G\* and 6-31G\* basis sets. Total of sixteen structures were optimized for  $Cu_B$  center corresponding to the eight protonation states in the oxidized and eight in the reduced state as given in



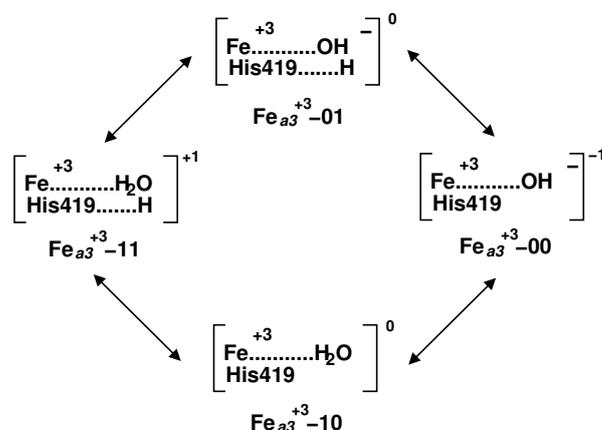
**Figure 5.3. Protonation equilibria of the  $\text{Cu}_B$  in the oxidized (+2) state.** Eight possible microscopic state and twelve possible proton equilibrium reactions of  $\text{Cu}_B^{+2}$  center are shown. Numbering of the ligands corresponds to the cytochrome *c* oxidase from *R. sphaeroides*. In the label,  $\text{Cu}_B^{+2}$  corresponds to the oxidized state of the  $\text{Cu}_B$  center. The protonation state of the corresponding  $\text{Cu}_B$  complex is indicated by protonation state vectors. The total charge of the complex is given in the top of the square bracket.



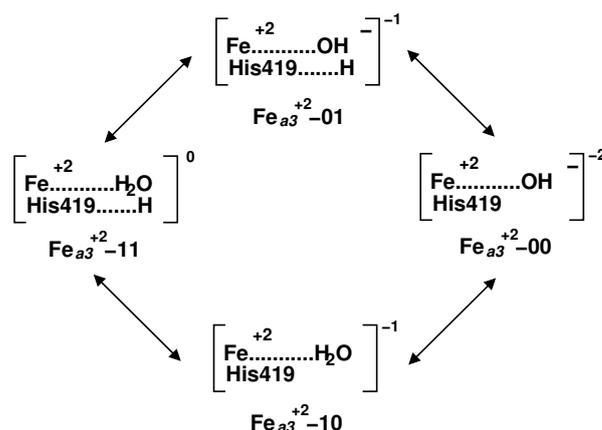
**Figure 5.4. Protonation equilibria of the  $\text{Cu}_B$  in the reduced (+1) state.** Eight possible microscopic state and twelve possible proton equilibrium reactions are shown. Numbering of the ligands corresponds to the cytochrome *c* oxidase from *R. sphaeroides*. In the label,  $\text{Cu}_B^{+1}$  corresponds to the reduced state of the  $\text{Cu}_B$  center. The protonation state of the corresponding  $\text{Cu}_B$  complex is indicated by protonation state vectors. The total charge of the complex is given in the top of the square bracket.

Figure 5.3 and 5.4 using the basis set mentioned above. The oxidized  $\text{Cu}_B$  center has  $S=1/2$  and the reduced  $\text{Cu}_B$  center has  $S=0$ . In total eight structures were optimized for heme  $a_3$  center corresponding to the four protonation states in the oxidized and four in the reduced state as given in Figure 5.5 and 5.6 respectively. The oxidized  $\text{Fe}_{a_3}$  center has  $S=1/2$  and the reduced  $\text{Fe}_{a_3}$  center has  $S=0$ . The total charges of all the  $\text{Cu}_B$  and  $\text{Fe}_{a_3}$  complexes are given in corresponding Figure 5.3, 5.4, 5.5 and 5.6. The geometry optimizations were performed by fixing the  $C_\beta$  and  $C_\gamma$  carbon atom of histidines and Tyr288 respectively. The geometry op-

timizations were also performed by completely relaxing the system without any constraints. The solvation energies were obtained from FDPB method using CHELPG charges.



**Figure 5.5. Protonation equilibria of the heme  $a_3$  in the oxidized (+3) state.** Four possible microscopic state and four possible proton equilibrium reactions are shown for the heme  $a_3$ . Numbering of the ligands corresponds to the cytochrome *c* oxidase from *R. sphaeroides*. In the label,  $Fe_{a_3}^{+3}$  corresponds to the oxidized state of the heme  $a_3$  center. The protonation state of the corresponding microscopic state is indicated by protonation state vectors. The total charge of the complex in given in the top of the square bracket.



**Figure 5.6. Protonation equilibria of the heme  $a_3$  in the reduced (+2) state.** Four possible microscopic state and four possible proton equilibrium reactions are shown for the heme  $a_3$  center. Numbering of the ligands corresponds to the cytochrome *c* oxidase from *R. sphaeroides*. In the label,  $Fe_{a_3}^{+2}$  corresponds to the reduced state of the heme  $a_3$  center. The protonation state of the corresponding microscopic state is indicated by protonation state vectors. The total charge of the complex in given in the top of the square bracket.

### 5.2.2 CHARGE FITTING

The CHELPG [79, 146] charges were derived by fitting the electrostatic potential obtained from PW91 calculations using the *chargefit* program developed in our group [189]. The three Cartesian components of the dipole moment were used as constraint combining with single value decomposition (SVD). In CHELPG, the outer atomic radius on each atom was set to 3.0 Å and grid spacing to 0.2 Å. The Bondi radii [175] used for various atom are as follows: Fe 1.3 Å; Cu 1.3 Å; C 1.7 Å; N 1.55 Å; H 1.2 Å; O 1.5 Å. Apart from the CHELPG charges the

multipole derived charges like  $MDC_m$ ,  $MDC_d$ ,  $MDC_q$  (see section 3.1.2) were used for solvation energies calculations.

The electrostatic potential from hybrid density function theory (B3LYP) were used to derive the ESP charges by different charge methods like CHELPG [79, 146] and MK [79, 147, 148] and RESP [156]. A detailed description of these methods are given in section 3.1.2. The wave functions based charges i.e., Mulliken and NPA charges are also used for solvation free energy calculations. The RESP [150] charges are obtained from the ESP obtained from MK and CHELPG method using both harmonic and hyperbolic restraint. The Mulliken charges and NPA charges are used as initial charges for harmonic and hyperbolic restraint.

### 5.2.3 SOLVATION FREE ENERGY CALCULATIONS

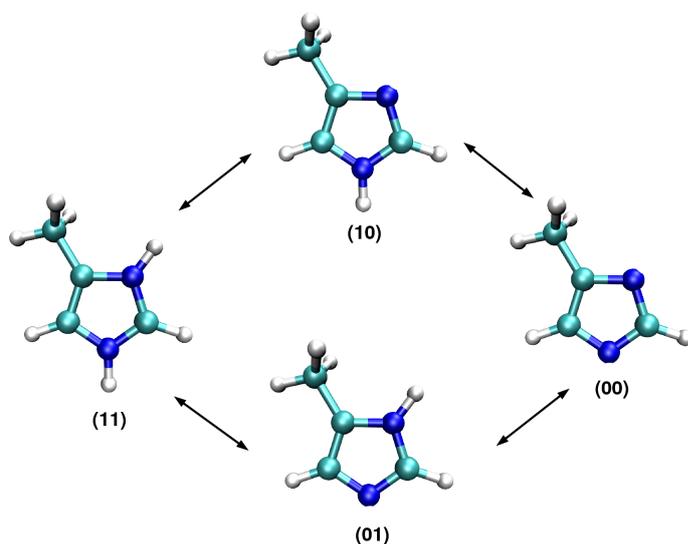
The solvation energies were calculated by both FDPB and C-PCM method. The FDPB stands for Poisson-Boltzmann equation by finite difference method where the solvation energies were obtained from electrostatic calculations by solving Poisson-Boltzmann equation by finite difference method [101–104]. In FDPB method, the solute charge distribution is considered as classical entity and does not include polarization of the solvent. The solvation free energy is obtained as the interaction of charge distribution of the solute with its reaction field. The reaction field arises due to the effective electrostatic field acting back on the charges of the solute. The solvation free energy is the free energy difference for charging the solute in vacuum and in solution. The solvation energy is given by

$$\Delta G_{solv} = \frac{1}{2} \sum_{i=1}^N q_i^{depro} \phi_i^{Rf,depro} - \frac{1}{2} \sum_{i=1}^N q_i^{pro} \phi_i^{Rf,pro} \quad (5.2)$$

where  $q_i^{depro}$  and  $q_i^{pro}$  denotes charges of the deprotonated and protonated state of the solute and  $\phi_i^{Rf,depro}$  and  $\phi_i^{Rf,pro}$  are the reaction field potential for the deprotonated and protonated state respectively. The solute is represented by a set of atomic charges and Bondi radii (see section 5.2.2). The solvation energies were calculated by the FDPB method using the program SOLVATE, which is a module in MEAD suite of programs developed by D. Bashford [65, 69, 101] which solves the Poisson-Boltzmann equation by finite difference method. All the reactants are represented by their atomic charges obtained from different charge methods and Bondi radii. The solvent is considered as a continuum dielectric medium with dielectric constant 80. The dielectric boundary between the interior and the exterior is defined by the surface of contact of a 1.4 Å radius sphere rolling over the superposition of spheres defined by the Bondi radii of the atoms. The free energy difference for charging the solute in vacuum and solution is calculated by solving the macroscopic Poisson-Boltzmann equation using finite difference method. A standard relaxation algorithm is used to obtain the reaction field potential  $\phi^{Rf}$  for each protonated and deprotonated forms [65, 69, 101] where the field points for the reaction field are evaluated at the point charge positions of protonated and deprotonated forms. The Poisson-Boltzmann equation was solved on two finer grids with spacing of 1.0 and 0.25 and sizes of  $91^3$  respectively. The C-PCM includes the polarization of the solute as well as that of the solvent; that is, the electrons of the solute are allowed to redistribute in the presence of the solvent reaction field. In C-PCM the polarization charges are included in the

solute Hamiltonian. The DFT calculations are repeated utilizing the augmented Hamiltonian of Eq. (2.85). The calculations (DFT for charge density of solute, Poisson for solvent reaction field) are repeated until the electronic energy of the solute converges to within a given tolerance value. The C-PCM solvation energies have been determined by the B3LYP method with TZVP, 6-31+G\* and 6-31G\* basis sets. The C-PCM is based on the polarizable continuum model (PCM) (see section 2.3.3) of Tomasi and co-workers [105, 122]. In C-PCM method, the cavities are modeled on molecular shape, using optimized parameters and both electrostatic and non-electrostatic contributions to the energies are included. The United Atom for Hartree-Fock (UAHF) topological model was used to build the cavity in the present calculations. In this model the vdW-surface is constructed from spheres located on heavy (i.e non-hydrogen) elements only. The total free energy is the sum of the free energies obtained for electrostatic interactions, formation of cavity in the continuum medium, dispersion interactions and repulsion interactions (see Eq. (2.86)). The dispersion and repulsion interaction terms are calculated following the procedure pioneered by Floris and Tomasi [190, 191].

**CALCULATION OF  $pK_a$  VALUES OF METHYLIMIDAZOLE.** Methylimidazoles are considered as a good model compound for histidines. The side chain of histidines has two protonation sites,  $N_{\delta_1}$  and  $N_{\epsilon_2}$ . The  $pK_a$  values for  $N_{\delta_1}$  and  $N_{\epsilon_2}$  deprotonation were found experimentally to be 7.0 and 6.6 [192] respectively. A  $pK_a$  value of  $\sim 14$  [193, 194] was reported for methylimidazolate anion. We calculated the  $pK_a$  values of the methylimidazole to bench mark the  $pK_a$  calculations. We also calculated the absolute  $pK_a$  value for all possible protonation reactions of methylimidazole. All possible protonation states and protonation reactions for methylimidazole are shown in Figure 5.7.



**Figure 5.7.** Scheme shows different protonation states of methylimidazole. The protonation state vectors are given in the bottom. The protonation state vector 11 represents the doubly protonated (both  $N_{\delta_1}$  and  $N_{\epsilon_2}$ ), the protonated state vector 10 and 01 represents singly protonated state at one of the nitrogen atoms  $N_{\delta_1}$  and  $N_{\epsilon_2}$  and 00 represents the doubly deprotonated state.

The structures were optimized by PW91 method with triple- $\zeta$  plus polarization (TZP) and B3LYP method with different basis set. In PW91 method, the gas phase energies were cal-

| React. State | Prod. State | $pK_a$           |                       |                           |
|--------------|-------------|------------------|-----------------------|---------------------------|
|              |             | (Exp.)           | PW91/TZV <sup>a</sup> | B3LYP/6-31G* <sup>b</sup> |
| 11           | 10          | 7.0 <sup>c</sup> | 8.1                   | 7.9                       |
| 11           | 01          | 6.6 <sup>c</sup> | 8.2                   | 7.4                       |
| 10           | 00          | $\sim 14^d$      | 13.2                  | 14.3                      |
| 01           | 00          | $\sim 14^d$      | 13.4                  | 13.8                      |

**Table 5.1. The  $pK_a$  values of methylimidazole.** <sup>a</sup> The  $pK_a$  values of methylimidazole obtained at PW91 method using TZV basis set. The solvation energies are obtained by using CHELPG charges. <sup>b</sup> The  $pK_a$  values of methylimidazole obtained at B3LYP method using 6-31G\* basis set. The solvation energies are obtained by using MK charges. The reactant and product states are denoted by protonation state vectors. <sup>c</sup> Ref [198] <sup>d</sup> Ref [194]

culated with TZP basis set. The solvation energies were calculated by FDPB method with CHELPG charges. The CHELPG charges were obtained by fitting the ESP obtained from the gas phase free energy calculations. The results shown in Table 5.1 are in reasonable agreement with the experimental data. In B3LYP method the gas phase free energies were calculated with 6-31G\* basis set. The solvation energies of the reactant and products were calculated using Merz-Kollman charges which are in good agreement with the experimental value compared to CHELPG and RESP charges. The Merz-Kollman charges obtained by fitting the electrostatic potential from B3LYP/6-31G\*/C-PCM level were used. All the  $pK_a$  values reported in this chapter are calculated as shown in Eq. (2.14) [75, 187]. The proton solvation free energy ( $\Delta G_{sol}^{H^+}$ ) is an elusive quantity, which will be very difficult to determine experimentally and to calculate reliably [195]. The available experimental [63, 101, 196] and calculated values reported [197] ( $\Delta G_{sol}^{H^+}$ ) vary between -252.6 and -264.6 kcal/mol. Liptak and Shields [63] used a experimental value of -264.61 kcal/mol for the proton solvation energy. They combined the complete Basis set models and the C-PCM solvation models to calculate the  $pK_a$  values of carboxylic acids and obtained the  $pK_a$  values accurate to less than half a  $pK_a$  unit. In present study, the  $pK_a$  values were calculated using an experimental solvation energy of proton -264.61 kcal/mol [63]. The calculated  $pK_a$  values for methylimidazole are given in Table 5.1 along with the experimental  $pK_a$  values for comparison.

### 5.3 RESULTS AND DISCUSSION

The  $pK_a$  values were calculated for the Cu-bound imidazole in various  $Cu_B$  and heme  $a_3$  complexes in aqueous solution. The  $pK_a$  values  $H_2O$ , His333 and His334 and  $H_2O$  were calculated both in the reduced and in the oxidized state of  $Cu_B$  center in aqueous solution. The gas phase energies were obtained using PW91 method, and the solvation free energies were obtained using different charges by FDPB method. The  $pK_a$  values obtained from PW91 method in combination with FDPB method are discussed first and the same obtained by B3LYP calculations are discussed later. The reactions are classified according to proton bound to each microscopic state. ‘Step 1’ corresponds to the removal of single proton from  $H_2O$ ,

| Step | Rea. State                         | Prod. State                        | TZP                     |        | TZ2P+ <sup>a</sup>      |        | xc-BLYP <sup>b</sup>    |        | unconstraint <sup>c</sup> |        |
|------|------------------------------------|------------------------------------|-------------------------|--------|-------------------------|--------|-------------------------|--------|---------------------------|--------|
|      |                                    |                                    | $\Delta\Delta G_{solv}$ | $pK_a$ | $\Delta\Delta G_{solv}$ | $pK_a$ | $\Delta\Delta G_{solv}$ | $pK_a$ | $\Delta\Delta G_{solv}$   | $pK_a$ |
| 1    | Cu <sub>B</sub> <sup>+2</sup> -111 | Cu <sub>B</sub> <sup>+2</sup> -011 | 98.77                   | 6.9    | 99.00                   | 7.9    | 97.56                   | 6.5    | 80.93                     | 6.5    |
|      | Cu <sub>B</sub> <sup>+2</sup> -111 | Cu <sub>B</sub> <sup>+2</sup> -101 | 96.19                   | 16.1   | 96.76                   | 17.6   | 96.38                   | 13.4   | 80.31                     | 15.8   |
|      | Cu <sub>B</sub> <sup>+2</sup> -111 | Cu <sub>B</sub> <sup>+2</sup> -110 | 97.60                   | 15.1   | 96.02                   | 14.9   | 96.33                   | 11.5   | 80.11                     | 13.3   |
| 2a   | Cu <sub>B</sub> <sup>+2</sup> -110 | Cu <sub>B</sub> <sup>+2</sup> -010 | 17.47                   | 6.4    | 22.21                   | 7.3    | 19.32                   | 7.6    | 24.10                     | 8.4    |
|      | Cu <sub>B</sub> <sup>+2</sup> -101 | Cu <sub>B</sub> <sup>+2</sup> -001 | 21.45                   | 4.4    | 19.22                   | 5.5    | 21.55                   | 5.0    | 25.13                     | 3.3    |
| 2b   | Cu <sub>B</sub> <sup>+2</sup> -110 | Cu <sub>B</sub> <sup>+2</sup> -100 | 20.42                   | 10.9   | 23.64                   | 13.9   | 21.09                   | 10.0   | 27.71                     | 12.9   |
|      | Cu <sub>B</sub> <sup>+2</sup> -011 | Cu <sub>B</sub> <sup>+2</sup> -001 | 18.87                   | 13.6   | 16.98                   | 15.2   | 20.37                   | 11.9   | 24.52                     | 12.7   |
| 2c   | Cu <sub>B</sub> <sup>+2</sup> -011 | Cu <sub>B</sub> <sup>+2</sup> -010 | 16.30                   | 14.6   | 19.23                   | 14.4   | 18.09                   | 12.6   | 23.28                     | 15.2   |
|      | Cu <sub>B</sub> <sup>+2</sup> -101 | Cu <sub>B</sub> <sup>+2</sup> -100 | 21.83                   | 9.9    | 22.90                   | 11.2   | 21.04                   | 8.1    | 27.50                     | 10.3   |
| 3    | Cu <sub>B</sub> <sup>+2</sup> -100 | Cu <sub>B</sub> <sup>+2</sup> -000 | -48.81                  | 4.6    | -49.54                  | 4.4    | -47.95                  | 5.6    | -26.60                    | 3.9    |
|      | Cu <sub>B</sub> <sup>+2</sup> -010 | Cu <sub>B</sub> <sup>+2</sup> -000 | -45.86                  | 9.1    | -48.11                  | 11.0   | -46.18                  | 8.0    | -23.00                    | 8.4    |
|      | Cu <sub>B</sub> <sup>+2</sup> -001 | Cu <sub>B</sub> <sup>+2</sup> -000 | -48.43                  | 10.1   | -45.86                  | 10.2   | -48.46                  | 8.7    | -24.23                    | 11.0   |

**Table 5.2. The solvation energy difference between deprotonated and protonated forms and the  $pK_a$  values of all possible protonation equilibrium reactions in the oxidized state of Cu<sub>B</sub> center.** The solvation energy difference between deprotonated and protonated ( $\Delta\Delta G_{solv}$ ) forms and the  $pK_a$  values of all possible proton equilibrium reactions are given. All solvation energies were obtained by FDPB method using CHELPG charges. The solvation energies are given in kcal/mol. <sup>a</sup> PW91 with TZ2P+ basis set (diffusion functions for iron and copper). <sup>b</sup> PW91 with BLYP exchange-correlation functionals and TZP basis set. <sup>c</sup> PW91 with TZP basis set where the geometry is completely relaxed.

His333 and His334. The second proton is removed from H<sub>2</sub>O in ‘step 2a’ and in ‘step 2b’ from His333 and from His334 in ‘step 2c’. In ‘Step 3’ the third proton is removed from H<sub>2</sub>O, His333 and His334. All the Tables reported in this section follow the same notation. The nomenclature used are shown in Figures 5.3, 5.4, 5.5 and 5.6 respectively. The  $pK_a$  values of all possible proton equilibrium reactions are calculated to check the internal consistency i.e., the deprotonation of the same proton binding site becomes more difficult when less protons are bound to the center. The solvation energies were calculated with different charges to obtain appropriate model  $pK_a$  for the Cu<sub>B</sub> and heme  $a_3$  center which is required to calculate the average  $pK_a$  values in the protein.

### 5.3.1 $pK_a$ VALUES OF THE Cu<sub>B</sub> CENTER

The enthalpy difference between the deprotonated and the protonated form ( $\Delta H_{vac}^{depro}$ ) were calculated for all microscopic states of Cu<sub>B</sub> center in the oxidized and in the reduced state using PW91 method with TZP, TZ2P+ basis set and also using BLYP as exchange and correlation functionals. Table 5.2 lists the solvation energy difference ( $\Delta\Delta G_{solv}$ ) between deprotonated and the protonated forms. The  $pK_a$  values are calculated based on Eq. (2.14). The  $pK_a$  values are very sensitive to the basis set used and also with different exchange correlation function-

| Step | Rea. State                         | Prod. State                        | PW91/TZP                |        |                         |        |                         |        |
|------|------------------------------------|------------------------------------|-------------------------|--------|-------------------------|--------|-------------------------|--------|
|      |                                    |                                    | MDC <sub>m</sub>        |        | MDC <sub>d</sub>        |        | MDC <sub>q</sub>        |        |
|      |                                    |                                    | $\Delta\Delta G_{solv}$ | $pK_a$ | $\Delta\Delta G_{solv}$ | $pK_a$ | $\Delta\Delta G_{solv}$ | $pK_a$ |
| 1    | Cu <sub>B</sub> <sup>+2</sup> -111 | Cu <sub>B</sub> <sup>+2</sup> -011 | 112.57                  | 17.1   | 97.55                   | 6.1    | 97.40                   | 5.9    |
|      | Cu <sub>B</sub> <sup>+2</sup> -111 | Cu <sub>B</sub> <sup>+2</sup> -101 | 117.00                  | 31.4   | 95.40                   | 15.5   | 93.70                   | 14.3   |
|      | Cu <sub>B</sub> <sup>+2</sup> -111 | Cu <sub>B</sub> <sup>+2</sup> -110 | 115.00                  | 27.9   | 95.60                   | 13.7   | 96.60                   | 14.4   |
| 2a   | Cu <sub>B</sub> <sup>+2</sup> -110 | Cu <sub>B</sub> <sup>+2</sup> -010 | 23.85                   | 11.1   | 20.49                   | 8.6    | 15.54                   | 5.0    |
|      | Cu <sub>B</sub> <sup>+2</sup> -101 | Cu <sub>B</sub> <sup>+2</sup> -001 | 29.28                   | 10.2   | 22.14                   | 5.0    | 20.30                   | 3.6    |
| 2b   | Cu <sub>B</sub> <sup>+2</sup> -110 | Cu <sub>B</sub> <sup>+2</sup> -100 | 36.45                   | 22.6   | 22.79                   | 12.6   | 17.40                   | 8.7    |
|      | Cu <sub>B</sub> <sup>+2</sup> -011 | Cu <sub>B</sub> <sup>+2</sup> -001 | 33.71                   | 24.5   | 19.99                   | 14.4   | 16.60                   | 11.9   |
| 2c   | Cu <sub>B</sub> <sup>+2</sup> -011 | Cu <sub>B</sub> <sup>+2</sup> -010 | 26.28                   | 21.9   | 18.54                   | 16.2   | 14.74                   | 13.4   |
|      | Cu <sub>B</sub> <sup>+2</sup> -101 | Cu <sub>B</sub> <sup>+2</sup> -100 | 34.45                   | 19.1   | 22.99                   | 10.7   | 20.30                   | 8.8    |
| 3    | Cu <sub>B</sub> <sup>+2</sup> -100 | Cu <sub>B</sub> <sup>+2</sup> -000 | -46.06                  | 6.6    | -46.12                  | 6.6    | -49.58                  | 4.1    |
|      | Cu <sub>B</sub> <sup>+2</sup> -010 | Cu <sub>B</sub> <sup>+2</sup> -000 | -33.46                  | 18.2   | -43.82                  | 10.6   | -47.72                  | 7.7    |
|      | Cu <sub>B</sub> <sup>+2</sup> -001 | Cu <sub>B</sub> <sup>+2</sup> -000 | -40.89                  | 15.6   | -45.27                  | 12.4   | -49.58                  | 9.2    |

**Table 5.3. The solvation energy difference between deprotonated and protonated forms and the  $pK_a$  values of all possible proton equilibrium reactions in the oxidized state of Cu<sub>B</sub> center.** The solvation energy difference between protonated and deprotonated forms ( $\Delta\Delta G_{solv}$ ) and the  $pK_a$  values of all possible proton equilibrium reactions in the oxidized state of Cu<sub>B</sub> center at PW91 method with TZP Basis set are given. The solvation energies are obtained from FDPB method using multipole derived charges. The solvation energies are given in kcal/mol.

als. Comparing the  $pK_a$  values obtained from the different basis set and exchange functionals, the variation between the  $pK_a$  values are found to be irregular. The major problem observed in the  $pK_a$  values trend is that the  $pK_a$  values correspond to ‘step 3’ where a third proton is removed from the doubly deprotonated state is 4.6 for by TZP basis set and 4.4 for TZP with diffused functions (TZ2P+) and 5.6 with BLYP exchange and correlation functionals. The trends in the  $pK_a$  values show that the removal of proton from the double deprotonated states is favored which is not energetically plausible. The protonation equilibrium should show the correct trend i.e., the deprotonation of the same proton binding site becomes more difficult when less protons are bound to the center.

On comparing the  $pK_a$  values of reactions which correspond to the removal of proton from H<sub>2</sub>O, Cu<sub>B</sub><sup>+2</sup>-111 → Cu<sub>B</sub><sup>+2</sup>-011, Cu<sub>B</sub><sup>+2</sup>-110 → Cu<sub>B</sub><sup>+2</sup>-010 and Cu<sub>B</sub><sup>+2</sup>-100 → Cu<sub>B</sub><sup>+2</sup>-000, the removal of third proton has the lower  $pK_a$  value compared to ‘step 1’ and ‘step 2a’. The solvation energy difference between the reactions are also differ very much. The solvation energy difference are very high for the reactions in ‘step 1’ where a single proton is removed from the fully protonated state. The total charges for the reactants are +2 and the product is +1 in ‘step

| Micro. State                       | TZP     | TZ2P <sup>a</sup> | xc-BLYP <sup>b</sup> | Unconstraint <sup>c</sup> | MDC <sub>m</sub> <sup>d</sup> | MDC <sub>d</sub> <sup>d</sup> | MDC <sub>q</sub> <sup>d</sup> |
|------------------------------------|---------|-------------------|----------------------|---------------------------|-------------------------------|-------------------------------|-------------------------------|
| Cu <sub>B</sub> <sup>+2</sup> -111 | -150.80 | -150.69           | -150.17              | -150.16                   | -175.60                       | -149.7                        | -154.90                       |
| Cu <sub>B</sub> <sup>+2</sup> -011 | -52.03  | -51.69            | -52.61               | -53.90                    | -63.03                        | -52.15                        | -57.50                        |
| Cu <sub>B</sub> <sup>+2</sup> -101 | -54.61  | -53.93            | -53.79               | -54.74                    | -58.60                        | -54.30                        | -61.20                        |
| Cu <sub>B</sub> <sup>+2</sup> -110 | -53.20  | -54.67            | -53.84               | -55.02                    | -60.60                        | -54.10                        | -58.30                        |
| Cu <sub>B</sub> <sup>+2</sup> -001 | -33.16  | -34.71            | -32.24               | -34.50                    | -29.32                        | -32.16                        | -40.90                        |
| Cu <sub>B</sub> <sup>+2</sup> -010 | -35.73  | -32.46            | -34.52               | -36.18                    | -36.75                        | -33.61                        | -42.76                        |
| Cu <sub>B</sub> <sup>+2</sup> -100 | -32.78  | -31.03            | -32.75               | -31.27                    | -24.15                        | -31.31                        | -40.90                        |
| Cu <sub>B</sub> <sup>+2</sup> -000 | -81.59  | -80.57            | -80.70               | -81.51                    | -70.21                        | -77.43                        | -90.48                        |

**Table 5.4. Solvation energies of each microscopic state of Cu<sub>B</sub> complex in oxidized state.** The solvation energies of the microscopic state of Cu<sub>B</sub> complex obtained by FDPB method. <sup>a</sup> PW91 with TZ2P+ basis set (diffusion functions for iron and copper). <sup>b</sup> PW91 with BLYP exchange-correlation functionals and TZP basis set. <sup>c</sup> PW91 with TZP basis set where the geometry is completely relaxed during optimization. <sup>d</sup> PW91 with TZP basis set and MDCs are used for solvation energy calculations. From <sup>a-c</sup> CHELPG charges are used for solvation energy calculations. All the solvation energies are given in kcal/mol.

1' reactions. Use of higher basis set and different exchange correlation functionals does not help very much to improved the trends in the  $pK_a$  values. To find the proper reason for such trends in the  $pK_a$  the solvation energies were calculated with different charges. The charges obtained from the multipole moment are used instead of CHELPG charges. Table 5.3 show the solvation energy difference between deprotonated and protonated forms where the solvation energies were obtained from FDPB method using multipole derived charges (MDC) (see section 3.1.2).

The  $pK_a$  values are very sensitive to the charges used. The solvation energies obtained from monopole derived charges are very high compared to the solvation energies obtained from dipole and quadrupole charges. Obviously, the  $pK_a$  values are very high with MDC<sub>m</sub>. On comparing the Tables 5.2 and 5.3, the  $pK_a$  values show almost the same trend.

The solvation energies obtained from each microscopic states of the Cu<sub>B</sub> center using FDPB methods with different basis set, exchange correlation functionals, different charges are given in Table 5.4. The solvation energies obtained from the the CHELPG charges are similar with  $\sim 1$ -2 kcal/mol difference with different basis set and exchange correlation functionals. The solvation energies from the MDC<sub>m</sub> charges are much higher compared to the solvation energies obtained from CHELPG charges. The  $pK_a$  values were calculated for the reduced state of the Cu<sub>B</sub> center to obtain the  $pK_a$  values of H<sub>2</sub>O, His333 and His334.

The gas phase free energies are obtained for all microscopic state of Cu<sub>B</sub> center in reduced state using PW91/TZP basis set. The enthalpy difference between deprotonated and protonated form ( $\Delta H_{vac}^{depro}$ ), the solvation energies of the deprotonated ( $\Delta G_{solv}^{A-}$ ) and protonated ( $\Delta G_{solv}^{AH}$ ), and the calculated  $pK_a$  values are given in Table 5.5. The  $pK_a$  values of the reduced state of the Cu<sub>B</sub> center also show the same trend like oxidized state of Cu<sub>B</sub> center. The  $pK_a$  values corresponding to 'step 2a' are lower than the reactions in 'step 1'. The solvation energies

| Step | Rea. State                         | Prod. State                        | PW91/TZP               |                        |                          |                         |        |
|------|------------------------------------|------------------------------------|------------------------|------------------------|--------------------------|-------------------------|--------|
|      |                                    |                                    | $\Delta G_{solv}^{AH}$ | $\Delta G_{solv}^{A-}$ | $\Delta H_{vac}^{depro}$ | $\Delta\Delta G_{solv}$ | $pK_a$ |
| 1    | Cu <sub>B</sub> <sup>+1</sup> -111 | Cu <sub>B</sub> <sup>+1</sup> -011 | -49.55                 | -34.35                 | -2.03                    | 15.20                   | 19.9   |
|      | Cu <sub>B</sub> <sup>+1</sup> -111 | Cu <sub>B</sub> <sup>+1</sup> -101 | -49.55                 | -36.06                 | -8.59                    | 13.59                   | 14.7   |
|      | Cu <sub>B</sub> <sup>+1</sup> -111 | Cu <sub>B</sub> <sup>+1</sup> -110 | -49.55                 | -33.54                 | -9.14                    | 16.01                   | 15.3   |
| 2a   | Cu <sub>B</sub> <sup>+1</sup> -110 | Cu <sub>B</sub> <sup>+1</sup> -010 | -34.35                 | -77.56                 | 67.82                    | -49.60                  | 9.0    |
|      | Cu <sub>B</sub> <sup>+1</sup> -101 | Cu <sub>B</sub> <sup>+1</sup> -001 | -33.54                 | -78.32                 | 47.80                    | -42.29                  | 13.0   |
| 2b   | Cu <sub>B</sub> <sup>+1</sup> -110 | Cu <sub>B</sub> <sup>+1</sup> -100 | -77.56                 | -197.25                | 48.48                    | -44.78                  | 27.5   |
|      | Cu <sub>B</sub> <sup>+1</sup> -011 | Cu <sub>B</sub> <sup>+1</sup> -001 | -82.98                 | -197.25                | 41.24                    | -43.90                  | 16.7   |
| 2c   | Cu <sub>B</sub> <sup>+1</sup> -011 | Cu <sub>B</sub> <sup>+1</sup> -010 | -78.32                 | -197.25                | 60.71                    | -48.79                  | 28.1   |
|      | Cu <sub>B</sub> <sup>+1</sup> -101 | Cu <sub>B</sub> <sup>+1</sup> -100 | -34.35                 | -82.98                 | 47.93                    | -42.36                  | 19.8   |
| 3    | Cu <sub>B</sub> <sup>+1</sup> -100 | Cu <sub>B</sub> <sup>+1</sup> -000 | -36.06                 | -77.56                 | 143.20                   | -118.93                 | 14.2   |
|      | Cu <sub>B</sub> <sup>+1</sup> -010 | Cu <sub>B</sub> <sup>+1</sup> -000 | -33.54                 | -82.98                 | 123.86                   | -114.11                 | 24.4   |
|      | Cu <sub>B</sub> <sup>+1</sup> -001 | Cu <sub>B</sub> <sup>+1</sup> -000 | -36.06                 | -78.32                 | 143.33                   | -119.00                 | 13.6   |

**Table 5.5. The enthalpy ( $\Delta H_{vac}^{depro}$ ) and the solvation energy difference ( $\Delta\Delta G_{solv}$ ) between protonated and deprotonated forms, the solvation energies of the protonated ( $\Delta G_{solv}^{AH}$ ) and deprotonated ( $\Delta G_{solv}^{A-}$ ) form, and the calculated  $pK_a$  values of Cu<sub>B</sub> center in the reduced state.** The enthalpy of difference ( $\Delta H_{vac}^{depro}$ ) between protonated and deprotonated form, the solvation energies of the protonated ( $\Delta G_{solv}^{AH}$ ) and deprotonated ( $\Delta G_{solv}^{A-}$ ) form, the solvation energy difference between deprotonated and protonated form ( $\Delta\Delta G_{solv}$ ), the calculated  $pK_a$  values of Cu<sub>B</sub> center in the reduced stateform at DFT method with TVP basis set are given. The solvation energies are obtained by FDPB method using CHELPG charges. The enthalpies and solvation free energies are given in kcal/mol.

are very low for reactions correspond to 'step 3'. The 'step 1' reactions show positive solvation energies. The enthalpy difference between the deprotonated and the protonated forms are positive for the reactions corresponding to 'step 3' where the third proton is removed from the doubly deprotonated state. The results show that the  $pK_a$  values depend on the basis sets, different exchange and correlation functionals and also the mode of the optimization. But the lower  $pK_a$  values obtained for the triply deprotonated reactions may be because of the solvation model used for the solvation energy calculations.

All the solvation energies from Table 5.2 to 5.4 were calculated using the FDPB method. Though this method was successful in many cases, [101–104] for the calculation of the  $pK_a$  values of Cu<sub>B</sub> complexes, this method seems to be insufficient. To have better understanding of the factors influencing the  $pK_a$  values, the  $pK_a$  calculations are also performed using the B3LYP method with 6-31+G\* and 6-31G\* basis sets. The solvation energies were calculated by C-PCM and FDPB method using different charges like Mulliken, NPA, CHELPG, MK and RESP. The  $pK_a$  values obtained at B3LYP level of theory with 6-31+G\* basis set and the solvation energy difference between deprotonated and protonated forms of the oxidized state of Cu<sub>B</sub>

| Step | Rea. State              | Prod. State             | B3LYP/6-31+G*           |        |                         |        |                         |        |
|------|-------------------------|-------------------------|-------------------------|--------|-------------------------|--------|-------------------------|--------|
|      |                         |                         | CHELPG                  |        | MK                      |        | RESP                    |        |
|      |                         |                         | $\Delta\Delta G_{solv}$ | $pK_a$ | $\Delta\Delta G_{solv}$ | $pK_a$ | $\Delta\Delta G_{solv}$ | $pK_a$ |
| 1    | $\text{Cu}_B^{+2}$ -111 | $\text{Cu}_B^{+2}$ -011 | 102.49                  | 6.9    | 101.64                  | 6.3    | 101.25                  | 6.0    |
|      | $\text{Cu}_B^{+2}$ -111 | $\text{Cu}_B^{+2}$ -101 | 105.21                  | 20.4   | 105.32                  | 20.5   | 105.39                  | 20.5   |
|      | $\text{Cu}_B^{+2}$ -111 | $\text{Cu}_B^{+2}$ -110 | 104.95                  | 19.0   | 104.94                  | 19.0   | 105.16                  | 19.1   |
| 2a   | $\text{Cu}_B^{+2}$ -110 | $\text{Cu}_B^{+2}$ -010 | 18.65                   | 2.1    | 18.31                   | 1.9    | 18.05                   | 1.7    |
|      | $\text{Cu}_B^{+2}$ -101 | $\text{Cu}_B^{+2}$ -001 | 22.20                   | -0.5   | 20.14                   | -2.0   | 20.47                   | -1.8   |
| 2b   | $\text{Cu}_B^{+2}$ -011 | $\text{Cu}_B^{+2}$ -001 | 24.92                   | 13.0   | 23.82                   | 12.2   | 24.61                   | 12.7   |
| 2c   | $\text{Cu}_B^{+2}$ -011 | $\text{Cu}_B^{+2}$ -010 | 21.11                   | 14.2   | 21.61                   | 14.5   | 21.96                   | 14.8   |
| 3    | $\text{Cu}_B^{+2}$ -010 | $\text{Cu}_B^{+2}$ -000 | -37.33                  | 9.5    | -38.06                  | 8.9    | -38.24                  | 8.8    |
|      | $\text{Cu}_B^{+2}$ -001 | $\text{Cu}_B^{+2}$ -000 | -41.14                  | 10.7   | -40.27                  | 11.3   | -40.89                  | 10.9   |

**Table 5.6. The solvation energy difference between deprotonated and protonated forms ( $\Delta\Delta G_{solv}$ ) and  $pK_a$  values of all possible proton equilibrium reactions in the oxidized state of  $\text{Cu}_B$  center.** The  $pK_a$  values are obtained at B3LYP/6-31+G\* basis set. The solvation energies are calculated with CHELPG, MK and RESP charges. The solvation energies are calculated using FDPB method. The solvation energies are given in kcal/mol.

center are given in Table 5.6. Initially the solvation energies were calculated with CHELPG, MK and RESP charges. The RESP charges are obtained by using hyperbolic restraint and ESP from MK method. The solvation energies were calculated using FDPB method. All the microscopic states of the  $\text{Cu}_B$  center were optimized with B3LYP/6-31+G\* basis set. The ESP charges CHELPG and MK charges were obtained by fitting the ESP obtained from B3LYP/6-31+G\* by CHELPG and MK method. The microscopic state  $\text{Cu}_B^{+2}$ -100 was unstable and did not converge. The reactions involving  $\text{Cu}_B^{+2}$ -100 microscopic state is not reported in the Table 5.6.

On comparing the solvation energy difference ( $\Delta\Delta G_{solv}$ ) between the reaction and product states, the solvation energies correspond to 'step 1' are more positive compared to other steps. The  $pK_a$  values of 'step 2a' are very low and even negative for the reaction  $\text{Cu}_B^{+2}$ -101  $\rightarrow$   $\text{Cu}_B^{+2}$ -001. According to the  $pK_a$  values, the removal of second proton from the singly deprotonated state is favored compared to the removal of single proton from the fully protonated state. The  $pK_a$  values obtained using CHELPG and MK are almost similar and in some cases they vary up to 1  $pK_a$  unit. The  $pK_a$  values obtained from RESP charges are close to the  $pK_a$  values obtained from MK charges. Either the level of theory or the basis set with diffusion and polarization function does not help to improve the  $pK_a$  trends. To check the effect of diffusion function on the  $pK_a$  calculations, the calculations were done at B3LYP level of theory with 6-31G\* basis set without diffusion functions. The solvation energy difference between deprotonated and protonated forms, the  $pK_a$  values of the oxidized state of  $\text{Cu}_B$  center are given in Table 5.7. The solvation energies were calculated with Mulliken, NPA, CHELPG and MK

| Step | Rea. State              | Prod. State             | B3LYP/6-31G*            |        |                         |        |                         |        |                         |        |
|------|-------------------------|-------------------------|-------------------------|--------|-------------------------|--------|-------------------------|--------|-------------------------|--------|
|      |                         |                         | Mulliken                |        | NPA                     |        | CHELPG                  |        | MK                      |        |
|      |                         |                         | $\Delta\Delta G_{solv}$ | $pK_a$ |
| 1    | $\text{Cu}_B^{+2}$ -111 | $\text{Cu}_B^{+2}$ -011 | 104.54                  | 11.3   | 103.52                  | 10.6   | 105.71                  | 12.1   | 105.91                  | 12.3   |
|      | $\text{Cu}_B^{+2}$ -111 | $\text{Cu}_B^{+2}$ -101 | 93.22                   | 20.0   | 86.91                   | 15.5   | 94.38                   | 20.9   | 94.36                   | 20.9   |
|      | $\text{Cu}_B^{+2}$ -111 | $\text{Cu}_B^{+2}$ -110 | 94.94                   | 19.4   | 87.92                   | 14.3   | 94.92                   | 19.4   | 95.71                   | 20.0   |
| 2a   | $\text{Cu}_B^{+2}$ -110 | $\text{Cu}_B^{+2}$ -010 | 33.24                   | 13.3   | 32.15                   | 12.5   | 31.66                   | 12.1   | 30.73                   | 11.5   |
|      | $\text{Cu}_B^{+2}$ -101 | $\text{Cu}_B^{+2}$ -001 | 32.22                   | 8.7    | 28.38                   | 5.9    | 32.89                   | 9.2    | 32.52                   | 8.9    |
| 2b   | $\text{Cu}_B^{+2}$ -110 | $\text{Cu}_B^{+2}$ -100 | 27.62                   | 15.8   | 22.07                   | 11.7   | 30.33                   | 17.7   | 27.73                   | 15.8   |
|      | $\text{Cu}_B^{+2}$ -011 | $\text{Cu}_B^{+2}$ -001 | 20.90                   | 17.4   | 11.77                   | 10.8   | 21.56                   | 17.9   | 20.97                   | 17.5   |
| 2c   | $\text{Cu}_B^{+2}$ -011 | $\text{Cu}_B^{+2}$ -010 | 23.64                   | 21.4   | 16.55                   | 16.2   | 20.87                   | 19.4   | 20.53                   | 19.1   |
|      | $\text{Cu}_B^{+2}$ -101 | $\text{Cu}_B^{+2}$ -100 | 29.34                   | 15.1   | 23.08                   | 10.6   | 30.87                   | 16.2   | 29.08                   | 14.9   |
| 3    | $\text{Cu}_B^{+2}$ -100 | $\text{Cu}_B^{+2}$ -000 | -32.74                  | 13.0   | -35.64                  | 10.9   | -35.53                  | 11.0   | -33.71                  | 12.3   |
|      | $\text{Cu}_B^{+2}$ -010 | $\text{Cu}_B^{+2}$ -000 | -38.36                  | 15.5   | -45.72                  | 10.2   | -36.86                  | 16.6   | -36.71                  | 16.7   |
|      | $\text{Cu}_B^{+2}$ -001 | $\text{Cu}_B^{+2}$ -000 | -35.62                  | 19.5   | -40.94                  | 15.7   | -37.55                  | 18.1   | -37.15                  | 18.4   |

**Table 5.7. The solvation energy difference between deprotonated and protonated forms ( $\Delta\Delta G_{solv}$ ) and  $pK_a$  values of all possible proton equilibrium reactions in the oxidized state of  $\text{Cu}_B$  center** The  $pK_a$  values are obtained at B3LYP/6-31G\* basis set. The solvation energies are calculated with Mulliken, NPA, CHELPG and MK charges. The solvation energies are calculated using FDPB method. The solvation energies are given in kcal/mol.

charges. The solvation energies were also calculated with the RESP charges with two different restraints hyperbolic and harmonic restraint using Mulliken charges as initial charges. The  $pK_a$  values obtained are much better compared to the  $pK_a$  values obtained at B3LYP/6-31+G\* basis set. But the  $pK_a$  values do not show correct trend i.e., the deprotonation of the same proton binding site becomes more difficult when less protons are bound to the center.

The  $pK_a$  values obtained using the CHELPG and MK charges are comparable. The solvation energy is more positive for the single deprotonated step compared to the others steps. The solvation energies are negative for the reactions where the third protons are removed from the doubly protonated state. The  $pK_a$  values obtained with either Mulliken or NPA charges show the same trend in the  $pK_a$  values as in  $pK_a$  values obtained from ESP based charges.

The solvation energy difference between deprotonated and deprotonated forms,  $pK_a$  values of all possible proton equilibrium reactions in the oxidized state, and the solvation energies calculated with RESP charges are given in Table 5.8. The solvation energies were calculated using FDPB method. The RESP charges are obtained by using the Mulliken and NPA charges as initial charges for harmonic and hyperbolic restraint. The electrostatic potential from MK charges were used. The  $pK_a$  values obtained by using RESP charges also show the same trend as other ESP charge methods. To get the proper understanding of the influence of the solvation energy on the  $pK_a$  calculations, the C-PCM was used to calculate the solvation free

| Step | Rea. State                         | Prod. State                        | B3LYP/6-31G*            |        |                         |        |                         |        |                         |        |
|------|------------------------------------|------------------------------------|-------------------------|--------|-------------------------|--------|-------------------------|--------|-------------------------|--------|
|      |                                    |                                    | RESP <sup>a</sup>       |        |                         |        | RESP <sup>b</sup>       |        |                         |        |
|      |                                    |                                    | qwt=0.005               |        | qwt=0.01                |        | qwt=0.005               |        | qwt=0.01                |        |
|      |                                    |                                    | $\Delta\Delta G_{solv}$ | $pK_a$ |
| 1    | Cu <sub>B</sub> <sup>+2</sup> -111 | Cu <sub>B</sub> <sup>+2</sup> -011 | 104.91                  | 11.6   | 104.59                  | 11.3   | 105.83                  | 12.2   | 105.64                  | 12.1   |
|      | Cu <sub>B</sub> <sup>+2</sup> -111 | Cu <sub>B</sub> <sup>+2</sup> -101 | 94.85                   | 21.2   | 94.23                   | 20.7   | 94.50                   | 21.0   | 94.52                   | 21.0   |
|      | Cu <sub>B</sub> <sup>+2</sup> -111 | Cu <sub>B</sub> <sup>+2</sup> -110 | 96.89                   | 20.8   | 96.76                   | 20.7   | 95.68                   | 19.9   | 95.70                   | 20.0   |
| 2a   | Cu <sub>B</sub> <sup>+2</sup> -110 | Cu <sub>B</sub> <sup>+2</sup> -010 | 30.06                   | 11.0   | 30.36                   | 11.2   | 31.02                   | 11.7   | 31.12                   | 11.7   |
|      | Cu <sub>B</sub> <sup>+2</sup> -101 | Cu <sub>B</sub> <sup>+2</sup> -001 | 33.14                   | 9.3    | 33.38                   | 9.5    | 33.37                   | 9.5    | 33.81                   | 9.8    |
| 2b   | Cu <sub>B</sub> <sup>+2</sup> -110 | Cu <sub>B</sub> <sup>+2</sup> -100 | 29.36                   | 17.0   | 29.65                   | 17.2   | 28.95                   | 16.7   | 29.16                   | 16.9   |
|      | Cu <sub>B</sub> <sup>+2</sup> -011 | Cu <sub>B</sub> <sup>+2</sup> -001 | 23.08                   | 19.0   | 23.02                   | 18.9   | 22.04                   | 18.2   | 22.69                   | 18.7   |
| 2c   | Cu <sub>B</sub> <sup>+2</sup> -011 | Cu <sub>B</sub> <sup>+2</sup> -010 | 22.04                   | 20.2   | 22.53                   | 20.6   | 20.87                   | 19.4   | 21.18                   | 19.6   |
|      | Cu <sub>B</sub> <sup>+2</sup> -101 | Cu <sub>B</sub> <sup>+2</sup> -100 | 31.40                   | 16.6   | 32.18                   | 17.2   | 30.13                   | 15.7   | 30.34                   | 15.8   |
| 3    | Cu <sub>B</sub> <sup>+2</sup> -100 | Cu <sub>B</sub> <sup>+2</sup> -000 | -36.33                  | 10.4   | -36.60                  | 10.3   | -34.31                  | 11.9   | -34.46                  | 11.8   |
|      | Cu <sub>B</sub> <sup>+2</sup> -010 | Cu <sub>B</sub> <sup>+2</sup> -000 | -37.03                  | 16.5   | -37.31                  | 16.3   | -36.38                  | 17.0   | -36.42                  | 16.9   |
|      | Cu <sub>B</sub> <sup>+2</sup> -001 | Cu <sub>B</sub> <sup>+2</sup> -000 | -38.07                  | 17.7   | -37.80                  | 17.9   | -37.55                  | 18.1   | -37.93                  | 17.8   |

**Table 5.8. The solvation energy difference between deprotonated and protonated forms ( $\Delta\Delta G_{solv}$ ) and  $pK_a$  values of all possible proton equilibrium reactions in the oxidized state of Cu<sub>B</sub> center.** The  $pK_a$  values are obtained at B3LYP/6-31G\* basis set. The solvation energies are calculated with RESP charges. The solvation energies are calculated using FDPB method. The solvation energies are given in kcal/mol <sup>a</sup> hyperbolic restraint with initial mulliken charges and electrostatic potential from MK method. <sup>b</sup> harmonic restraint with mulliken charges as initial charges and electrostatic potential from MK method. Restraint factor (qwt) of 0.005 and 0.01 are used for RESP calculations.

energies. The solvation energy difference between deprotonated and protonated forms,  $pK_a$  values obtained at B3LYP with TZVP, 6-31+G\* and 6-31G\* basis set are given in Table 5.9.

The  $pK_a$  calculations with C-PCM at B3LYP level with TZVP and 6-31+G\* also show lower  $pK_a$  values for the deprotonation reaction which are not energetically plausible. The  $pK_a$  values obtained at B3LYP with 6-31G\* show the proper trend. The protonation equilibrium show the correct trend, i.e., the deprotonation of the same proton binding site becomes more difficult when less protons are bound to the center. The enthalpy of difference ( $\Delta H_{vac}^{depro}$ ), the solvation energy difference ( $\Delta\Delta G_{solv}$ ) and the calculated  $pK_a$  values of Cu<sub>B</sub> center in the reduced state at B3LYP level with 6-31G\* are given in Table 5.10. The solvation energies are obtained by using C-PCM. The  $pK_a$  in the reduced Cu<sub>B</sub> center also show correct trend. The solvation energies obtained for all the eight microscopic states are given in Table 5.12 for comparison.

The  $pK_a$  values obtained from B3LYP/6-31G\* level of theory show that the solvation energies should be calculated in a self-consistency manner as in C-PCM. The FDPB model treats the interactions in a simpler way. The solute charge distribution is considered as classical entity and the electronic polarization due to solvent effects is discarded. The FDPB model suffers

| Step | Rea. State              | Prod. State             | C-PCM                   |        |                         |        |                         |        |
|------|-------------------------|-------------------------|-------------------------|--------|-------------------------|--------|-------------------------|--------|
|      |                         |                         | B3LYP/TZVP              |        | B3LYP/6-31+G*           |        | B3LYP/6-31G*            |        |
|      |                         |                         | $\Delta\Delta G_{solv}$ | $pK_a$ | $\Delta\Delta G_{solv}$ | $pK_a$ | $\Delta\Delta G_{solv}$ | $pK_a$ |
| 1    | $\text{Cu}_B^{+2}$ -111 | $\text{Cu}_B^{+2}$ -011 | 98.47                   | 7.0    | 97.20                   | 3.1    | 98.55                   | 7.4    |
|      | $\text{Cu}_B^{+2}$ -111 | $\text{Cu}_B^{+2}$ -101 | 96.45                   | 16.4   | 96.85                   | 14.3   | 86.72                   | 15.9   |
|      | $\text{Cu}_B^{+2}$ -111 | $\text{Cu}_B^{+2}$ -110 | 95.96                   | 14.9   | 96.24                   | 12.7   | 90.25                   | 15.2   |
| 2a   | $\text{Cu}_B^{+2}$ -110 | $\text{Cu}_B^{+2}$ -010 | 23.96                   | 8.9    | 27.01                   | 8.1    | 29.96                   | 13.0   |
|      | $\text{Cu}_B^{+2}$ -101 | $\text{Cu}_B^{+2}$ -001 | 27.60                   | 6.0    | 26.52                   | 2.6    | 35.84                   | 10.6   |
| 2b   | $\text{Cu}_B^{+2}$ -110 | $\text{Cu}_B^{+2}$ -100 | -a-                     | -a-    | -a-                     | -a-    | 27.84                   | 18.1   |
|      | $\text{Cu}_B^{+2}$ -011 | $\text{Cu}_B^{+2}$ -001 | 25.58                   | 15.4   | 26.17                   | 13.9   | 24.01                   | 19.1   |
| 2c   | $\text{Cu}_B^{+2}$ -011 | $\text{Cu}_B^{+2}$ -010 | 21.45                   | 16.8   | 26.05                   | 17.7   | 21.66                   | 20.8   |
|      | $\text{Cu}_B^{+2}$ -101 | $\text{Cu}_B^{+2}$ -100 | -a-                     | -a-    | -a-                     | -a-    | 31.37                   | 17.4   |
| 3    | $\text{Cu}_B^{+2}$ -100 | $\text{Cu}_B^{+2}$ -000 | -a-                     | -a-    | -a-                     | -a-    | -33.22                  | 14.3   |
|      | $\text{Cu}_B^{+2}$ -010 | $\text{Cu}_B^{+2}$ -000 | -32.48                  | 16.3   | -35.40                  | 10.9   | -35.34                  | 19.4   |
|      | $\text{Cu}_B^{+2}$ -001 | $\text{Cu}_B^{+2}$ -000 | -36.61                  | 17.7   | -35.52                  | 14.8   | -37.69                  | 21.0   |

**Table 5.9. The solvation energy difference between deprotonated and protonated forms ( $\Delta\Delta G_{solv}$ ) and  $pK_a$  values of all possible proton equilibrium reactions in the oxidized state of  $\text{Cu}_B$  center.** The  $pK_a$  values are obtained at B3LYP with TZVP, 6-31+G\* and 6-31G\* basis set. The solvation energies are calculated with C-PCM.

*a* The structure of  $\text{Cu}_B^{+2}$ -100 was unstable and did not converge at the given level of theory and the basis set used.

by severe limitations to obtain the solvation energy of the  $\text{Cu}_B$  complexes. The present results show that in the solvation energy calculations the electronic polarization due to the solvent should be included explicitly. In C-PCM, the interaction energies between the solvent and the solute are included in the Hamiltonian itself. In C-PCM the electron distribution of the solute is obtained with the fixed nuclei in the presence of the reaction field of the solvent.

Using the C-PCM for the oxidized state of the  $\text{Cu}_B$  model in the aqueous solution, a  $pK_a$  value of 15.9 and 15.2 are obtained for the first deprotonation reaction of His333 and His334 and the  $pK_a$  value of 7.4 for the  $\text{H}_2\text{O}$  coordinating the copper. Obviously, the  $pK_a$  values of the two histidines are too high to allow their deprotonation in aqueous solution. The  $pK_a$  values of the reduced state is even higher because the negative charge of the electron is stabilizing the proton. The  $pK_a$  trend is obtained correctly only in the combination of B3LYP/6-31G\*/C-PCM level.

| Step | Rea. State                         | Prod. State                        | B3LYP/6-31G*           |                        |                          |                         |        |
|------|------------------------------------|------------------------------------|------------------------|------------------------|--------------------------|-------------------------|--------|
|      |                                    |                                    | $\Delta G_{solv}^{AH}$ | $\Delta G_{solv}^{A-}$ | $\Delta H_{vac}^{depro}$ | $\Delta\Delta G_{solv}$ | $pK_a$ |
| 1    | Cu <sub>B</sub> <sup>+1</sup> -111 | Cu <sub>B</sub> <sup>+1</sup> -011 | -36.17                 | -15.89                 | 304.45                   | 20.28                   | 32.5   |
|      | Cu <sub>B</sub> <sup>+1</sup> -111 | Cu <sub>B</sub> <sup>+1</sup> -101 | -36.17                 | -12.90                 | 280.73                   | 23.27                   | 17.7   |
|      | Cu <sub>B</sub> <sup>+1</sup> -111 | Cu <sub>B</sub> <sup>+1</sup> -110 | -36.17                 | -12.37                 | 277.77                   | 23.80                   | 15.9   |
| 2a   | Cu <sub>B</sub> <sup>+1</sup> -110 | Cu <sub>B</sub> <sup>+1</sup> -010 | -a-                    | -a-                    | -a-                      | -a-                     | -a-    |
|      | Cu <sub>B</sub> <sup>+1</sup> -101 | Cu <sub>B</sub> <sup>+1</sup> -001 | -12.90                 | -63.16                 | 379.80                   | -50.26                  | 36.2   |
| 2b   | Cu <sub>B</sub> <sup>+1</sup> -110 | Cu <sub>B</sub> <sup>+1</sup> -100 | -12.37                 | -47.50                 | 344.09                   | -35.13                  | 21.3   |
|      | Cu <sub>B</sub> <sup>+1</sup> -011 | Cu <sub>B</sub> <sup>+1</sup> -001 | -15.89                 | -63.16                 | 356.08                   | -47.27                  | 21.1   |
| 2c   | Cu <sub>B</sub> <sup>+1</sup> -011 | Cu <sub>B</sub> <sup>+1</sup> -010 | -a-                    | -a-                    | -a-                      | -a-                     | -a-    |
|      | Cu <sub>B</sub> <sup>+1</sup> -101 | Cu <sub>B</sub> <sup>+1</sup> -100 | -12.90                 | -47.50                 | 341.13                   | -34.60                  | 19.5   |
| 3    | Cu <sub>B</sub> <sup>+1</sup> -100 | Cu <sub>B</sub> <sup>+1</sup> -000 | -47.50                 | -162.00                | 460.90                   | -114.50                 | 48.4   |
|      | Cu <sub>B</sub> <sup>+1</sup> -010 | Cu <sub>B</sub> <sup>+1</sup> -000 | -a-                    | -a-                    | -a-                      | -a-                     | -a-    |
|      | Cu <sub>B</sub> <sup>+1</sup> -001 | Cu <sub>B</sub> <sup>+1</sup> -000 | -63.16                 | -162.00                | 422.23                   | -98.84                  | 31.7   |

**Table 5.10. The enthalpy ( $\Delta H_{vac}^{depro}$ ) and the solvation energy difference ( $\Delta\Delta G_{solv}$ ) between protonated and deprotonated form and the calculated  $pK_a$  values of Cu<sub>B</sub> center in the reduced state.** The enthalpy difference  $\Delta H_{vac}^{depro}$  between protonated and deprotonated, solvation energy difference between deprotonated and protonated form  $\Delta\Delta G_{solv}$  and the calculated  $pK_a$  values of Cu<sub>B</sub> center in the reduced state at B3LYP/6-31G\* level are given. The solvation energies are obtained by using C-PCM.

*a* The structure of Cu<sub>B</sub><sup>+1</sup>-010 was unstable and did not converge at the given level of theory and the basis set used.

### 5.3.2 $pK_a$ VALUES OF THE HEME $a_3$ CENTER

The enthalpy of difference  $\Delta H_{vac}^{depro}$  between deprotonated ( $H_{vac}^{AH}$ ) and protonated ( $H_{vac}^{A-}$ ), the solvation energies of the deprotonated ( $\Delta G_{solv}^{A-}$ ) and protonated ( $\Delta G_{solv}^{AH}$ ), solvation energy difference between deprotonated and protonated form  $\Delta\Delta G_{solv}$  and the calculated  $pK_a$  values of Cu<sub>B</sub> center in the reduced and oxidized form of heme  $a_3$  center are given in Table 5.11. The solvation energies were calculated using CHELPG charges in case of PW91 method and C-PCM solvation model was used in B3LYP/6-31G\* level. The  $pK_a$  values obtained at B3LYP/6-31G\* level are higher than the  $pK_a$  values obtained at PW91/TZP basis set. The  $pK_a$  values 19.3 and 10.7 correspond to the deprotonation of His419 and H<sub>2</sub>O respectively. The  $pK_a$  value of the doubly deprotonated and reduced heme  $a_3$  center are even too high. The  $pK_a$  values of the histidine are too high to allow their deprotonation in aqueous solution.

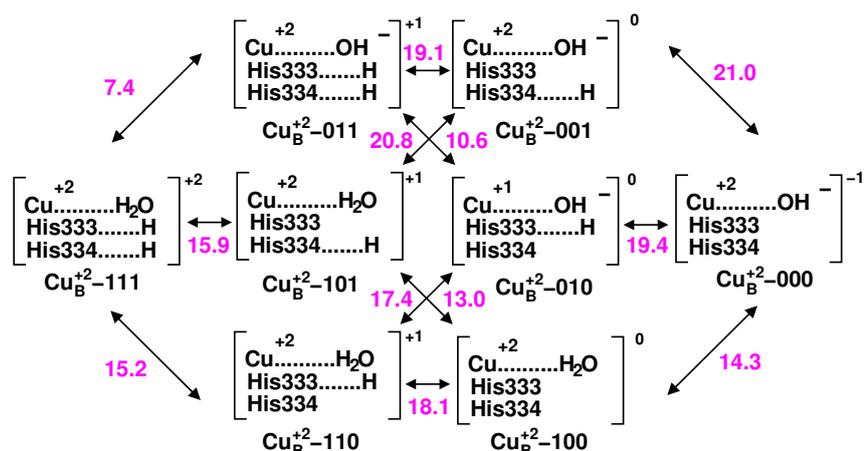
| Rea. State                         | Prod. State                        | PW91/TZP <sup>a</sup>    |                         |        | B3LYP/6-31G*             |                         |        |
|------------------------------------|------------------------------------|--------------------------|-------------------------|--------|--------------------------|-------------------------|--------|
|                                    |                                    | $\Delta H_{vac}^{depro}$ | $\Delta\Delta G_{solv}$ | $pK_a$ | $\Delta H_{vac}^{depro}$ | $\Delta\Delta G_{solv}$ | $pK_a$ |
| Fe <sub>a3</sub> <sup>+3</sup> -11 | Fe <sub>a3</sub> <sup>+3</sup> -01 | 268.65                   | 17.33                   | 7.8    | 268.65                   | 21.47                   | 10.7   |
| Fe <sub>a3</sub> <sup>+3</sup> -11 | Fe <sub>a3</sub> <sup>+3</sup> -10 | 277.26                   | 19.7                    | 10.4   | 277.26                   | 24.56                   | 19.3   |
| Fe <sub>a3</sub> <sup>+3</sup> -01 | Fe <sub>a3</sub> <sup>+3</sup> -00 | 343.72                   | -43                     | 16.5   | 343.72                   | -33.52                  | 25.4   |
| Fe <sub>a3</sub> <sup>+3</sup> -10 | Fe <sub>a3</sub> <sup>+3</sup> -00 | 335.11                   | -45.37                  | 13.8   | 335.11                   | -36.61                  | 16.8   |
| Fe <sub>a3</sub> <sup>+2</sup> -11 | Fe <sub>a3</sub> <sup>+2</sup> -01 | 364.65                   | -43.78                  | 18.3   | 364.65                   | -45.85                  | 31.7   |
| Fe <sub>a3</sub> <sup>+2</sup> -11 | Fe <sub>a3</sub> <sup>+2</sup> -10 | 349.27                   | -38.39                  | 13.5   | 349.27                   | -36.00                  | 27.7   |
| Fe <sub>a3</sub> <sup>+2</sup> -01 | Fe <sub>a3</sub> <sup>+2</sup> -00 | 402.75                   | -93.08                  | 19.9   | 402.75                   | -79.27                  | 35.2   |
| Fe <sub>a3</sub> <sup>+2</sup> -10 | Fe <sub>a3</sub> <sup>+2</sup> -00 | 418.13                   | -98.47                  | 24.7   | 418.13                   | -89.12                  | 39.3   |

**Table 5.11. The enthalpy ( $\Delta H_{vac}^{depro}$ ) and the solvation energy difference between deprotonated and protonated form  $\Delta\Delta G_{solv}$  and the calculated  $pK_a$  values of Cu<sub>B</sub> center in the reduced and oxidized states of heme *a*<sub>3</sub> center.** The  $pK_a$  values are calculated using PW91/TZP and B3LYP/6-31G\* basis sets. The solvation energies are calculated using CHELPG charges in case of PW91 method and C-PCM solvation model was used in B3LYP/6-31G\* level. The solvation energies and enthalpies are given in kcal/mol.

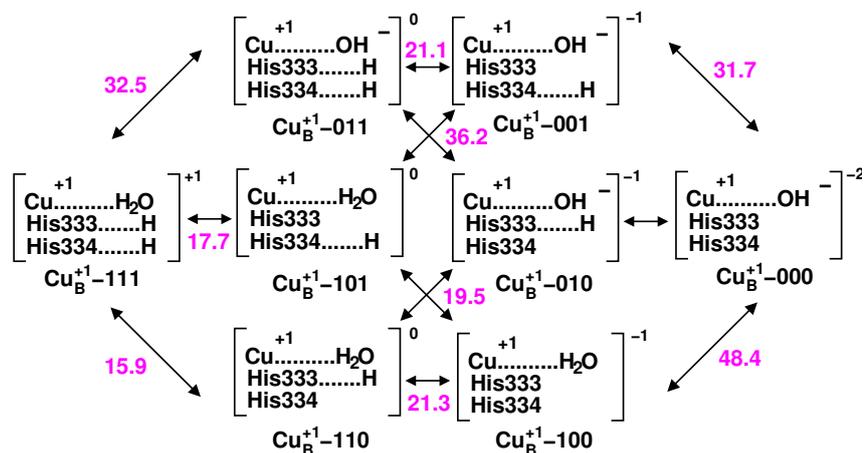
## 5.4 CONCLUSIONS

In the present study, we calculated all microscopic  $pK_a$  values of the Cu<sub>B</sub> center combining PW91 and B3LYP with both FDPB and C-PCM solvation models, with the purpose of assessing the validity of the recently proposed pumping mechanism of cytochrome *c* oxidase, based on the deprotonation of the His334 [1, 2]. The  $pK_a$  values of H<sub>2</sub>O, His333 and His334 were calculated in the aqueous solution to find the possibility of these ligands role in the proton pumping mechanism of cytochrome *c* oxidase. Extensive studies were done to understand the influence of various factors on the  $pK_a$  values of the Cu<sub>B</sub> center. The  $pK_a$  calculations performed by two different density functional methods PW91 and B3LYP and solvation models (FDPB and C-PCM) show that in solvation energy calculations the electronic polarization due to the solvent should be included explicitly, which is lacking in FDPB method. Due to this limitations, the  $pK_a$  values calculated using the solvation energies obtained from FDPB method do not show the trend correctly i.e., the removal of the third proton is favored compared to other reactions [193, 194] which is energetically not plausible. The B3LYP/6-31G\*/C-PCM level should be used to get correct  $pK_a$  trend for Cu<sub>B</sub> where solvation energies were calculated by C-PCM in which the electronic polarization of the solute due to the solvent reaction field is obtain by self-consistent manner.

The experimental  $pK_a$  value for the deprotonation of imidazole is approximately 14 [193, 194]. Our calculations, in aqueous solution on small Cu<sub>B</sub> models do not show any significant change of the  $pK_a$  due to the ligation of the imidazoles to Cu ion. Further more, the results obtained from reduced Cu<sub>B</sub> center reveals a small increase in the  $pK_a$  values of imidazoles in aqueous solution. The  $pK_a$  values of 15.9 and 15.2 for the first deprotonation reaction of His333 and



His334 are too high to allow their deprotonation in aqueous solution. The doubly and triply deprotonation reactions are not energetically plausible in aqueous solution due to their high  $pK_a$  values.



The  $pK_a$  values of His333 and His334 in the reduced state of  $\text{Cu}_B$  center are even higher to allow their deprotonation in aqueous solution. The  $pK_a$  value of 7.4 and 10.7 for the deprotonation of  $\text{H}_2\text{O}$  in oxidized  $\text{Cu}_B$  and  $\text{Fe}_{a3}$  center show that may be the  $\text{H}_2\text{O}$  can deprotonate in aqueous solution. In the oxidized state, Stuchebrukhov et al. found aqueous  $pK_a$  value in the range of 6.3-9.3. In the present study, we obtained the  $pK_a$  values of 15.2 for His334 in aqueous solution. The high  $pK_a$  values show that the deprotonation of His333 not favored in the aqueous solution. The  $pK_a$  values in the range of 15.2-21 demonstrat that His334 and His334 are protonated in aqueous solution. The charges of the  $\text{Cu}_B$  complexes play crucial role when they are transferred into protein environment. The charges obtained from the C-PCM are incorporated to the protein to find the influence of the protein environment on the  $pK_a$  values of the different proton equilibrium reactions in  $\text{Cu}_B$  center. The microscopic  $pK_a$  values obtained for  $\text{Cu}_B$  center both in oxidized and reduced state are used as a model  $pK_a$  values in the protein.

| Micro. State            | Solvation Energies (in kcal/mol) |         |         |         |               |         |         |         |              |         |         |         |         |         |
|-------------------------|----------------------------------|---------|---------|---------|---------------|---------|---------|---------|--------------|---------|---------|---------|---------|---------|
|                         | B3LYP/TZVP                       |         |         |         | B3LYP/6-31+G* |         |         |         | B3LYP/6-31G* |         |         |         |         |         |
|                         | CHELPG                           | MK      | RESP    | C-PCM   | CHELPG        | MK      | RESP    | C-PCM   | Mulliken     | NPA     | CHELPG  | MK      | RESP    | C-PCM   |
| $\text{Cu}_B^{+2}$ -111 | -158.69                          | -160.85 | -160.70 | -147.42 | -159.04       | -161.45 | -161.80 | -147.25 | -165.5       | -188.44 | -160.35 | -162.48 | -162.68 | -137.35 |
| $\text{Cu}_B^{+2}$ -011 | -54.70                           | -57.51  | -56.05  | -48.95  | -56.55        | -59.81  | -60.55  | -50.05  | -61.02       | -84.92  | -54.64  | -56.57  | -56.85  | -38.80  |
| $\text{Cu}_B^{+2}$ -101 | -54.80                           | -56.30  | -55.56  | -50.97  | -53.83        | -56.13  | -56.41  | -50.40  | -72.34       | -101.53 | -65.97  | -68.12  | -68.18  | -50.63  |
| $\text{Cu}_B^{+2}$ -110 | -53.82                           | -56.14  | -55.54  | -51.46  | -54.09        | -56.51  | -38.59  | -51.01  | -70.62       | -100.52 | -65.43  | -66.77  | -67.0   | -47.10  |
| $\text{Cu}_B^{+2}$ -001 | -30.49                           | -34.43  | -33.12  | -23.37  | -31.63        | -35.99  | -35.94  | -23.88  | -40.12       | -73.15  | -33.08  | -35.60  | -34.81  | -14.79  |
| $\text{Cu}_B^{+2}$ -010 | -34.46                           | -37.35  | -34.32  | -27.50  | -35.44        | -38.20  | -38.59  | -24.00  | -37.38       | -68.37  | -33.77  | -36.04  | -35.98  | -17.14  |
| $\text{Cu}_B^{+2}$ -100 | -a-                              | -a-     | -a-     | -a-     | -a-           | -a-     | -a-     | -a-     | -43.0        | -78.45  | -35.10  | -39.04  | -38.05  | -19.26  |
| $\text{Cu}_B^{+2}$ -000 | -72.02                           | -75.70  | -74.71  | -59.98  | -72.77        | -76.26  | -76.83  | -59.40  | -75.74       | -114.09 | -70.63  | -36.04  | -72.36  | -52.48  |

**Table 5.12. Solvation energies of each microscopic state of  $\text{Cu}_B$  complex in oxidized state.** The solvation energies of the microscopic state obtained by FDPB method using CHELPG, MK and RESP charges and C-PCM model are given. The solvation energies are given in kcal/mol.

-a- structure is unstable and did not converge.



# $pK_a$ CALCULATIONS OF $C_{uB}$ LIGANDS IN CYTOCHROME *c* OXIDASE

---

Solution pH is a major determinant of protein stability as well as of protein function. For this reason, understanding the factors that govern the  $pK_a$  values of amino acids in proteins has been a classical problem in biophysical chemistry [128, 199]. The  $pK_a$  values of titrating groups in proteins play an essential role in their stability and function. In recent years, renewed interest in this problem has been due to both theoretical advances in the treatment of the electrostatic properties of proteins as well as the improved experimental probes of electrostatic effects on protein stability [128].

The basic theoretical problem in calculating  $pK_a$  values is, to predict how the protein environment shifts the  $pK_a$  of an amino acid in the protein from that of the isolated amino acid in solution. The underlying factors affecting these shifts are well understood, at least in a qualitative sense. For instance, the equilibrium between the ionized and neutral states in aqueous solution is determined in part by the different solvation free energies of the two species. In proteins, although interactions with water are restricted, they are replaced by new interactions with the electrons, permanent dipoles of the protein and other titratable groups.

Advances in numerical methods now make it possible to solve the Poisson-Boltzmann equation for molecules of arbitrary shape and charge distribution. This approach provides a detailed description of the full range of electrostatic interactions including those due to permanent and induced dipoles in the protein, the dielectric effects of the solvent and the effects of ionic strength.

A macroscopic electrostatic model was used to calculate the  $pK_a$  values of the titratable groups in lysozyme. The model uses a detailed structural information and solvation self-energies and interactions arising from permanent partial charges and titratable charges [65]. The continuum dielectric model and finite difference technique were used to calculate the  $pK_a$  values of the titrating groups of lysozyme. The same model was used to study the  $pK_a$  values of the ionizable groups in bacteriorhodopsin [69].

The redox states and protonation states of the protein play a key role in the reaction mechanism of enzymes. To study the function of the enzyme, the electrostatic interactions between the redox states and protonation states of the protein are required. To estimate such interactions and thereby to get insights into the possible mechanism of the coupled reactions, computer simulations can be performed. A bacterial cytochrome *c* oxidase [179] and

other membrane proteins such as bacterial photosynthetic reaction center [200–202], bacteriorhodopsin [203, 204] and cytochrome *bc*<sub>1</sub> complex that take part in energy transduction in cells were recently investigated by such methods.

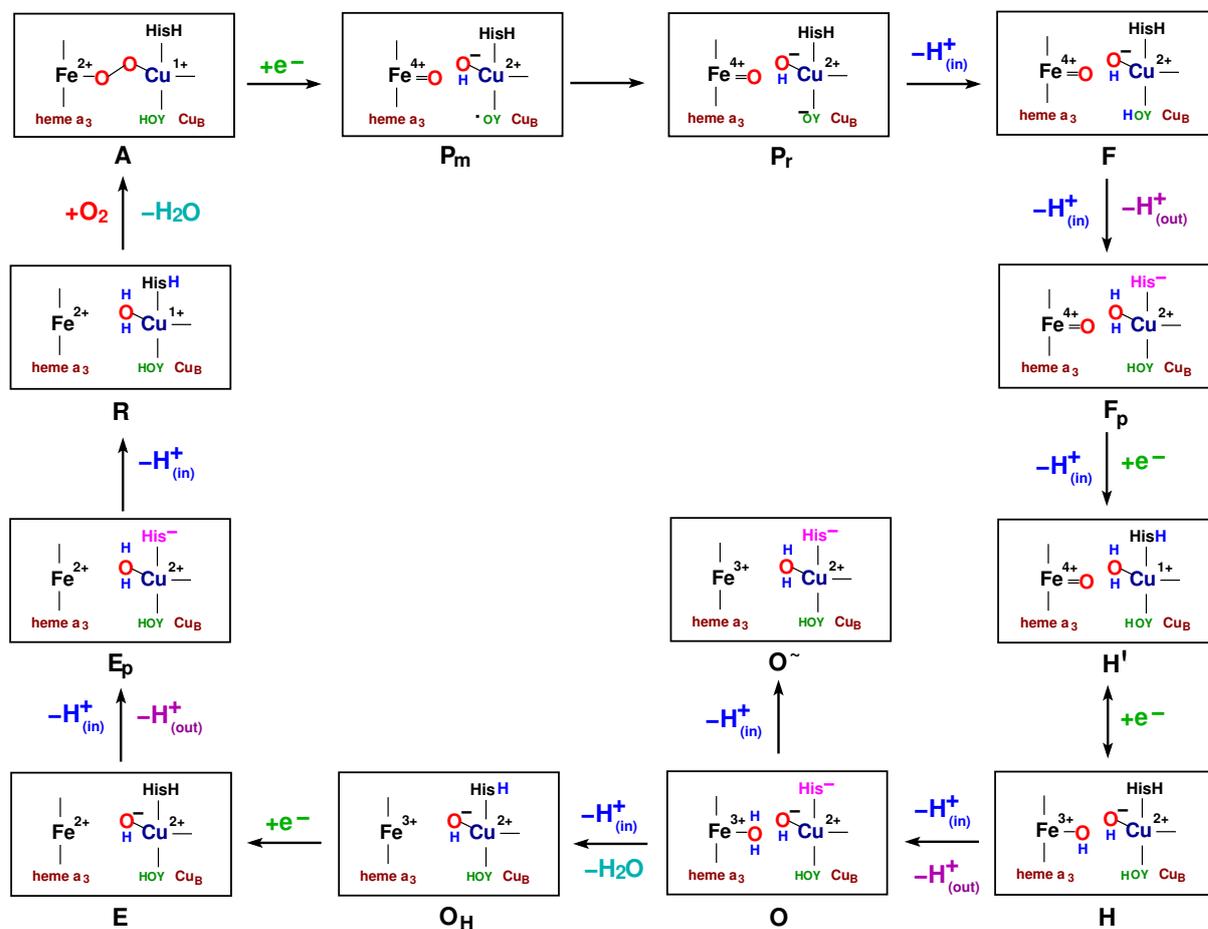
## 6.1 PREVIOUS COMPUTATIONAL WORK ON CYTOCHROME *c* OXIDASE

Michel and co-workers performed electrostatic calculations to study the reaction mechanism of cytochrome *c* oxidase from *P. denitrificans*. They calculated the electrostatic potential and the interaction energies of ionizable groups and analyzed the response of the protein environment to redox changes in *P. denitrificans* cytochrome *c* oxidase around the heme *a*<sub>3</sub> and Cu<sub>B</sub>. They identified a cluster of 18 strongly interacting residues that account for the most of the proton uptake linked to the electron transfer [179].

Siegbahn and co-workers [184] used hybrid density functional theory with B3LYP functional to study the energetics of proton translocation in cytochrome *c* oxidase. They computed the redox potentials of the metal centers and the tyrosyl radical and also calculated the p*K*<sub>a</sub> values of important groups along the translocation path. The main finding of their work is that heme propionate A of heme *a*<sub>3</sub> can act as a potential proton-loading site [184].

Stuchebrukhov and Popović performed continuum electrostatic calculations to evaluate the electrostatic potential, energies and protonation state of bovine heart cytochrome *c* oxidase for different redox states of the enzyme along its catalytic cycle. They studied the energetics and the pH dependence of the P→F, F→H and O→E steps of the catalytic cycle. On the basis of the electrostatic calculations, they proposed two possible schemes for the redox-linked proton pumping. The first scheme involves His291 in bovine heart or His334 in *R. sphaeroides* as a pump element and the second scheme involves a group linked to the propionate D of the heme *a*<sub>3</sub> [1]. The authors proposed a possible scheme based on their calculations, where His334 (one of the ligands of the Cu<sub>B</sub> center) deprotonates at its N<sub>δ1</sub> and plays a role of the proton loading site in the proton pump [2]. Based on these calculations they claimed that His334 can act as a proton loading site in the pumping mechanism of cytochrome *c* oxidase.

Stuchebrukhov et al. [1] performed DFT and continuum electrostatic calculations to evaluate the p*K*<sub>a</sub> values of His334 and Glu286 in bovine heart or Glu286 in *R. sphaeroides*, the two proposed key residues in the reaction mechanism of cytochrome *c* oxidase. They calculated the p*K*<sub>a</sub> values for different redox states of the enzyme and studied the influence of different factors on the p*K*<sub>a</sub> values. From DFT calculations on the models of the Cu<sub>B</sub> in aqueous solution, these authors found that the p*K*<sub>a</sub> value of His334 is 8.6 for the oxidized and 13.2 for the reduced Cu<sub>B</sub> center. In the protein, the p*K*<sub>a</sub> value of 5.4 is reported for the oxidized state and 21.9 for the reduced state [3]. Based on these results they proposed that the N<sub>δ1</sub> proton of His334 is deprotonated when the chemical protons (proton involved in oxygen reduction) enter the binuclear center of enzyme where the hydroxyl group is converted to water molecule and thus His334 can act as a proton loading site in the pumping mechanism of cytochrome *c* oxidase. Fadda et al. used DFT in combination with continuum electrostatic calculations to calculate the p*K*<sub>a</sub> values of His334 in various Cu<sub>B</sub> and Fe*a*<sub>3</sub>-Cu<sub>B</sub> complexes in bulk water and within the protein [182]. They obtained p*K*<sub>a</sub> value of His334 in aqueous solution in the range



**Figure 6.1. Proposed main steps of the reduction of  $O_2$  to  $H_2O$  in the binuclear center of cytochrome *c* oxidase.** The heme  $a_3$  is represented by the central Fe atom and by a line defining the porphyrin ring. Only the  $Cu_B$  histidine ligand (His334) side chain undergoing deprotonation is shown. The protonated and deprotonated forms of imidazole are indicated as HisH and His<sup>-</sup>. Catalytic states are defined commonly used nomenclature. The protons entering and exiting the binuclear center are indicated as  $H^+_{(in)}$  and  $H^+_{(out)}$ . The intermediate states are further detailed in their electronic configuration at the binuclear center. A tyrosine side chain of subunit I is shown.  $YO^\bullet$  stands for the neutral radical of the His284-Y288 cross-link and  $YO^-$  for the tyrosinate.

of 14.8-16.6 and in the protein the  $pK_a$  values range between 14.9-26.9. The results obtained by Fadda et al. were inconsistent with the proposed role of the Cu-bound histidine as a key element in the pumping mechanism. The discrepancy between these two studies lead to some debate in the field [4, 183].

A schematic depiction of main steps of oxygen reduction in the binuclear center, and the proposed connection to the His334 deprotonation [1] are shown in Figure 6.1. The main steps  $A \rightarrow P_m \rightarrow P_r$  states are explained in detailed in chapter 1. The steps associated with the His334 deprotonation ( $F \rightarrow F_p$ ,  $H \rightarrow O$ ,  $O \rightarrow O^\sim$  and  $E \rightarrow E_p$ ) in the catalytic cycle are discussed here.

In the transition between the  $F \rightarrow F_p$  states, a chemical proton (proton involved in oxygen reduction) enters the binuclear center and converts the hydroxyl  $Cu_B$  ligand into water thereby increasing the total charge of the site from +1 to +2. This state of the catalytic cycle is supposedly connected to the first deprotonation of the His334 imidazole ring.

The coupled electron-proton transfer reaction leads to  $H'$  state. It was proposed [1] that  $H'$  is likely in dynamic equilibrium with the  $H$  state in which iron is in the +3 state and copper ion is in the +2 state and each of the metals is ligated to the hydroxyl group.

The transition between the  $H$  and  $O$  states is connected to a second deprotonation of His334. The protonation of the hydroxyl group bound to heme  $a_3$  leads to the formation of  $O$  state, which is associated with the pumping event. The  $O$  state can evolve in two different ways. The proton entering the binuclear center can protonate the hydroxyl group of  $Cu_B$  center which is energetically least favorable ( $O^\sim$  state). The another possible pathway ( $O_H$ ) results in the formation of  $E$  state.

The  $E$  state was proposed [1] to be thermodynamically unstable and converted to  $E_p$ . In  $E$  state, the redox state of heme  $a_3$  changes from +3 to +2. In  $E_p$  state, the chemical proton enters the active site and converts the hydroxyl group to water, triggering the third deprotonation of His334. The next incoming proton reloads the His334 site, closing the catalytic cycle with removal of water molecule.

In the two studies [1–3, 5, 6, 182], the crystal coordinates of the  $Cu_B$  center was not fully relaxed during the geometry optimizations which can lead to unrealistic bond lengths. Moreover in later studies [182], it was not considered that some residues of the cytochrome *c* oxidase can adopt a protonation deviating from their usual protonation in aqueous solution. In the two studies, only the deprotonation of His334 was determined and not all possible microscopic proton equilibrium reactions in the protein are considered. There is a possibility that His333, another ligand of  $Cu_B$  center can serve as a proton loading site for subsequent pumping which was not examined.

To understand the coupling of the electron transfer and proton translocation, DFT and continuum electrostatic calculations were performed on *R. sphaeroides* cytochrome *c* oxidase. Initially all the microscopic  $pK_a$  values of the bound  $H_2O$ , His333 and His334 in  $Cu_B$  center were calculated in aqueous solution by combining DFT calculations with PCM model (see chapter 5). In order to take the influence of the protein environment into account, the average  $pK_a$  values of  $H_2O$ , His333 and His334 ligands bound to  $Cu_B$  were calculated in the protein by considering the protonation of the protein at  $pH=7$ . The  $Cu_B$  center was considered in both reduced and oxidized state. The heme  $a_3$  center was considered with all possible ligand states and ligand unbound states. The electrostatic calculations were performed by solving the Poisson Boltzmann Equation by finite difference method.

## 6.2 STRUCTURE PREPARATION AND MODELS

Over the past years, substantial progress has been made in the understanding of the structure and function of cytochrome *c* oxidase. A landmark in the field of cytochrome *c* oxidase research was the determination of the three-dimensional structures of the bacterial cytochrome *c* oxidase from the soil bacterium *P. denitrificans* [12, 14] and its mammalian counterpart from

bovine heart mitochondria [16–18]. The first crystals of the bacterial enzyme were obtained using the four subunit cytochrome *c* oxidase [12]. Later, an improved crystal form of the functionally active two-subunit cytochrome *c* oxidase complex with the F<sub>v</sub> antibody fragment as the four-subunit cytochrome *c* oxidase was observed and the structure was determined at 2.7 Å resolution [14]. The structures of the bacterial and the mitochondrial enzymes are surprisingly similar. The core parts (subunits I, II, and III) of the two crystal structures look nearly identical at the atomic level.

The following section deals with the structures and necessary preparations for the electrostatic calculations of cytochrome *c* oxidase.

### 6.2.1 PREPARATION OF X-RAY STRUCTURE OF PROTEIN

The electrostatic calculations were performed on the X-ray structure of *R. sphaeroides* (PDB ID:1M56). The structure of cytochrome *c* oxidase have been determined at 2.8 Å resolution by Iwata et al. [15]. The crystal structures were determined for wild type and mutant by the replacement of glutamate-286 of subunit I by glutamine. The electrostatic calculations were performed on the wild type enzyme. The overall structure of the subunit I-IV of *R. sphaeroides* is very similar to those of the *P. denitrificans* while subunit I-III are similar to the corresponding subunits of the bovine heart cytochrome *c* oxidase. The environments around the redox-metal centers (the binuclear center, heme *a* and Cu<sub>A</sub>) and the non-redox-active centers (Mg<sup>2+</sup> and Ca<sup>2+</sup>) are structurally very similar to those of both the bovine heart and *P. denitrificans*. Six phospholipid molecules were identified in the X-ray structure of *R. sphaeroides* and are assigned as phosphatidylethanolamines from the shapes of the electron density. These phospholipid molecules are also included in the electrostatic calculations. The electrostatic calculations were performed in the monomer of cytochrome *c* oxidase.

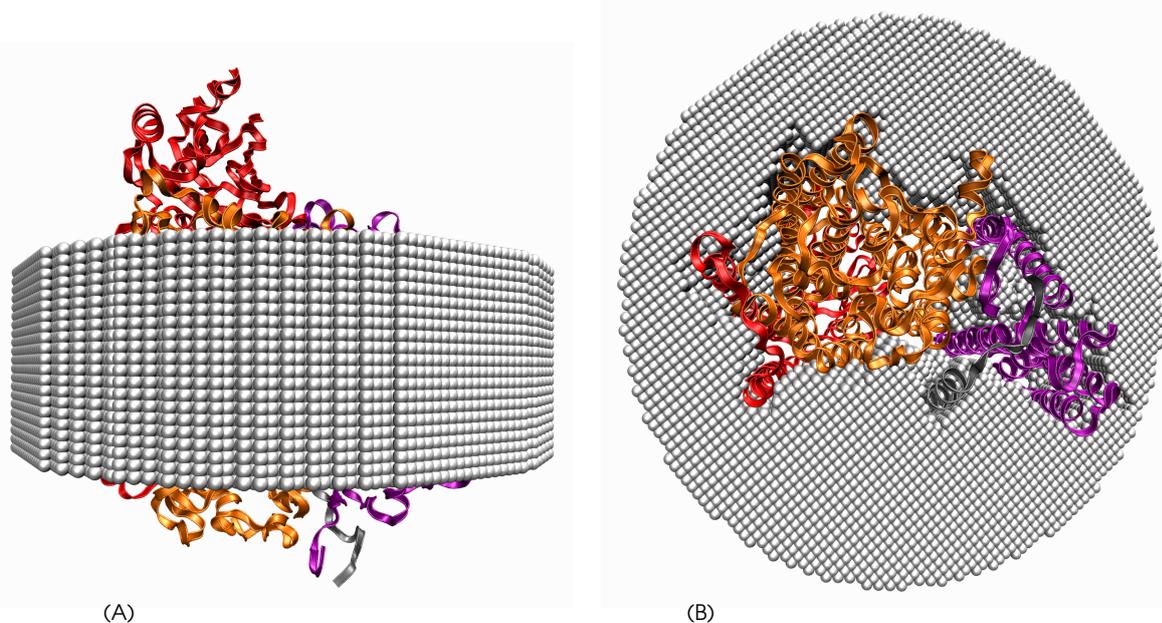
Structures were prepared for electrostatic calculations using the CHARMM [153] molecular modeling package. Hydrogen atom positions were generated using the HBUILD algorithm implemented in the CHARMM program. All the heavy atoms were fixed and energy was minimized using the CHARMM forcefield. The energy minimizations were performed by 500 steepest decent (SD) steps, followed by 500 conjugate gradient (CG) steps.

Partial atomic charges for atoms of the standard amino acids were taken from the CHARMM parameter set. The partial atomic charges for the redox-metal centers (the binuclear center, heme *a* and Cu<sub>A</sub>) and the non-redox-active centers (Mg<sup>2+</sup> and Ca<sup>2+</sup>) were obtained from quantum chemical calculations (see section 6.3). The protein structures were minimized with the crystallographic water molecules. In electrostatic calculations the crystallographic water molecules were removed since the construction of water hydrogens would arbitrarily assign a certain orientation to the water molecules that would affect electrostatic calculations. The C-terminus of subunits I-IV of cytochrome *c* oxidase were not resolved and a neutral blocking group (acetyl group) was added to the C-terminus.

CONSTRUCTION OF A MEMBRANE MODEL. In order to include the effect of the membrane environment into the electrostatic calculations, the protein complex was embedded into a cylindrical shaped belt of uncharged dummy atoms where the dummy atoms model the hydrophobic membrane core (see Figure 6.2). The dummy atoms which mimics the membrane environ-

ment were assigned a low dielectric constant in the electrostatic calculations. The Bondi radii were used for the protein atoms [175]. The cavities within the protein that show no connection to the membrane environment were not be filled with the dummy atoms. These cavities are considered to contain distorted water molecules and assigned a high dielectric constant.

The structure of cytochrome *c* oxidase with membrane was obtained from the orientations of protein in membranes (OPM) database [205–207]. The OPM database currently includes all unique structures of transmembrane protein complexes and selected monotopic, peripheral proteins and membrane-bound peptides from PDB with their calculated membrane boundaries. Coordinate files of the proteins with calculated membrane boundaries are available for download. The coordinates of cytochrome *c* oxidase with calculated membrane boundaries were superimposed with the crystal structure of *R. sphaeroides* using the program *superimpose*. The membrane was constructed around this superimposed structures. The *Vol* program [208] was used to construct a membrane around the protein. The thickness of the membrane model is 30 Å. The radius of the contracted membrane is 50 Å. A probe radius of 1.4 Å was used for generating the protein surface. The membrane model was constructed for all structures considered in the electrostatic calculations.



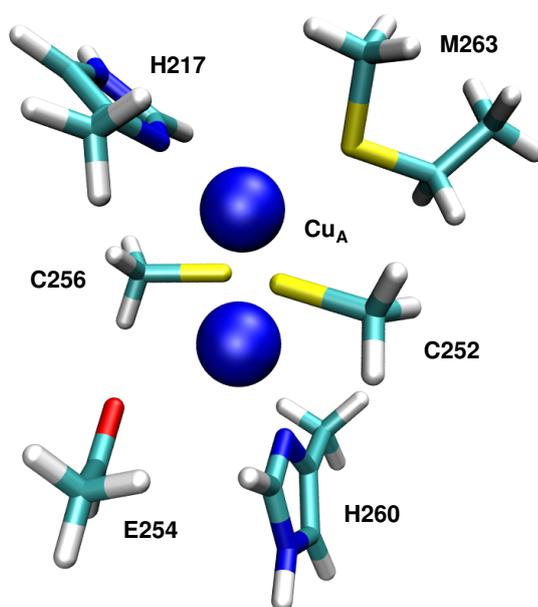
**Figure 6.2. Membrane model.** (A): View of cytochrome *c* oxidase from *R. sphaeroides* along the membrane plane. The dummy atoms which mimics the membrane environment were assigned a low dielectric constant. (B): View from the cytoplasm.

### 6.2.2 REDOX CENTER MODELS

In the reaction mechanism of cytochrome *c* oxidase, four redox centers are involved. The reduction of molecular oxygen to water takes place in the binuclear center where the electrons and protons are delivered for oxygen reduction and four protons are translocated across the inner-mitochondrial membrane, a process which results in a membrane electrochemical pro-

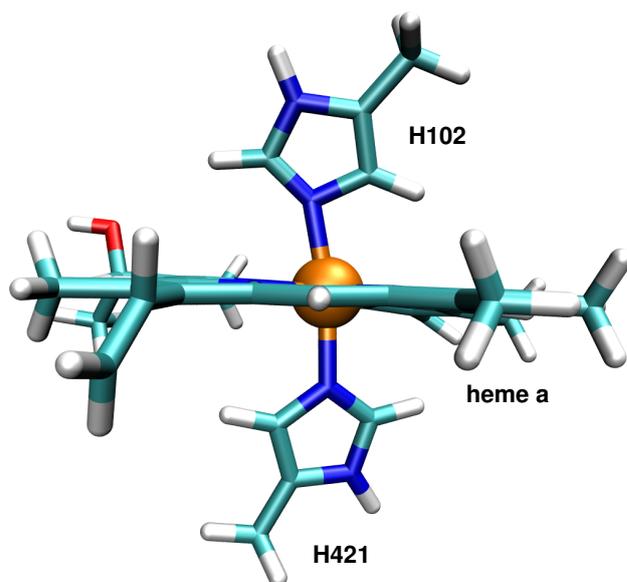
ton gradient. The proton translocation is coupled with the electron transfer which makes the reaction difficult to study. To incorporate the redox centers to the protein, the redox centers were considered for the quantum chemical calculations to get the charges with different oxidized and reduced states as well these models were used for the DFT calculations to obtain the gas phase and solvation energies. The models used for the redox centers are described below. DFT calculations were also performed in non-redox-active centers ( $Mg^{2+}$  and  $Ca^{2+}$ ) to obtain partial atomic charges.

**MODEL FOR  $Cu_A$  CENTER.** The  $Cu_A$  center contains two copper ions (see Figure 6.3) and are bridged by two cysteins sulfur atoms. The two copper ions of the  $Cu_A$  center are coordinated by two His, one Met, a backbone carbonyl oxygen of Glu, and two bridging Cys residues. The ligated cysteins were simplified as methyl thiolates. The histidines were modeled by the methyl imidazoles. The backbone carbonyl oxygen of Glu was included in the model compound. The amino acid side chains were cut at the  $C_\alpha$  atom and their  $C_\beta$  atoms were fixed in their crystal structure positions. The initial coordinates were obtained from the crystal structure of *R. sphaeroides*. All the hydrogens were added using `babel` program. The  $Cu_A$  center model was used to obtain the charges for the calculations. The charges were obtained for both reduced and oxidized states of  $Cu_A$  center.



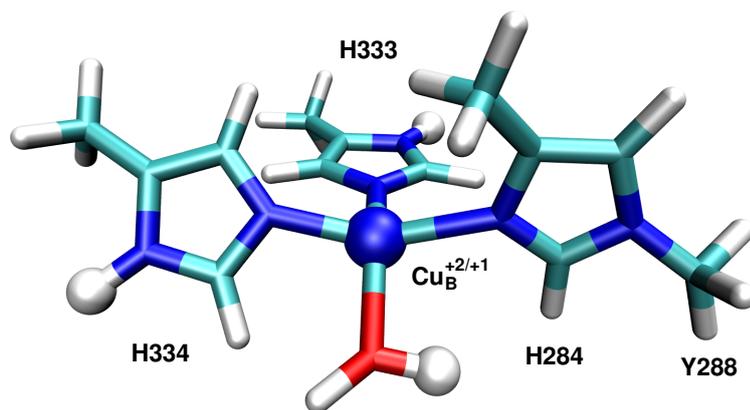
**Figure 6.3. The model compound of  $Cu_A$  center** The cysteins are simplified as methyl thiolates. The histidines are modeled by methyl imidazole.

**MODEL FOR HEME *a*** Heme *a* center is with two histidine residues as axial iron-ligands (see Figure 6.4). The histidine residues were modeled by methyl imidazole. The heme propionates were cut off and substituted by hydrogen atoms. The hydrophobic hydroxyethyl-farnesyl group was truncated next to the hydroxyl group. The calculations were performed both in the reduced and in the oxidized state.



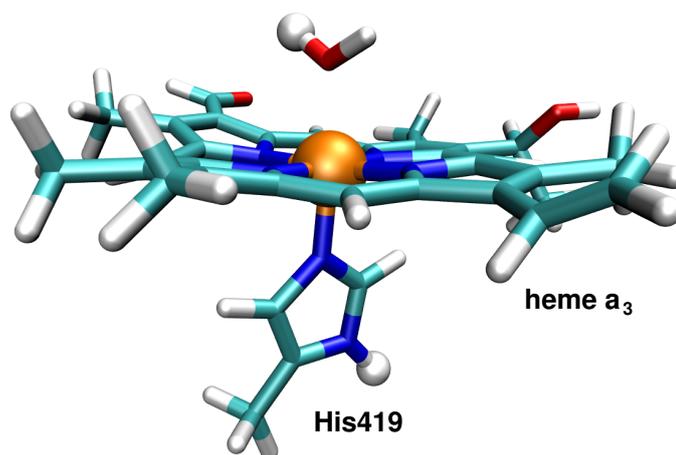
**Figure 6.4. The model compound of heme *a*.** The histidines coordinating the iron are modeled by methyl imidazole. The hydrophobic hydroxyethyl-farnesyl group was truncated next to the hydroxy group. The heme propionates were cut off and substituted by hydrogen atoms.

MODEL FOR Cu<sub>B</sub> CENTER. The model for the Cu<sub>B</sub> center consists of Cu ion, methylimidazole model for coordinating histidines 284, 333, and 334 and the tyrosine 288 was modeled by methyl group (see Figure 6.5). The fourth coordination was modeled with H<sub>2</sub>O molecule and a hydroxyl group. The calculations were performed both in the reduced and the in the oxidized state.



**Figure 6.5. The model compound of Cu<sub>B</sub> center.** The histidines coordinating the copper ion are modeled by methyl imidazole. The cross linked tyrosine 288 is replaced by methyl group. The fourth coordination is the H<sub>2</sub>O molecule. The proton binding sites are indicated by white spheres.

MODEL FOR HEME  $a_3$  CENTER. The model for the heme  $a_3$  center consists of the heme  $a_3$  and a axial histidine which was modeled by methylimidazole (see Figure 6.6). The sixth coordination was modeled with  $H_2O$  molecule ( $Fe^{III}-H_2O$  and  $Fe^{II}-H_2O$ ), hydroxyl group ( $Fe^{III}-OH$  and  $Fe^{II}-OH$ ), oxo-ferryl ( $Fe^{IV}=O$ ) and ligand unbound state ( $Fe^{III}$  and  $Fe^{II}$ ). The hydrophobic hydroxyethyl-farnesyl group was truncated next to the hydroxyl group. The heme  $a_3$  center was optimized both in oxidized and reduced states with all ligands mentioned above.



**Figure 6.6. The model compound of heme  $a_3$  center.** The histidines coordinating the iron are modeled by methyl imidazole. The hydrophobic hydroxyethyl-farnesyl group was truncated next to the hydroxyl group. The heme propionates were cut off and substituted by hydrogen atoms. The  $H_2O$  molecule is considered in the six coordination position is shown. The proton binding sites are indicated by white spheres.

### 6.3 DENSITY FUNCTIONAL CALCULATIONS

The DFT calculations [81, 172] were performed to obtain partial atomic charges of the redox centers  $Cu_A$ , heme  $a$ , heme  $a_3$  and  $Cu_B$  respectively. For redox centers  $Cu_A$ , heme  $a$  and heme  $a_3$  the Perdrew-Wang 91 (PW91) calculations were performed with ADF 2004.01 [173] program. The partial atomic charges of  $Cu_B$  center were obtained by B3LYP method using 6-31G\* basis sets using GAUSSIAN 03 program. The DFT calculations were performed on the oxidized states of  $Cu_A$  and heme  $a$ . Following seven states were considered for heme  $a_3$ : aqua ferric state ( $Fe^{III}-H_2O$ , charge=+1,  $S=5/2$ ), aqua ferrous state ( $Fe^{II}-H_2O$ , charge=0,  $S=2$ ), hydroxyl state ( $Fe^{III}-OH$ , charge=0,  $S=5/2$  and  $Fe^{II}-OH$ , charge=-1,  $S=2$ ), oxo-ferryl state ( $Fe^{IV}=O$ , charge=0,  $S=2$ ) and ligand unbound states ( $Fe^{III}$ , charge=+1,  $S=5/2$  and  $Fe^{II}$ , charge=0,  $S=2$ ). The local density approximation (LDA) for exchange and correlation are based on the parametrization of Vosko, Wilk and Nausair [84]. The Perdrew-Wang 91 (PW91) [86] exchange and correlation functionals were used for the generalized gradient approximation (GGA). The numerical integration scheme used in this calculation was Voronoi polyhedron method developed by te Velde et al. with the accuracy parameter ACCINT set to its default value. A set of triple  $\zeta$  Slater type orbital (STO) was employed with single polarization function. The inner core shells were treated by the frozen core approximation. All the calculations were

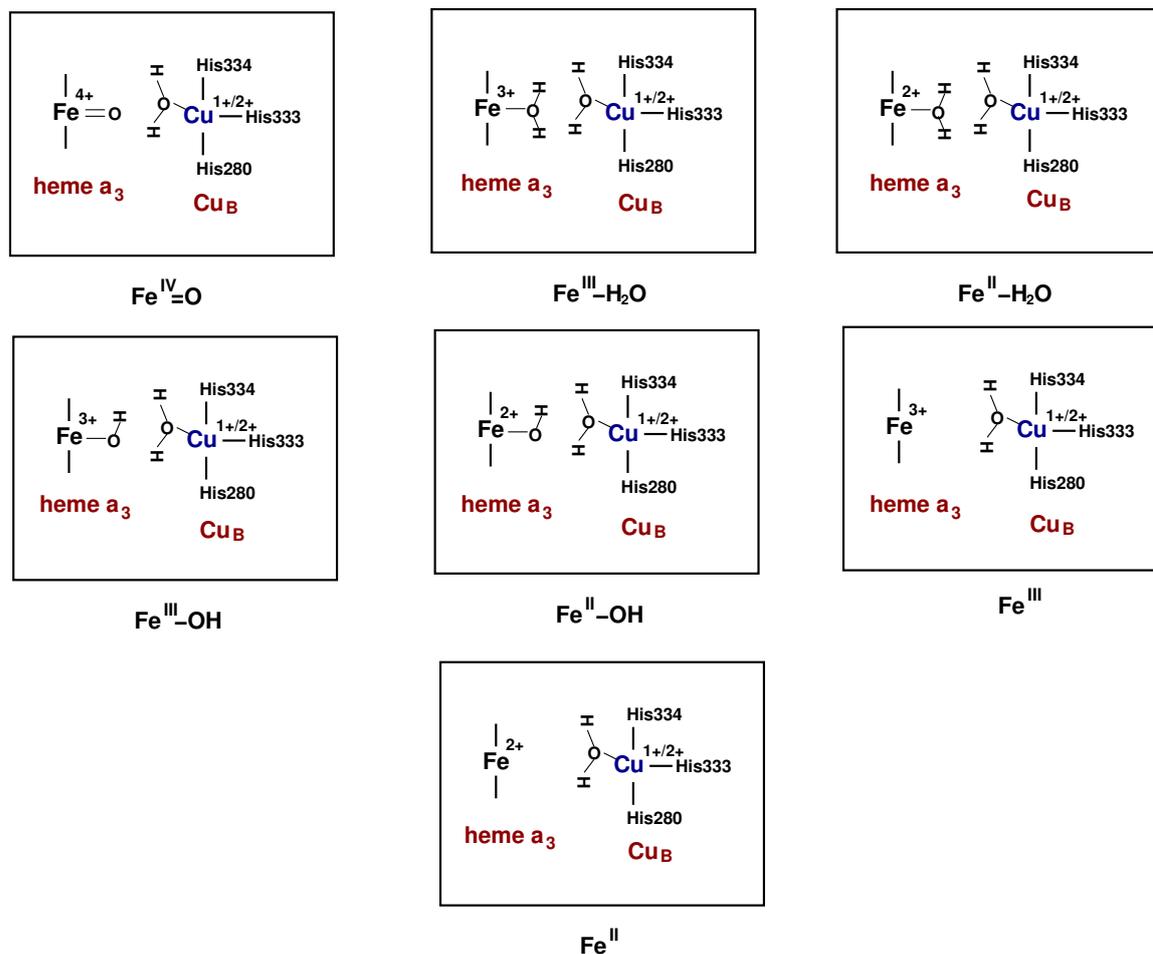
done with a spin-unrestricted scheme. The optimization were performed by the quasi Newton method and the Hessian was updated with the Broyden-Fletcher-Goldfarb-Shanno strategy.

The distributions of partial atomic point charges were computed by fitting the molecular electrostatic potentials calculated by ADF 2004.01 program [173]. The program `chargefit` was used to obtain point charges which is based on CHELPG algorithm [174]. The net charge of the molecule and the three Cartesian dipole moment components from PW91 calculations were adopted as constraints for the chargefit. Point charges were then computed by determining the electrostatic potential, how well the point charges effectively reproduce the electrostatic potential. The ESP charges were calculated on the cubic grid with uniform spacing of 0.2 and 3 Å outer boundary around each atom of the molecule. The atoms were assigned the Bondi values of 1.7 for carbon, 1.2 for hydrogen, 1.55 for nitrogen, 1.5 for oxygen, 1.3 for iron and 1.8 for sulfur. To minimize the uncertainties in the fitting procedure the single value decomposition (SVD) method [209] was used to obtain a model with stable atomic charges. Partial atomic charges for the amino acids are taken from the CHARMM22 parameter set. The partial atomic charges for all the redox centers were derived from PW91 calculations. In addition to atomic coordinates and atomic charges, electrostatic calculations require radii. The radii used were taken from Bondi [175].

All microscopic states of  $\text{Cu}_B$  center were optimized with hybrid density functional calculations (B3LYP) using 6-31G\* basis sets. The geometry optimizations were performed using GAUSSIAN 03 program. The single point calculations were performed on the optimized structures of the  $\text{Cu}_B$  center using B3LYP/6-31G\*/C-PCM level. The electrostatic potential obtained from C-PCM self-consistent reaction field were fitted by Merz-Kollman method to obtain the point charges.

## 6.4 ELECTROSTATIC CALCULATIONS

The electrostatic calculations were performed on  $\text{Fe}a_3\text{-Cu}_B$  complexes (see Figure 6.7). The  $\text{Cu}_B$  center was allowed to adopt each of the eight states shown in Figure 5.3 and 5.4 respectively. Following seven possible states of heme  $a_3$  center were considered:  $\text{Fe}^{\text{IV}}=\text{O}$ ,  $\text{Fe}^{\text{III}}\text{-H}_2\text{O}$ ,  $\text{Fe}^{\text{II}}\text{-H}_2\text{O}$ ,  $\text{Fe}^{\text{III}}\text{-OH}$ ,  $\text{Fe}^{\text{II}}\text{-OH}$ ,  $\text{Fe}^{\text{III}}$  and  $\text{Fe}^{\text{II}}$ . These seven states with different  $\text{Cu}_B$  center are shown in Figure 6.7. The  $\text{Cu}_A$  center and heme  $a$  were considered oxidized throughout the studies. Continuum electrostatic calculations were performed with the QMPB which solve the linear Poisson-Boltzmann equation by numerical finite difference method. The whole system is divided into three regions with three different dielectric constant of  $\epsilon_i = 1$  for the active site (quantum region) in this case the  $\text{Cu}_B$  center,  $\epsilon_s = 80$  for the solvent region, and  $\epsilon_p = 4$  for the protein. Compared to the purely electrostatic dielectric constant  $\epsilon = 2$ ,  $\epsilon_p = 4$  was adopted for the protein, which allows some mobility of the protein dipoles and accounts for some re-orientational relaxation of the protein in an approximate way. The ionic strength was set to 0.1 mol/l and the absolute temperature is set to 300 K. The boundary between the interior and the exterior is defined as the solvent contact and re-entrant surfaces of 1.4 Å spherical solvent probe rolling over the van der Waals surface of the protein. The atomic radii for the protein's atom and the charges for the protein's non-titrating atoms were taken from the polar hydrogen parameter set of CHARMM22. Electrostatic potential was calculated by first focusing the grid on the protein model compound and the second grid was centered on the titratable



**Figure 6.7. The  $\text{Fea}_3\text{-Cu}_B$  complexes considered in the present study.** The  $\text{Fea}_3\text{-Cu}_B$  complexes considered for the present study are shown. The nomenclature of the corresponding complexes are given in the bottom.

group. A coarse grid with 1 Å grid spacing and a finer grid with 0.25 Å grid spacing were used. Aspartates, glutamates, lysine, histidines (two sites for each histidines) cystein, tyrosine and N- and C-termini are treated as titratable groups.  $\text{p}K_a^{\text{model}}$  values of the titratable groups are given in Table 6.1. The  $\text{p}K_a$  values calculated for the  $\text{Cu}_B$  center both in the reduced and in the oxidized states in the aqueous solution are taken as a model  $\text{p}K_a$  values of the  $\text{Cu}_B$  center in protein.

#### 6.4.1 CALCULATION OF AVERAGE $\text{p}K_a$ IN PROTEIN

The electrostatic calculations were performed on the protein structure of *R. sphaeroides*. The electrostatic energies obtained from the QMPB calculations were decomposed into Born, background and the interaction energies. The difference in the Born energy ( $\Delta\Delta G^{\text{Born}}$ ) (see Eq. 2.36) and the background energy ( $\Delta\Delta G^{\text{back}}$ ) (see Eq. 2.37) for a protonation reaction of

| Titratable group         | p <i>K</i> <sub>a</sub> <sup>model</sup> |
|--------------------------|--|
| aspartate                | 4.0                                      |
| glutamate                | 4.4                                      |
| arginine                 | 10.4                                     |
| lysine                   | 10.4                                     |
| histidine N <sub>δ</sub> | 6.6                                      |
| histidine N <sub>ε</sub> | 7.0                                      |
| tyrosine                 | 9.6                                      |
| cysteine                 | 9.5                                      |
| C-terminus               | 3.8                                      |
| N-terminus               | 7.5                                      |

**Table 6.1.** p*K*<sub>a</sub><sup>model</sup> values of the titratable groups. The p*K*<sub>a</sub><sup>model</sup> values are taken from Ref. [70] and Ref. [71].

the site can be directly obtained from the electrostatic calculations (see section 2.2.2). The average p*K*<sub>a</sub> values of Cu<sub>B</sub> ligand in protein were obtained from Eq. (6.1).

$$pK_{a,i} = pK_{a,i}^{intr} + \sum_{j=1}^N (\langle x_j \rangle - x_j^{(0)}) W_{ij} \quad (6.1)$$

where the intrinsic p*K*<sub>a</sub> value (*pK*<sub>a,*i*<sup>intr</sup>) is the p*K*<sub>a</sub> value that the particular titratable group would have, if all other titratable groups are in their reference form. This term includes the solvation energy and the interaction with non-titrating residues and the protein backbone. The term *W*<sub>*ij*</sub> represents the interaction energy between the titratable groups *i* and *j* in their charged form;  $\langle x_j \rangle$  represents the protonation probability of the group *j* which is obtained by a thermodynamic average over all possible protonation states and *x*<sub>*j*</sub><sup>(0)</sup> is the reference protonation form of site *j*. The  $\langle x_j \rangle$  are obtained from MC calculations. These average p*K*<sub>a</sub> values do not represent an equilibrium situation, but they are a good approximation of the real p*K*<sub>a</sub> value of the site.</sub>

The protonation probabilities were obtained by Metropolis Monte Carlo calculations. The titration curves were calculated by Monte Carlo program, GMCT developed in our group. The protonation probabilities were computed at pH 7. The temperature was set to 300 K. The double flip and triple flips were set to 2 and 3 pH units respectively. The number of full MC scans were set to 30,000 and 100 equilibrium MC scans were performed before the MC full scan.

## 6.5 RESULTS AND DISCUSSION

The  $pK_a$  values calculated for the Cu-bound ligands in various  $Cu_B$  and heme  $a_3$  complexes in cytochrome *c* oxidase are discussed in this section. The  $pK_a$  values discussed in this section are average  $pK_a$  values of the deprotonation reactions at pH=7. The effect of the protein environment on the  $Cu_B$  ligands were studied by combining the DFT and continuum electrostatic calculations. The  $pK_a$  values obtained for the  $Cu_B$  center both in oxidized and reduced state in aqueous solution are discussed in detail in chapter 5.

$pK_a$  VALUES OF MICROSCOPIC STATES OF  $Cu_B$  LIGANDS IN CYTOCHROME *c* OXIDASE. The  $pK_a$  values of all possible proton equilibrium reactions of oxidized ( $Cu_B^{+2}$ ) and reduced ( $Cu_B^{+1}$ )  $Cu_B$  center in the presence of different heme  $a_3$  states of cytochrome *c* oxidase are given in Table 6.2 and 6.3 respectively.

The  $pK_a$  value of  $H_2O$  ligand was calculated both in oxidized and reduced states of the  $Cu_B$  center. There is no experimental  $pK_a$  value available for  $H_2O$  ligand in the  $Cu_B$  center. The  $pK_a$  values of  $H_2O$  ligand obtained in the aqueous solution are used as model  $pK_a$  values. Compared to the  $pK_a$  values of the  $H_2O$  in aqueous solution, the protein environment shifts the  $pK_a$  values of the  $H_2O$  both in oxidized and reduced state of  $Cu_B$  center to higher values. The  $pK_a$  values are very high for the reduced state of the  $Cu_B$  center compared to the oxidized state. The  $pK_a$  values are in the range of 51-60 which show that the deprotonation of the  $H_2O$  in the reduced state is highly unfavorable. The  $pK_a$  values calculated in the  $Fe^{II}$ -OH state are obviously high due to -1 charge on the heme  $a_3$  center. The  $pK_a$  values increase with the change in charge state of heme  $a_3$  center from  $+1 \rightarrow 0 \rightarrow -1$ . The high  $pK_a$  values for deprotonation of  $H_2O$  show that it is protonated within the physiological pH.

There is a possibility that His333 another ligand of  $Cu_B$  center, can serve as a proton loading site for subsequent pumping. The  $pK_a$  values of His333 ligand obtained in the aqueous solution is used as model  $pK_a$  values. There is no experimental  $pK_a$  value available for His333 ligand in the  $Cu_B$  center. The low dielectric protein environment shifts the  $pK_a$  values of the His333 both in oxidized and reduced state of  $Cu_B$  center to higher values compared to results in aqueous solution. The experimental  $pK_a$  value for the deprotonation of imidazole, which produces an anionic imidazole, is approximately 14 [193, 194]. In present calculations, the  $pK_a$  values for His333 deprotonation are in the range of 37-48 in the oxidized state and 44.5-53.8 in the reduced state. The high  $pK_a$  values in the protein show that His333 is likely to be protonated within the physiological pH both in the oxidized and in the reduced state of the  $Cu_B$  center. The protonated  $N_{\delta 1}$  of His333 is hydrogen bonded to Thr352. As a result, the deprotonation of His333 may be significantly hindered.

The  $pK_a$  value of His334 obtained both in aqueous solution and protein ( $\langle pK_{a,i} \rangle_{prot}$ ) with different heme  $a_3$  ligand are given in Table 6.2 and 6.3 respectively. The  $pK_a$  values of His334 ligand obtained in the aqueous solution are used as model  $pK_a$  values. There is no experimental  $pK_a$  value available for His334 ligand in the  $Cu_B$  center. The protein environment shifts most of the  $pK_a$  values of the His334 both in oxidized and reduced state of  $Cu_B$  center. The  $pK_a$  values are shifted to higher values in the oxidized state of the  $Cu_B$  center and are in the range of 27.5-39. The  $pK_a$  values are even shifted to higher values in the reduced state of the  $Cu_B$  center. The  $pK_a$  value of 33.5 is obtained in the case of oxo-ferryl state ( $Fe^{IV} = O$ ) whereas

| Step | Rea. State                         | Prod. State                        | p <i>K<sub>a</sub></i> <sup>model</sup> | ⟨p <i>K<sub>a,i</sub></i> ⟩ <sub>prot</sub> at pH=7 |                   |                  |                                    |                       |                     |                      |
|------|------------------------------------|------------------------------------|---|---|-------------------|------------------|------------------------------------|-----------------------|---------------------|----------------------|
|      |                                    |                                    |   | Fe <sup>III</sup> -H <sub>2</sub> O                 | Fe <sup>III</sup> | Fe <sup>II</sup> | Fe <sup>II</sup> -H <sub>2</sub> O | Fe <sup>III</sup> -OH | Fe <sup>IV</sup> =O | Fe <sup>II</sup> -OH |
|      |                                    |                                    |   | [+1]  | [+1]              | [0]              | [0]                                | [0]                   | [0]                 | [-1]                 |
| 1    | Cu <sub>B</sub> <sup>+2</sup> -111 | Cu <sub>B</sub> <sup>+2</sup> -011 | 7.4                                     | 15.9  | 15.7              | 19.9             | 20.1                               | 21.4                  | 21.5                | 26.3                 |
|      | Cu <sub>B</sub> <sup>+2</sup> -111 | Cu <sub>B</sub> <sup>+2</sup> -101 | 15.9                                    | 37.2  | 36.7              | 41.6             | 42.1                               | 42.6                  | 42.3                | 48.2                 |
|      | Cu <sub>B</sub> <sup>+2</sup> -111 | Cu <sub>B</sub> <sup>+2</sup> -110 | 15.2                                    | 27.8  | 27.5              | 32.7             | 32.9                               | 33.6                  | 33.5                | 39.4                 |
| 2a   | Cu <sub>B</sub> <sup>+2</sup> -110 | Cu <sub>B</sub> <sup>+2</sup> -010 | 13.0                                    | 28.8  | 28.6              | 32.6             | 32.9                               | 34.1                  | 34.0                | 38.7                 |
|      | Cu <sub>B</sub> <sup>+2</sup> -101 | Cu <sub>B</sub> <sup>+2</sup> -001 | 10.6                                    | 17.9  | 17.7              | 21.7             | 22.0                               | 23.0                  | 22.9                | 27.7                 |
| 2b   | Cu <sub>B</sub> <sup>+2</sup> -110 | Cu <sub>B</sub> <sup>+2</sup> -100 | 18.1                                    | 47.9  | 47.4              | 52.3             | 2.8                                | 53.1                  | 52.8                | 58.7                 |
|      | Cu <sub>B</sub> <sup>+2</sup> -011 | Cu <sub>B</sub> <sup>+2</sup> -001 | 19.1                                    | 39.2  | 38.7              | 43.5             | 4.0                                | 44.1                  | 43.8                | 49.6                 |
| 2c   | Cu <sub>B</sub> <sup>+2</sup> -011 | Cu <sub>B</sub> <sup>+2</sup> -010 | 20.8                                    | 40.7  | 40.4              | 45.4             | 5.7                                | 46.2                  | 46.0                | 51.9                 |
|      | Cu <sub>B</sub> <sup>+2</sup> -101 | Cu <sub>B</sub> <sup>+2</sup> -100 | 17.4                                    | 38.5  | 38.2              | 43.3             | 3.6                                | 44.2                  | 44.0                | 49.9                 |
| 3    | Cu <sub>B</sub> <sup>+2</sup> -100 | Cu <sub>B</sub> <sup>+2</sup> -000 | 14.3                                    | 27.7  | 27.5              | 31.8             | 32.0                               | 33.1                  | 33.0                | 37.9                 |
|      | Cu <sub>B</sub> <sup>+2</sup> -010 | Cu <sub>B</sub> <sup>+2</sup> -000 | 19.4                                    | 46.9  | 46.4              | 51.4             | 51.9                               | 52.1                  | 51.8                | 57.8                 |
|      | Cu <sub>B</sub> <sup>+2</sup> -001 | Cu <sub>B</sub> <sup>+2</sup> -000 | 21.0                                    | 48.4  | 48.1              | 53.3             | 53.6                               | 54.3                  | 54.1                | 60.0                 |

**Table 6.2.** The p*K<sub>a</sub>* values of all possible proton equilibrium reactions of oxidized (Cu<sub>B</sub><sup>+2</sup>) Cu<sub>B</sub> center in the presence of difference heme *a*<sub>3</sub> states of cytochrome *c* oxidase.

The p*K<sub>a</sub>* values of all possible proton equilibrium reactions of oxidized (Cu<sub>B</sub><sup>+2</sup>) Cu<sub>B</sub> center in aqueous solution (p*K<sub>a</sub>*<sup>model</sup>) as well as in the protein (⟨p*K<sub>a,i</sub>*⟩<sub>prot</sub>) are given. The total charge of the heme *a*<sub>3</sub> states are given in square brackets. The reactions are classified according to proton bound to each microscopic state. ‘Step 1’ corresponds to the removal of single proton from H<sub>2</sub>O, His333 and His334. The second proton is removed from H<sub>2</sub>O in ‘step 2a’ and in ‘step 2b’ from His333 and from His334 in ‘step 2c’. In ‘Step 3’ the third proton is removed from H<sub>2</sub>O, His333 and His334. Each of these protonation states can be described by a protonation state vectors in which ‘1’ indicates that the proton is bound and ‘0’ indicates that the proton is not bound. The order of state vector is H<sub>2</sub>O, His333, His334. For example state vector ‘111’ denotes that all the sites (H<sub>2</sub>O, His333, His334) are protonated and state vector ‘011’ represents deprotonated H<sub>2</sub>O and protonated His333 and His334 respectively.

Stuchebrukhov et al. [3] obtained the p*K<sub>a</sub>* value of the 5.4 for oxidized Cu<sub>B</sub> center and 26.2 for the reduced Cu<sub>B</sub> center [3]. The present calculations show that there is no downshift in the p*K<sub>a</sub>* values (experimental p*K<sub>a</sub>* value for the deprotonation of imidazole, which produces an anionic imidazole is approximately 14) of the His334 depending on the redox state of the Cu<sub>B</sub> center. High p*K<sub>a</sub>* values are observed for the reactions corresponding to double and triple deprotonation reactions. These high p*K<sub>a</sub>* values show that such reactions are unfavored in the protein environment. The p*K<sub>a</sub>* values obtained for the Cu<sub>B</sub> ligands in the protein show that H<sub>2</sub>O, His333 and His334 is likely to be protonated within the physiological pH.

Within the framework of a recently proposed pumping mechanism of cytochrome *c* oxidase, [1–3, 5, 6] the deprotonation of the Cu<sub>B</sub> ligand His334 is presumed to be triggered by the entry of chemical proton in the binuclear center, which converts a hydroxyl ligand into a water molecule [1, 2]. According to this mechanism the repulsion between the chemical proton and the proton on His334 is strong enough for the deprotonation of His334 and that act as a proton loading site in the pumping mechanism of cytochrome *c* oxidase, whereas present

| Step | Rea. State                         | Prod. State                        | $pK_a^{model}$ | $\langle pK_{a,i} \rangle_{prot}$ at pH=7 |                   |                  |   |                       |                     |                                   |
|------|------------------------------------|------------------------------------|----------------|---|-------------------|------------------|---|-----------------------|---------------------|-----------------------------------|
|      |                                    |                                    |                | Fe <sup>III</sup> -H <sub>2</sub> O       | Fe <sup>III</sup> | Fe <sup>II</sup> | Fe <sup>II</sup> -H <sub>2</sub> O <sup>a</sup> | Fe <sup>III</sup> -OH | Fe <sup>IV</sup> =O | Fe <sup>II</sup> -OH <sup>a</sup> |
|      |                                    |                                    |                | [+1]                                      | [+1]              | [0]              | [0]   | [0]                   | [0]                 | [-1]                              |
| 1    | Cu <sub>B</sub> <sup>+1</sup> -111 | Cu <sub>B</sub> <sup>+1</sup> -011 | 32.5           | 51.4                                      | 51.1              | 54.9             | 55.3  | 56.5                  | 56.4                | 59.9                              |
|      | Cu <sub>B</sub> <sup>+1</sup> -111 | Cu <sub>B</sub> <sup>+1</sup> -101 | 17.7           | 45.0                                      | 44.5              | 49.0             | 49.5  | 49.7                  | 49.3                | 53.8                              |
|      | Cu <sub>B</sub> <sup>+1</sup> -111 | Cu <sub>B</sub> <sup>+1</sup> -110 | 15.9           | 41.4                                      | 41.0              | 45.9             | 46.3  | 46.6                  | 46.4                | 50.0                              |
| 2a   | Cu <sub>B</sub> <sup>+1</sup> -101 | Cu <sub>B</sub> <sup>+1</sup> -001 | 36.2           | 55.3                                      | 55.1              | 58.9             | 59.2  | 60.4                  | 60.4                | 63.8                              |
| 2b   | Cu <sub>B</sub> <sup>+1</sup> -110 | Cu <sub>B</sub> <sup>+1</sup> -100 | 21.3           | 49.1                                      | 48.4              | 53.4             | 54.0  | 54.3                  | 54.0                | 58.9                              |
|      | Cu <sub>B</sub> <sup>+1</sup> -011 | Cu <sub>B</sub> <sup>+1</sup> -001 | 21.1           | 48.9                                      | 48.5              | 53.0             | 53.4  | 53.6                  | 53.3                | 57.8                              |
| 2c   | Cu <sub>B</sub> <sup>+1</sup> -101 | Cu <sub>B</sub> <sup>+1</sup> -100 | 19.5           | 45.5                                      | 45.0              | 50.3             | 50.78   | 51.2                  | 50.9                | 55.1                              |
| 3    | Cu <sub>B</sub> <sup>+1</sup> -100 | Cu <sub>B</sub> <sup>+1</sup> -000 | 48.4           | 67.0                                      | 67.1              | 70.8             | 70.8  | 71.9                  | 72.1                | 74.9                              |
|      | Cu <sub>B</sub> <sup>+1</sup> -001 | Cu <sub>B</sub> <sup>+1</sup> -000 | 31.7           | 57.2                                      | 57.0              | 62.2             | 62.4  | 62.7                  | 62.6                | 66.2                              |

**Table 6.3. The  $pK_a$  values of all possible proton equilibrium reactions of reduced (Cu<sub>B</sub><sup>+1</sup>) Cu<sub>B</sub> center in the presence of difference heme  $a_3$  states of cytochrome  $c$  oxidase.** The  $pK_a$  values of all possible proton equilibrium reactions of oxidized (Cu<sub>B</sub><sup>+1</sup>) Cu<sub>B</sub> center in aqueous solution ( $pK_a^{model}$ ) as well as in the protein ( $\langle pK_{a,i} \rangle_{prot}$ ) are given. The total charge of the heme  $a_3$  states are given in square brackets. The reactions are classified according to proton bound to each microscopic state. ‘Step 1’ corresponds to the removal of single proton from H<sub>2</sub>O, His333 and His334. The second proton is removed from H<sub>2</sub>O in ‘step 2a’ and in ‘step 2b’ from His333 and from His334 in ‘step 2c’. In ‘Step 3’ the third proton is removed from H<sub>2</sub>O, His333 and His334. Each of these protonation states can be described by a protonation state vectors in which ‘1’ indicates that the proton is bound and ‘0’ indicates that the proton is not bound. The order of state vector is H<sub>2</sub>O, His333, His334. For example state vector ‘111’ denotes that all the sites (H<sub>2</sub>O, His333, His334) are protonated and state vector ‘011’ represents deprotonated H<sub>2</sub>O and protonated His333 and His334 respectively.

study show that the protein environment shifts the  $pK_a$  values of the Cu<sub>B</sub> ligands to higher values both in oxidized and reduced state of Cu<sub>B</sub> center. The higher  $pK_a$  values show that the His333 and His334 protonated forms are favored in protein environment.

The  $pK_a$  values of H<sub>2</sub>O and His334 reported in previous work and from the present study are given in Table 6.4. The protonation probabilities of titratable residues of cytochrome  $c$  oxidase in different heme  $a_3$  redox states at pH=7 are given in Table 6.5. Stuchebrukhov et al. obtained [3] obtained a  $pK_a$  value of 8.6 and Fadda et al. [182] obtained a  $pK_a$  value of 13.5 for the deprotonation of His334 in the aqueous solution when Cu<sub>B</sub> center is oxidized. In present study, a  $pK_a$  value of 15.2 is obtained for the deprotonation of His334 in the aqueous solution when Cu<sub>B</sub> center is oxidized. A  $pK_a$  value of 14.4 was obtained by Fadda et al. and 13.2 was obtained by Stuchebrukhov et al. for the deprotonation of His334 when Cu<sub>B</sub> center is reduced. The  $pK_a$  value obtained from the present study is 15.9 which is  $\sim 2$   $pK_a$  units higher than previously reported values. All the results show that the deprotonation of His334 is not favored in the reduced state of Cu<sub>B</sub> center.

Stuchebrukhov et al. obtained the  $pK_a$  value of the 5.4 when Cu<sub>B</sub> center is oxidized and  $pK_a$  value of 21.9 when the Cu<sub>B</sub> center is reduced. Based upon these calculations they proposed that the  $pK_a$  values of His334 depends on the redox states of heme  $a_3$ -Cu<sub>B</sub> and

| Reac. State          | Prod. State     | Redox State of heme $a_3$ | Struchebukhov et al. <sup>a</sup> | Fadda et al. <sup>b</sup> | Present Study |                      |
|----------------------|-----------------|---------------------------|-----------------------------------|---------------------------|---------------|----------------------|
| In aqueous solution: |                 |                           |                                   |                           |               |                      |
| $Cu_B^{+2}-111$      | $Cu_B^{+2}-011$ |                           | –                                 | –                         | 7.4           |                      |
| $Cu_B^{+2}-111$      | $Cu_B^{+2}-110$ |                           | 8.6 <sup>c</sup>                  | 13.5                      | 15.2          |                      |
| $Cu_B^{+1}-111$      | $Cu_B^{+1}-110$ |                           | 13.2 <sup>c</sup>                 | 14.4                      | 15.9          |                      |
| $Cu_B^{+1}-011$      | $Cu_B^{+1}-010$ |                           | –                                 | 14.4                      | 20.8          |                      |
| In protein:          |                 |                           |                                   |                           |               |                      |
| $Cu_B^{+2}-111$      | $Cu_B^{+2}-011$ | $Fe^{IV}=O$               | 8.5                               | –                         | 21.5          |                      |
| $Cu_B^{+2}-011$      | $Cu_B^{+2}-010$ | $Fe^{IV}=O$               | –                                 | –                         | 46.0          | <b>F</b>             |
| $Cu_B^{+2}-111$      | $Cu_B^{+2}-110$ | $Fe^{IV}=O$               | 5.4                               | 16.0                      | 33.5          | <b>F<sub>p</sub></b> |
| $Cu_B^{+1}-111$      | $Cu_B^{+1}-110$ | $Fe^{IV}=O$               | 21.9                              | –                         | 46.4          | <b>H'</b>            |
| $Cu_B^{+2}-011$      | $Cu_B^{+2}-010$ | $Fe^{III}-OH$             | –                                 | –                         | 46.2          | <b>H</b>             |
| $Cu_B^{+2}-011$      | $Cu_B^{+2}-010$ | $Fe^{III}-H_2O$           | –                                 | 14.9                      | 40.7          | <b>O</b>             |
| $Cu_B^{+2}-111$      | $Cu_B^{+2}-110$ | $Fe^{III}$                | –                                 | –                         | 27.5          | <b>O~</b>            |
| $Cu_B^{+2}-011$      | $Cu_B^{+2}-010$ | $Fe^{III}$                | –                                 | 16.2                      | 40.4          | <b>O<sub>H</sub></b> |
| $Cu_B^{+2}-011$      | $Cu_B^{+2}-010$ | $Fe^{II}$                 | –                                 | 24.9                      | 45.4          | <b>E</b>             |
| $Cu_B^{+2}-111$      | $Cu_B^{+2}-110$ | $Fe^{II}$                 | –                                 | –                         | 32.7          | <b>E<sub>p</sub></b> |
| $Cu_B^{+1}-111$      | $Cu_B^{+1}-110$ | $Fe^{II}$                 | –                                 | –                         | 45.9          | <b>R</b>             |

**Table 6.4. The  $pK_a$  values reported in previous work and the  $pK_a$  values obtained for catalytic intermediates from present studies are given.**

<sup>a</sup> Ref [3] <sup>b</sup> Ref [183] <sup>c</sup> Ref [6]

thus have a key role in the proton pumping mechanism of cytochrome *c* oxidase. Fadda et al. obtained a  $pK_a$  value of 16.0 for the deprotonation of His334 in the protein. From our present calculations, we obtained  $pK_a$  values of 33.5 and 46.4 for the deprotonation of His334 when the  $Cu_B$  in oxidized and reduced state respectively. The  $pK_a$  values of 14.9, 16.2 and 24.9 were obtained for deprotonation of His334 in O, O<sub>H</sub> and E states by Fadda et al. and the  $pK_a$  values show that His334 is protonated in all these states. In present study we obtained very high  $pK_a$  values for the same states showing that His334 deprotonation reactions are unfavored in protein. Interestingly the  $pK_a$  values of the His334 both in oxidized and reduced state of the  $Cu_B$  center show significant increase in the protein environment. The high  $pK_a$  values of His334 in the protein show that His334 remains protonated irrespective of  $Cu_B$  and heme  $a_3$  redox states. The  $pK_a$  values of His334 are analyzed within the framework of the cytochrome *c* oxidase catalytic cycle shown in Figure 6.1. The first deprotonation of His334 is expected after the entry of the chemical proton in F to form F<sub>p</sub> state. A  $pK_a$  value of 46 is obtained for the deprotonation of His334 from F→F<sub>p</sub> state transition. The high  $pK_a$  value of His334 show that the His334 is protonated when a chemical proton arrives the  $Fea_3-Cu_B$  center which does not support the proposed His334 model as a proton loading site for the proton pumping mechanism of cytochrome *c* oxidase [1, 2].

**Table 6.5. The protonation probabilities ( $\langle x \rangle$ ) of titratable residues of cytochrome *c* oxidase in different heme  $a_3$  redox states at pH=7 are given.**

| Subunit | Residue | $\langle x \rangle$ at pH=7                 |                           |                         |   |                              |                            |                              |
|---------|---------|---|---------------------------|-------------------------|---|------------------------------|----------------------------|------------------------------|
|         |         | Fe <sup>III</sup> -H <sub>2</sub> O<br>[+1] | Fe <sup>III</sup><br>[+1] | Fe <sup>II</sup><br>[0] | Fe <sup>II</sup> -H <sub>2</sub> O<br>[0] | Fe <sup>III</sup> -OH<br>[0] | Fe <sup>IV</sup> =O<br>[0] | Fe <sup>II</sup> -OH<br>[-1] |
| I       | ARG-14  | 1.000                                       | 1.000                     | 1.000                   | 1.000                                     | 1.000                        | 1.000                      | 1.000                        |
|         | ARG-19  | 1.000                                       | 1.000                     | 1.000                   | 1.000                                     | 1.000                        | 1.000                      | 1.000                        |
|         | TRP-20  | 1.000                                       | 1.000                     | 1.000                   | 1.000                                     | 1.000                        | 1.000                      | 1.000                        |
|         | HIS-26  | 0.006                                       | 0.007                     | 0.006                   | 0.010                                     | 0.010                        | 0.010                      | 0.011                        |
|         | LYG-27  | 0.993                                       | 0.991                     | 0.993                   | 0.988                                     | 0.989                        | 0.990                      | 0.988                        |
|         | ASP-28  | 0.000                                       | 0.000                     | 0.000                   | 0.000                                     | 0.000                        | 0.000                      | 0.000                        |
|         | TYR-33  | 1.000                                       | 1.000                     | 1.000                   | 1.000                                     | 1.000                        | 1.000                      | 1.000                        |
|         | TYR-50  | 1.000                                       | 1.000                     | 1.000                   | 1.000                                     | 1.000                        | 1.000                      | 1.000                        |
|         | ARG-52  | 0.168                                       | 0.162                     | 0.666                   | 0.676                                     | 0.666                        | 0.659                      | 0.955                        |
|         | CYS-64  | 1.000                                       | 1.000                     | 1.000                   | 1.000                                     | 1.000                        | 1.000                      | 1.000                        |
|         | GLU-66  | 0.007                                       | 0.008                     | 0.005                   | 0.005                                     | 0.006                        | 0.007                      | 0.007                        |
|         | HIS-67  | 0.942                                       | 0.942                     | 0.929                   | 0.935                                     | 0.927                        | 0.935                      | 0.920                        |
|         | GLU-69  | 0.007                                       | 0.008                     | 0.011                   | 0.009                                     | 0.008                        | 0.010                      | 0.009                        |
|         | LYS-74  | 0.000                                       | 0.000                     | 0.000                   | 0.000                                     | 0.000                        | 0.000                      | 0.000                        |
|         | TRP-81  | 1.000                                       | 1.000                     | 1.000                   | 1.000                                     | 1.000                        | 1.000                      | 1.000                        |
|         | GLU-86  | 0.231                                       | 0.240                     | 0.221                   | 0.231                                     | 0.228                        | 0.226                      | 0.221                        |
|         | CYS-88  | 1.000                                       | 1.000                     | 1.000                   | 1.000                                     | 1.000                        | 1.000                      | 1.000                        |
|         | HIS-93  | 0.286                                       | 0.294                     | 0.325                   | 0.326                                     | 0.339                        | 0.322                      | 0.382                        |
|         | TRP-95  | 1.000                                       | 1.000                     | 1.000                   | 1.000                                     | 1.000                        | 1.000                      | 1.000                        |
|         | HIS-102 | 1.000                                       | 1.000                     | 1.000                   | 1.000                                     | 1.000                        | 1.000                      | 1.000                        |
|         | TYR-122 | 1.000                                       | 1.000                     | 1.000                   | 1.000                                     | 1.000                        | 1.000                      | 1.000                        |
|         | HIS-127 | 0.000                                       | 0.000                     | 0.000                   | 0.000                                     | 0.000                        | 0.000                      | 0.000                        |
|         | ASP-132 | 0.003                                       | 0.003                     | 0.003                   | 0.004                                     | 0.003                        | 0.004                      | 0.006                        |
|         | ARG-137 | 1.000                                       | 1.000                     | 1.000                   | 1.000                                     | 1.000                        | 1.000                      | 1.000                        |
|         | TYR-143 | 1.000                                       | 1.000                     | 1.000                   | 1.000                                     | 1.000                        | 1.000                      | 1.000                        |
|         | TRP-144 | 1.000                                       | 1.000                     | 1.000                   | 1.000                                     | 1.000                        | 1.000                      | 1.000                        |

*continued on next page*

| Subunit | Residue | <i>x</i> at pH=7                    |                   |                  |                                    |                       |                     |                      |
|---------|---------|-------------------------------------|-------------------|------------------|------------------------------------|-----------------------|---------------------|----------------------|
|         |         | Fe <sup>III</sup> -H <sub>2</sub> O | Fe <sup>III</sup> | Fe <sup>II</sup> | Fe <sup>II</sup> -H <sub>2</sub> O | Fe <sup>III</sup> -OH | Fe <sup>IV</sup> =O | Fe <sup>II</sup> -OH |
|         |         | [+1]                                | [+1]              | [0]              | [0]                                | [0]                   | [0]                 | [-1]                 |
|         | TYR-146 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TRP-172 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TYR-175 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | GLU-182 | 0.051                               | 0.049             | 0.064            | 0.064                              | 0.066                 | 0.069               | 0.097                |
|         | TYR-185 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | ASP-188 | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
|         | HIS-195 | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
|         | ARG-216 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | HIS-223 | 0.000                               | 0.000             | 0.001            | 0.001                              | 0.000                 | 0.000               | 0.001                |
|         | LYS-224 | 0.999                               | 0.998             | 0.998            | 0.998                              | 0.999                 | 0.998               | 0.999                |
|         | TRP-230 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TRP-237 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | ASP-256 | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
|         | ARG-257 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | ASP-271 | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
|         | TYR-275 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | HIS-277 | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
|         | TRP-280 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | HIS-284 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | GLU-286 | 0.991                               | 0.994             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | HIS-300 | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.001               | 0.001                |
|         | LYS-307 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | LYS-308 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TYR-313 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TYR-318 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TRP-331 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TYR-336 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TYR-347 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | LYS-362 | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
|         | TRP-366 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TRP-371 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | GLU-376 | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |

*continued on next page*

| Subunit | Residue | $\langle x \rangle$ at pH=7         |                   |                  |                                    |                       |                     |                      |
|---------|---------|-------------------------------------|-------------------|------------------|------------------------------------|-----------------------|---------------------|----------------------|
|         |         | Fe <sup>III</sup> -H <sub>2</sub> O | Fe <sup>III</sup> | Fe <sup>II</sup> | Fe <sup>II</sup> -H <sub>2</sub> O | Fe <sup>III</sup> -OH | Fe <sup>IV</sup> =O | Fe <sup>II</sup> -OH |
|         |         | [+1]                                | [+1]              | [0]              | [0]                                | [0]                   | [0]                 | [-1]                 |
|         | LYS-378 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TRP-383 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | ASP-407 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | ARG-408 | 0.979                               | 0.980             | 0.995            | 0.996                              | 0.995                 | 0.995               | 0.999                |
|         | TYR-409 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TYR-410 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TYR-414 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TYR-415 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TYR-437 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TRP-439 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | LYS-442 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | ARG-446 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TYR-448 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | GLU-450 | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
|         | TRP-451 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | LYS-454 | 0.996                               | 0.997             | 0.997            | 0.997                              | 0.997                 | 0.997               | 0.997                |
|         | HIS-456 | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
|         | TRP-458 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | HIS-472 | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
|         | ARG-476 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | ARG-481 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | ARG-482 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TYR-483 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | ASP-485 | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
|         | TYR-486 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | GLU-488 | 0.008                               | 0.007             | 0.007            | 0.006                              | 0.008                 | 0.004               | 0.004                |
|         | TRP-493 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TYR-517 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | ARG-521 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | ARG-524 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TYR-530 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TRP-531 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |

*continued on next page*

| Subunit | Residue               | $\langle x \rangle$ at pH=7         |                   |                  |                                    |                       |                     |                      |
|---------|-----------------------|-------------------------------------|-------------------|------------------|------------------------------------|-----------------------|---------------------|----------------------|
|         |                       | Fe <sup>III</sup> -H <sub>2</sub> O | Fe <sup>III</sup> | Fe <sup>II</sup> | Fe <sup>II</sup> -H <sub>2</sub> O | Fe <sup>III</sup> -OH | Fe <sup>IV</sup> =O | Fe <sup>II</sup> -OH |
|         |                       | [+1]                                | [+1]              | [0]              | [0]                                | [0]                   | [0]                 | [-1]                 |
|         | GLU-533               | 0.002                               | 0.001             | 0.002            | 0.002                              | 0.002                 | 0.002               | 0.002                |
|         | HIS-534               | 0.411                               | 0.409             | 0.436            | 0.437                              | 0.426                 | 0.434               | 0.461                |
|         | ASP-536               | 0.001                               | 0.003             | 0.003            | 0.002                              | 0.003                 | 0.004               | 0.004                |
|         | GLU-539               | 0.001                               | 0.001             | 0.001            | 0.002                              | 0.002                 | 0.002               | 0.003                |
|         | TRP-540               | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | GLU-548               | 0.004                               | 0.003             | 0.003            | 0.003                              | 0.003                 | 0.003               | 0.003                |
|         | HIS-549               | 0.107                               | 0.118             | 0.123            | 0.119                              | 0.122                 | 0.117               | 0.128                |
|         | GLU-552               | 0.002                               | 0.003             | 0.004            | 0.004                              | 0.003                 | 0.004               | 0.003                |
|         | LYS-556               | 0.997                               | 0.998             | 0.999            | 0.997                              | 0.998                 | 0.997               | 0.999                |
|         | ARG-557               | 0.999                               | 1.000             | 0.999            | 0.999                              | 0.999                 | 0.999               | 1.000                |
|         | GLU-558               | 0.003                               | 0.003             | 0.003            | 0.003                              | 0.002                 | 0.002               | 0.002                |
|         | ASP-559               | 0.002                               | 0.002             | 0.002            | 0.001                              | 0.002                 | 0.001               | 0.002                |
|         | TRP-560               | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | HEA $\alpha$ -PROPA   | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
|         | HEA $\alpha$ -PROPD   | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
|         | HEA $\alpha_3$ -PROPA | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
|         | HEA $\alpha_3$ -PROPD | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
| II      | GLU-31                | 0.019                               | 0.015             | 0.016            | 0.015                              | 0.021                 | 0.016               | 0.015                |
|         | ARG-35                | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | HIS-55                | 0.004                               | 0.003             | 0.005            | 0.003                              | 0.005                 | 0.004               | 0.007                |
|         | TRP-56                | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | ASP-58                | 0.000                               | 0.000             | 0.000            | 0.001                              | 0.001                 | 0.000               | 0.005                |
|         | TYR-78                | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TRP-81                | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | ARG-82                | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | HIS-84                | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
|         | GLU-85                | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
|         | LYS-86                | 0.924                               | 0.921             | 0.923            | 0.925                              | 0.920                 | 0.925               | 0.924                |
|         | ARG-87                | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | LYS-89                | 0.366                               | 0.355             | 0.363            | 0.362                              | 0.356                 | 0.354               | 0.371                |
|         | ARG-93                | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | HIS-96                | 0.045                               | 0.046             | 0.048            | 0.051                              | 0.048                 | 0.053               | 0.052                |

*continued on next page*

| Subunit | Residue | $\langle x \rangle$ at pH=7         |                   |                  |                                    |                       |                     |                      |
|---------|---------|-------------------------------------|-------------------|------------------|------------------------------------|-----------------------|---------------------|----------------------|
|         |         | Fe <sup>III</sup> -H <sub>2</sub> O | Fe <sup>III</sup> | Fe <sup>II</sup> | Fe <sup>II</sup> -H <sub>2</sub> O | Fe <sup>III</sup> -OH | Fe <sup>IV</sup> =O | Fe <sup>II</sup> -OH |
|         |         | [+1]                                | [+1]              | [0]              | [0]                                | [0]                   | [0]                 | [-1]                 |
|         | GLU-101 | 1.000                               | 0.999             | 1.000            | 1.000                              | 0.999                 | 1.000               | 1.000                |
|         | TRP-104 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | GLU-128 | 0.011                               | 0.013             | 0.012            | 0.012                              | 0.014                 | 0.014               | 0.016                |
|         | GLU-131 | 0.011                               | 0.011             | 0.011            | 0.014                              | 0.009                 | 0.012               | 0.009                |
|         | ASP-133 | 0.001                               | 0.001             | 0.001            | 0.001                              | 0.001                 | 0.001               | 0.001                |
|         | LYS-137 | 0.999                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TYR-141 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TRP-143 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TYR-144 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TRP-145 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TYR-147 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | GLU-148 | 0.030                               | 0.027             | 0.031            | 0.034                              | 0.029                 | 0.031               | 0.031                |
|         | TYR-149 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | ASP-151 | 0.006                               | 0.008             | 0.006            | 0.009                              | 0.009                 | 0.009               | 0.008                |
|         | GLU-152 | 0.005                               | 0.005             | 0.006            | 0.005                              | 0.004                 | 0.007               | 0.006                |
|         | GLU-153 | 0.010                               | 0.010             | 0.009            | 0.008                              | 0.010                 | 0.010               | 0.012                |
|         | GLU-157 | 0.080                               | 0.080             | 0.080            | 0.086                              | 0.075                 | 0.078               | 0.081                |
|         | TYR-159 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | ASP-169 | 0.004                               | 0.004             | 0.005            | 0.006                              | 0.005                 | 0.005               | 0.005                |
|         | ARG-171 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | GLU-175 | 0.002                               | 0.003             | 0.002            | 0.002                              | 0.002                 | 0.001               | 0.002                |
|         | GLU-177 | 0.011                               | 0.009             | 0.009            | 0.010                              | 0.011                 | 0.012               | 0.007                |
|         | GLU-182 | 0.013                               | 0.013             | 0.011            | 0.012                              | 0.015                 | 0.017               | 0.015                |
|         | TYR-185 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | ARG-187 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | ASP-188 | 0.001                               | 0.001             | 0.001            | 0.001                              | 0.001                 | 0.001               | 0.001                |
|         | GLU-189 | 0.272                               | 0.271             | 0.262            | 0.263                              | 0.254                 | 0.268               | 0.274                |
|         | ASP-195 | 0.001                               | 0.000             | 0.000            | 0.001                              | 0.001                 | 0.001               | 0.001                |
|         | LYS-204 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | ASP-214 | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
|         | TRP-219 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | LYS-227 | 0.984                               | 0.984             | 0.996            | 0.997                              | 0.997                 | 0.997               | 1.000                |

*continued on next page*

| Subunit | Residue | $\langle x \rangle$ at pH=7         |                   |                  |                                    |                       |                     |                      |
|---------|---------|-------------------------------------|-------------------|------------------|------------------------------------|-----------------------|---------------------|----------------------|
|         |         | Fe <sup>III</sup> -H <sub>2</sub> O | Fe <sup>III</sup> | Fe <sup>II</sup> | Fe <sup>II</sup> -H <sub>2</sub> O | Fe <sup>III</sup> -OH | Fe <sup>IV</sup> =O | Fe <sup>II</sup> -OH |
|         |         | [+1]                                | [+1]              | [0]              | [0]                                | [0]                   | [0]                 | [-1]                 |
|         | ASP-229 | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
|         | ARG-234 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TRP-239 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | ARG-241 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | GLU-243 | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
|         | ARG-244 | 0.998                               | 0.999             | 0.998            | 0.999                              | 0.999                 | 0.998               | 0.999                |
|         | GLU-245 | 0.040                               | 0.043             | 0.040            | 0.039                              | 0.042                 | 0.041               | 0.042                |
|         | TYR-262 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | LYS-268 | 0.975                               | 0.981             | 0.979            | 0.977                              | 0.980                 | 0.975               | 0.977                |
|         | GLU-272 | 0.013                               | 0.016             | 0.017            | 0.015                              | 0.014                 | 0.014               | 0.017                |
|         | GLU-273 | 0.007                               | 0.005             | 0.004            | 0.005                              | 0.006                 | 0.004               | 0.007                |
|         | TYR-275 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TRP-278 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | GLU-280 | 0.016                               | 0.012             | 0.015            | 0.014                              | 0.015                 | 0.015               | 0.014                |
|         | ARG-283 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TYR-287 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | GLU-288 | 0.054                               | 0.055             | 0.053            | 0.052                              | 0.053                 | 0.057               | 0.052                |
| III     | HIS-3   | 0.245                               | 0.241             | 0.236            | 0.233                              | 0.232                 | 0.241               | 0.239                |
|         | LYS-5   | 0.997                               | 0.998             | 0.998            | 0.998                              | 0.998                 | 0.997               | 0.997                |
|         | HIS-7   | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
|         | ASP-8   | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
|         | TYR-9   | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | HIS-10  | 0.001                               | 0.001             | 0.001            | 0.001                              | 0.002                 | 0.002               | 0.002                |
|         | TRP-17  | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TRP-35  | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | HIS-37  | 0.460                               | 0.462             | 0.486            | 0.493                              | 0.491                 | 0.490               | 0.510                |
|         | TRP-42  | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TYR-53  | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TRP-58  | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TRP-59  | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | ASP-61  | 0.004                               | 0.004             | 0.004            | 0.003                              | 0.005                 | 0.005               | 0.004                |
|         | GLU-65  | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |

*continued on next page*

| Subunit | Residue | $\langle x \rangle$ at pH=7         |                   |                  |                                    |                       |                     |                      |
|---------|---------|-------------------------------------|-------------------|------------------|------------------------------------|-----------------------|---------------------|----------------------|
|         |         | Fe <sup>III</sup> -H <sub>2</sub> O | Fe <sup>III</sup> | Fe <sup>II</sup> | Fe <sup>II</sup> -H <sub>2</sub> O | Fe <sup>III</sup> -OH | Fe <sup>IV</sup> =O | Fe <sup>II</sup> -OH |
|         |         | [+1]                                | [+1]              | [0]              | [0]                                | [0]                   | [0]                 | [-1]                 |
| GLU-68  | 0.046   | 0.045                               | 0.042             | 0.040            | 0.041                              | 0.043                 | 0.042               |                      |
| ASP-70  | 0.837   | 0.836                               | 0.847             | 0.848            | 0.848                              | 0.845                 | 0.845               |                      |
| HIS-71  | 0.000   | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               |                      |
| ARG-76  | 1.000   | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               |                      |
| ARG-80  | 1.000   | 0.999                               | 0.999             | 0.999            | 1.000                              | 0.999                 | 1.000               |                      |
| TRP-81  | 1.000   | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               |                      |
| GLU-90  | 0.969   | 0.975                               | 0.979             | 0.977            | 0.981                              | 0.975                 | 0.984               |                      |
| TRP-97  | 1.000   | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               |                      |
| TRP-99  | 1.000   | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               |                      |
| LYS-103 | 1.000   | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               |                      |
| HIS-104 | 0.000   | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               |                      |
| TYR-107 | 1.000   | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               |                      |
| GLU-112 | 0.018   | 0.018                               | 0.019             | 0.017            | 0.017                              | 0.021                 | 0.020               |                      |
| ASP-117 | 0.020   | 0.021                               | 0.025             | 0.025            | 0.022                              | 0.024                 | 0.031               |                      |
| GLU-123 | 0.057   | 0.060                               | 0.057             | 0.059            | 0.055                              | 0.060                 | 0.057               |                      |
| ASP-129 | 0.000   | 0.001                               | 0.000             | 0.001            | 0.000                              | 0.001                 | 0.000               |                      |
| TRP-131 | 1.000   | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               |                      |
| HIS-132 | 0.013   | 0.012                               | 0.016             | 0.015            | 0.014                              | 0.013                 | 0.017               |                      |
| CYS-143 | 1.000   | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               |                      |
| CYS-146 | 1.000   | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               |                      |
| TRP-150 | 1.000   | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               |                      |
| HIS-152 | 0.000   | 0.000                               | 0.000             | 0.001            | 0.000                              | 0.000                 | 0.000               |                      |
| HIS-153 | 0.035   | 0.035                               | 0.031             | 0.034            | 0.035                              | 0.037                 | 0.034               |                      |
| HIS-157 | 0.417   | 0.433                               | 0.430             | 0.408            | 0.413                              | 0.416                 | 0.419               |                      |
| GLU-158 | 0.072   | 0.069                               | 0.068             | 0.073            | 0.073                              | 0.065                 | 0.071               |                      |
| ARG-161 | 0.998   | 0.997                               | 0.998             | 0.997            | 0.998                              | 0.998                 | 0.997               |                      |
| ARG-162 | 1.000   | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               |                      |
| ASP-163 | 0.000   | 0.001                               | 0.000             | 0.001            | 0.001                              | 0.001                 | 0.000               |                      |
| TRP-166 | 1.000   | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               |                      |
| TYR-184 | 1.000   | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               |                      |
| GLU-185 | 1.000   | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               |                      |
| TYR-186 | 1.000   | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               |                      |

*continued on next page*

| Subunit | Residue | $\langle x \rangle$ at pH=7         |                   |                  |                                    |                       |                     |                      |
|---------|---------|-------------------------------------|-------------------|------------------|------------------------------------|-----------------------|---------------------|----------------------|
|         |         | Fe <sup>III</sup> -H <sub>2</sub> O | Fe <sup>III</sup> | Fe <sup>II</sup> | Fe <sup>II</sup> -H <sub>2</sub> O | Fe <sup>III</sup> -OH | Fe <sup>IV</sup> =O | Fe <sup>II</sup> -OH |
|         |         | [+1]                                | [+1]              | [0]              | [0]                                | [0]                   | [0]                 | [-1]                 |
|         | HIS-188 | 0.982                               | 0.983             | 0.978            | 0.985                              | 0.981                 | 0.983               | 0.982                |
|         | TYR-198 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | HIS-209 | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
|         | HIS-212 | 0.029                               | 0.025             | 0.021            | 0.022                              | 0.019                 | 0.025               | 0.016                |
|         | CYS-223 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | ARG-226 | 0.321                               | 0.332             | 0.342            | 0.250                              | 0.256                 | 0.254               | 0.361                |
|         | ARG-229 | 0.997                               | 0.996             | 0.997            | 0.997                              | 0.997                 | 0.997               | 0.996                |
|         | HIS-231 | 0.001                               | 0.001             | 0.001            | 0.001                              | 0.000                 | 0.001               | 0.001                |
|         | GLU-235 | 0.002                               | 0.002             | 0.002            | 0.001                              | 0.001                 | 0.002               | 0.001                |
|         | LYS-236 | 0.290                               | 0.275             | 0.275            | 0.299                              | 0.298                 | 0.302               | 0.282                |
|         | HIS-237 | 0.521                               | 0.512             | 0.505            | 0.562                              | 0.548                 | 0.561               | 0.512                |
|         | GLU-241 | 0.140                               | 0.138             | 0.150            | 0.156                              | 0.156                 | 0.150               | 0.148                |
|         | TRP-245 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TYR-246 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TRP-247 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | HIS-248 | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
|         | ASP-251 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TRP-254 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TYR-262 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
|         | TRP-264 | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |
| IV      | HIS-11  | 0.316                               | 0.318             | 0.320            | 0.320                              | 0.317                 | 0.316               | 0.319                |
|         | ASP-17  | 0.000                               | 0.000             | 0.000            | 0.000                              | 0.000                 | 0.000               | 0.000                |
|         | GLU-22  | 0.002                               | 0.002             | 0.002            | 0.003                              | 0.001                 | 0.002               | 0.003                |
|         | LYS-23  | 0.993                               | 0.996             | 0.994            | 0.995                              | 0.995                 | 0.994               | 0.995                |
|         | ARG-30  | 0.999                               | 1.000             | 0.999            | 0.999                              | 0.999                 | 0.999               | 1.000                |
|         | TRP-34  | 1.000                               | 1.000             | 1.000            | 1.000                              | 1.000                 | 1.000               | 1.000                |

The p*K<sub>a</sub>* value of 46.4 and 46.2 are obtained for His334 in H' and H states respectively. The high p*K<sub>a</sub>* values show that His334 remains protonated in these states. The next deprotonation is expected in connection with the formation of the O state and a p*K<sub>a</sub>* of 40.7 is obtained for His334 which is associated with proton pumping event. The high p*K<sub>a</sub>* for the His334 show that it is protonated in O state. The O state was proposed [1] to evolve in two different ways. The protonation of the hydroxyl group of the Cu<sub>B</sub> center leads to O~ state and a p*K<sub>a</sub>* value of 27.5 is obtained for the deprotonation of His334 and in next possible step (O<sub>H</sub>) a p*K<sub>a</sub>* value of

40.4 is obtained for His334. The  $pK_a$  values of 45.4, 32.7 and 45.9 are obtained for His334 in E,  $E_p$  and R states respectively. All these high  $pK_a$  values show that His334 is protonated in O, E,  $E_p$  and R states.

In the present study, high  $pK_a$  values were obtained compared to the  $pK_a$  values reported by Stuchebrukhov and Fadda et al. This difference may be due to how the  $pK_a$  calculations were done. For example the model  $pK_a$  values used for the  $Cu_B$  ligands and the reliability of the charges used for the electrostatic calculations. In present study the average  $pK_a$  values of the  $Cu_B$  ligands were calculated by Tanford-Roxby approximation [77]. Stuchebrukhov and Fadda et al. calculated the  $pK_a$  values by taking only the two instances like protonated and deprotonated states of the titratable groups and not all instances were considered by them. In our calculations, we considered all instances of the titratable groups, for example four instances were considered for the tyrosines and tryptophanes. An extensive bench mark studies were performed to obtain more accurate charges and solvation energies. In the two studies [1-3, 5, 6, 182], the crystal coordinates of the  $Cu_B$  center was not fully relaxed during the geometry optimization which can lead to unrealistic bond lengths. Fadda et al. only optimized the protonated structures, while the deprotonated was not optimized i.e., same geometry was used for both protonated and deprotonated states. The homogeneous dielectric model of the protein was used in the calculations. The charges of the  $Cu_B$  center used in the calculations were found without regard to the reaction field i.e., not in a self-consistent manner. The protonation state of the protein is not properly defined in their calculations where the charges of the titratable groups are critical for electrostatic calculations. The protonation depends on the redox states of the protein. This issues were completely ignored in Fadda et al. work.

## 6.6 CONCLUSIONS

The DFT in combination with continuum electrostatic calculations were used to calculate the  $pK_a$  values of the  $Cu_B$  ligands,  $H_2O$ , His333 and His334 in various  $Cu_B$  and  $Fea_3-Cu_B$  complexes to find the role of  $Cu_B$  ligands in the reaction mechanism of cytochrome *c* oxidase. The His334 which acts as ligand for the  $Cu_B$  center was proposed [1] to be involved in the reaction mechanism of cytochrome *c* oxidase.

In present study all the possible redox combination of  $Cu_B$ -heme  $a_3$  center of cytochrome *c* oxidase are considered and all microscopic  $pK_a$  values of  $H_2O$ , His333 and His334 were calculated for different states of the  $Cu_B$ -heme  $a_3$  center. The  $pK_a$  values of His334 in  $F \rightarrow O$ ,  $E \rightarrow E_p$  transitions of the catalytic cycle were studied. Our calculations are based on the detailed study of all the  $pK_a$  values of the ligands in the  $Cu_B$  center considering all redox states of  $Cu_B$  and heme  $a_3$  center, which allows us to analyze the effect of the redox state of  $Cu_B$ -heme  $a_3$  center. According to the proposed scheme [3, 5, 6], His334 undergoes deprotonation whenever the chemical proton enters the binuclear center and convert the hydroxyl ligand to water molecule. This model was proposed based on the  $pK_a$  value of His334 where the  $pK_a$  value of His334 was shifted to lower  $pK_a$  value around 5  $pK_a$  units.

In the present study, the  $pK_a$  values of  $Cu_B$  ligands are significantly shifted to higher values in protein compared to aqueous solution. The  $pK_a$  values of  $Cu_B$  ligands are in the range of 15-60. The high  $pK_a$  values of His334 show that His334 is protonated during all steps of the catalytic cycle. The  $pK_a$  values of His334 in protein increase significantly compared to the

aqueous solution when the  $Cu_B$  center is present in the oxidized states which is inconsistent with the  $pK_a$  values reported by the Stuchebrukhov et al. The protein environment shifts the  $pK_a$  values of all complexes to higher values, independently of the redox state of the metals. According to the  $pK_a$  values of His334 the proton pumping model as suggested by Stuchebrukhov [1] might not be possible.

# CONCLUDING REMARKS AND OUTLOOK

---

This thesis is focused on the  $pK_a$  calculations of  $Cu_B$  and heme  $a_3$  center of cytochrome  $c$  oxidase using density functional theory and continuum electrostatic models to understand the role of the  $Cu_B$ -bound histidines in reaction mechanism of cytochrome  $c$  oxidase. The cytochrome  $c$  oxidase energetically couples the electron transfer reactions associated with the reduction of oxygen to water, to pump protons across the membrane. Although a vast amount of structural and functional information of cytochrome  $c$  oxidase is available from experimental and theoretical data, the actual step of coupling the redox reactions to the proton translocation is poorly understood.

Recently Stuchebrukhov et al. [1–3, 5, 6] suggested that the deprotonation of the  $Cu_B$  ligand His334 plays a central role in the proton pumping mechanism of cytochrome  $c$  oxidase. According to this suggestion, His334 deprotonates at its  $N_{\delta 1}$  when the  $Cu_B$  center gets oxidized. From our results, His334 is protonated during all steps of the catalytic cycle in cytochrome  $c$  oxidase and concluding that a simple deprotonation of His334 in the protein environment is impossible in an equilibrium situation due to the high  $pK_a$  value that this group shows when it is bound to  $Cu_B$  center. A role of this residue in the mechanism of proton pumping might not be possible.

The  $pK_a$  calculations on  $Cu_B$ -bound His333 and His334 in aqueous solution (Chapter 5) lead to the following conclusions:

The  $pK_a$  values for the first deprotonation reactions of His333 and His334 are too high ( $\sim 15$ ) to allow their deprotonation in aqueous solution. The experimental  $pK_a$  value for the deprotonation of imidazole is approximately 14 [193, 194]. The double and triple deprotonation reactions are not energetically plausible in aqueous solution due to their high  $pK_a$  values. The  $pK_a$  values in the range of 15.2–21 demonstrate that His333 and His334 are protonated at physiological pH.

The  $pK_a$  values of  $Cu_B$  bound His333 and His334 in oxidized state do not show any significant change of the  $pK_a$  value due to their coordination to Cu ion in aqueous solution. The results obtained from reduced  $Cu_B$  center reveals a small increase in the  $pK_a$  values of imidazoles. The  $pK_a$  values of His333 and His334 in the reduced state of  $Cu_B$  center are even higher to allow their deprotonation in aqueous solution. The results show that the solvation energies needed for  $pK_a$  calculations should be calculated by including the electronic polarization due to the solvent explicitly i.e., in self-consistent manner to obtain proper trend in the  $pK_a$  values of the  $Cu_B$  center.

The  $pK_a$  calculations on  $Cu_B$ -bound His333 and His334 in cytochrome *c* oxidase (Chapter 6) at pH=7 lead to the following conclusions:

The  $pK_a$  values of  $Cu_B$  ligands are significantly shifted to higher values in protein compared to aqueous solution. The  $pK_a$  values of  $Cu_B$  ligands are in the range of 15-60. The  $Cu_B$  bound His334 is protonated during all steps of the catalytic cycle proving that the Fe and Cu ion oxidation states do not lower the  $pK_a$  values of  $Cu_B$  ligands in the protein. The  $pK_a$  values of  $Cu_B$  ligands are significantly shifted to higher values dependent on the redox states of the  $Cu_B$  and heme  $a_3$  center. The protein environment increases the  $pK_a$  values of all complexes to higher values independent of the redox state of the metals. The His333 and His334 are likely to be protonated at physiological pH. The double deprotonation and triple deprotonation reactions in protein are energetically unfavorable. The  $pK_a$  values of His333 show that this residue is likely to be protonated in the protein and an involvement of this residue as proton loading site in the reaction mechanism of cytochrome *c* oxidase can therefore be ruled out. These results are inconsistent with the proposed role of His334 as a key element in the pumping mechanism. The proton pumping model as suggested by Stuchebrukhov [1] with the involvement of His334 might not be possible.

A few issues could not be completely solved in the framework of this study and should be the topic of future research:

In the present work, Tyr280 which is cross-linked with His284 is replaced by methyl group in the  $Cu_B$  model. This cross-linked His-Tyr residue can be represent by cross-linked imidazole-phenol model and the  $pK_a$  values and redox potentials of the Tyr280 can be calculated to understand the role of cross-linked Try280 in the reaction mechanism of cytochrome *c* oxidases.

To understand the redox coupled protonation reactions in cytochrome *c* oxidases, the mid-point potential of the redox centers  $Cu_A$ , heme *a*, heme  $a_3$  and  $Cu_B$  can be calculated using accurate theory which can be used to study redox behavior of the cytochrome *c* oxidase.

In the present work, minimal models are used for the  $Cu_B$  and heme  $a_3$  centers. The interactions between the  $Cu_B$  and heme  $a_3$  centers in the protein are treated as electrostatic interactions. Large models can be used for the binuclear center including the  $Cu_B$  center with the two histidine ligands and the cross-linked His-Tyr residues and heme  $a_3$  center with substituted porphyrin. High level quantum chemical methods can be performed on this complete binuclear center, to study the influence of the heme  $a_3$  redox states on  $Cu_B$  ligands.

The crystal structures of both mammalian and bacterial cytochrome *c* oxidases do not contain all water molecules, that could be involved in the proton transfer pathways. Computer algorithms can be used to determine the likely positions and model internal water molecules in the protein cavities to understand the hydrogen bond networks in proton transfer pathways. We plan to use `Vol` program (developed in our group) to place water molecules and to find the hydrogen bond networks in cytochrome *c* oxidase.

The complex coupling of electron and proton transfer reactions in cytochrome *c* oxidases must be elucidated in detail. In the future, we plan to use the DMC (Dynamic Monte Carlo) program to study the charge transfer and proton transfer rates in cytochrome *c* oxidase.

---

## BIBLIOGRAPHY

---

- [1] D M Popović and A A Stuchebrukhov. Electrostatic study of the proton pumping mechanism in bovine heart cytochrome *c* oxidase. *J Am Chem Soc*, 126:1858–1871, 2004.
- [2] D M Popović and A A Stuchebrukhov. Proton pumping mechanism and catalytic cycle of cytochrome *c* oxidase: Coulomb pump model with kinetic gating. *FEBS Lett*, 556:126–130, 2004.
- [3] J Quenneville, D M Popović, and A A Stuchebrukhov. Combined DFT and electrostatics study of the proton pumping mechanism in cytochrome *c* oxidase. *Biochim Biophys Acta*, 1757:1035–1046, 2006.
- [4] A A Stuchebrukhov and D M Popović. Comment on “Acidity of a Cu-bound histidine in the binuclear center of cytochrome *c* oxidase”. *J Phys Chem*, 110:17286–17287, 2006.
- [5] J Quenneville, D M Popović, and A A Stuchebrukhov. Redox-dependent  $pK_a$  of Cu<sub>B</sub> histidine ligand in cytochrome *c* oxidase. *J Phys Chem B*, 108:18383–18389, 2004.
- [6] D M Popović, J Quenneville, and A A Stuchebrukhov. DFT/Electrostatic calculations of  $pK_a$  values in cytochrome *c* oxidase. *J Phys Chem B*, 109:3616–3626, 2005.
- [7] O M H Richter and B Ludwig. Cytochrome *c* oxidase - structure, function, and physiology of a redox-driven molecular machine. *Rev Physiol Biochem Pharmacol*, 147:47–74, 2003.
- [8] M Wikström. Identification of the electron transfers in cytochrome oxidase that are coupled to proton-pumping. *Nature*, 338:776–778, 1989.
- [9] P Mitchell. Coupling of phosphorylation to electron and hydrogen transfer by a chemiosmotic type of mechanism. *Nature*, 356:301–309, 1961.
- [10] D Voet and J G Voet. *Biochemistry*. John Wiley and Sons, Ltd, Chichester, 3rd edition, 2004.
- [11] B Ludwig and G Suhatz. A two-subunit cytochrome *c* oxidase (cytochrome *aa*<sub>3</sub>) from *Paracoccus denitrificans*. *Proc Natl Acad Sci USA*, 77:196–200, 1980.
- [12] S Iwata, C Ostermeier, B Ludwig, and H Michel. Structure at 2.8 Å resolution of cytochrome *c* oxidase from *Paracoccus denitrificans*. *Nature*, 376:660–669, 1995.
- [13] C Ostermeier, S Iwata, B Ludwig, and H Michel. F<sub>v</sub> fragment-mediated crystallization of the membrane protein bacterial cytochrome *c* oxidase. *Nat Struct Biol*, 2:842–846, 1995.

- [14] C Ostermeier, A Harrenga, U Ermler, and H Michel. Structure at 2.7 Å resolution of the *Paracoccus denitrificans* two-subunit cytochrome *c* oxidase complexed with an F<sub>v</sub> fragment. *Proc Natl Acad Sci USA*, 95:10547–10553, 1997.
- [15] M Svensson-Ek, J Abramson, G Larsson, S Tornroth, P Brzezinski, and S Iwata. The X-ray crystal structures of wild-type and EQ(I-286) mutant cytochrome *c* oxidase from *Rhodobacter sphaeroides*. *J Mol Biol*, 321:329–339, 2002.
- [16] S Yoshikawa, K Shinzawa-Itoh, R Nakashima, R Yaona, E Yamashita, N Inoue, M Yao, M J Fei, C Peters-Libeu, T Mizushima, H Yamaguchi, T Tomizaki, and T Tsukihara. Redox-coupled crystal structural changes in bovine heart cytochrome *c* oxidase. *Science*, 280:1723–1729, 1998.
- [17] T Tsukihara, H Aoyama, E Yamashita, T Tomizaki, H Yamaguchi, K Shinzawa-Itoh, R Nakashima, R Yaona, and S Yoshikawa. Structure of metal sites of oxidized bovine heart cytochrome *c* oxidase at 2.8 Å. *Science*, 269:1069–1074, 1995.
- [18] T Tsukihara, H Aoyama, E Yamashita, T Tomizaki, H Yamaguchi, K Shinzawa-Itoh, R Nakashima, R Yaona, and S Yoshikawa. The whole structure of the 13-subunit oxidized cytochrome *c* oxidase at 2.8 Å. *Science*, 272:1136–1144, 1996.
- [19] K C Minghetti, V C Goswitz, N E Gabriel, J J Hill, C A Barassi, C D Georgiou, S I Chan, and R B Gennis. Modified, large-scale purification of the cytochrome *o* complex (*bo*-type oxidase) of *Escherichia coli* yields a two heme/one copper terminal oxidase with high specific activity. *Biochemistry*, 31:6917–6924, 1992.
- [20] A Puustinen, J E Morgan, M Verkhovskii, J W Thomas, R B Gennis, and M Wikström. The low-spin heme site of cytochrome *o* from *Escherichia coli* is promiscuous with respect to heme type. *Biochemistry*, 31:10363–10369, 1992.
- [21] M Lauraeus, M Wikström, C Varotsis, M M J Tecklenburg, and G T Babcock. Optical and resonance Raman spectroscopy of the heme groups of the quinol-oxidizing cytochrome *aa*<sub>3</sub> of *Bacillus subtilis*. *Biochemistry*, 31:10054–10060, 1992.
- [22] M Santana, F Kunst, M F Hullo, G Rapoport, A Danchin, and P Glaser. Molecular cloning, sequencing, and physiological characterization of the *qox* operon from *Bacillus subtilis* encoding the *aa*<sub>3</sub>-600 quinol oxidase. *J Biol Chem*, 267:10225–10231, 1992.
- [23] Y C Fann, I Ahmed, N J Blackburn, J S Boswell, M L Verkhovskaya, B M Hoffman, and M Wikström. Structure of CuB in the binuclear heme-copper center of the cytochrome *aa*<sub>3</sub>-type quinol oxidase from *Bacillus subtilis*: An ENDOR and EXAFS study. *Biochemistry*, 34:10245–10255, 1995.
- [24] T Soulimane, G Buse, G P Bourenkov, H D Bartunik, R Huber, and M E Than. Structure and mechanism of the aberrant *ba*<sub>3</sub> cytochrome *c* oxidase from *Thermus thermophilus*. *EMBO J*, 19:1766–1776, 2000.
- [25] L M Hunsicker-Wang, R L Pacoma, Y Chen, J A Fee, and C D Stout. A novel cryoprotection scheme for enhancing the diffraction of crystals of recombinant cytochrome *ba*<sub>3</sub> oxidase from *Thermus thermophilus*. *Acta Crystallogr Sect.D*, 61:340–343, 2005.

- [26] J Abramson, S Riistama, G Larsson, A Jasaitis, M Svensson-EK, L Laakkonen, A Puustinen, S Iwata, and M Wikström. The structure of the ubiquinone oxidase from *Escherichia coli* and its ubiquinone binding site. *Nat Struct Biol*, 7, 2000. 910-917.
- [27] B Epel, C S Slutter, F Neese, P M H Kroneck, W G Zumft, I Pecht, O Farver, Y Lu, and D Goldfarb. Electron-mediating  $Cu_A$  centers in proteins: A comparative high field H ENDOR study. *J Am Chem Soc*, 124:8152-8162, 2002.
- [28] M R Bratton, M A Pressler, and J P Hosler. Suicide inactivation of cytochrome *c* oxidase: catalytic turnover in the absence of subunit III alters the active site. *Biochemistry*, 38: 16236-16245, 1999.
- [29] S Riistama, A Puustinen, A Garcia-Horsman, S Iwata, H Michel, and M Wikström. Channelling of dioxygen into respiratory enzyme. *Biochim Biophys Acta*, 1275:1-4, 1996.
- [30] D Zaslavsky and R B Gennis. Proton pumping by cytochrome oxidase progress, problems and postulates. *Biochim Biophys Acta*, 1458:164-179, 2000.
- [31] F MacMillan, A Kannig, J Behr, T Prisner, and H Michel. Direct evidence for a tyrosine radical in the reaction of cytochrome *c* oxidase with hydrogen peroxide. *Biochemistry*, 38:9179-9184, 1999.
- [32] J Wang, J Rumbley, Y C Ching, S Takahashi, and R B Gennis. Reaction of cytochrome  $bo_3$  with oxygen: extra redox center(s) are present in the protein. *Biochemistry*, 34:15504-15511, 1995.
- [33] S E J Rigby, S Jünemann, P R Rich, and P Heathcote. Reaction of bovine cytochrome *c* oxidase with hydrogen peroxide produces a tryptophan cation and a porphyrin cation radical. *Biochemistry*, 39:5921-5928, 2000.
- [34] H Käss, F MacMillan, B Ludwig, and T F Prisner. Investigation of the Mn site in cytochrome *c* oxidase from *Paracoccus denitrificans* by high-frequency EPR. *J Phys Chem B*, 104:5362-5371, 2000.
- [35] A Seelig, B Ludwig, J Seeling, and G Schatz. Copper and manganese electron spin resonance studies of cytochrome *c* oxidase from *Paracoccus denitrificans*. *Biochim Biophys Acta*, 636:162-167, 1981.
- [36] H Witt, A Wittershagen, E Bill, B O Kolbesen, and B Ludwig. Asp-193 and Glu-218 of subunit II are involved in the  $Mn^{2+}$ -binding of *Paracoccus denitrificans* cytochrome *c* oxidase. *FEBS Lett*, 409:128-130, 1997.
- [37] U Pfitzner, A Kirichenko, A A Konstantinov, M Mertens, A Wittershagen, B O Kolbesen, G C M Steffens, A Harrenga, H Michel, and B Ludwig. Mutations in the  $Ca^{2+}$  binding site of the *Paracoccus denitrificans* cytochrome *c* oxidase. *FEBS Lett*, 456:365-369, 1999.
- [38] S Riistama, L Laakkonen, M Wikström, M Verkhovskiy, and A Puustinen. The calcium binding site in cytochrome  $aa_3$  from *Paracoccus denitrificans*. *Biochemistry*, 38:10670-10677, 1999.

- [39] A Lee, A Kirichenko, T Vygodina, S A Siletsky, T K Das, D L Rousseau, R B Gennis, and A A Konstantinov.  $\text{Ca}^{2+}$ -binding site in *Rhodobacter sphaeroides* cytochrome *c* oxidase. *Biochemistry*, 41:8886–8898, 2002.
- [40] A D Mills and S Ferguson-Miller. Understanding the mechanism of proton movement linked to oxygen reduction in cytochrome *c* oxidase: lessons from other protons. *FEBS Lett*, 545:47–51, 2003.
- [41] S Yoshikawa, K Shinzawa-Itoh, and T Tsukihara. X-ray structure and the reaction mechanism of bovine heart cytochrome *c* oxidase. *J Inorg Biochem*, 82:1–7, 2000.
- [42] R M Nyquist, D Heitbrink, C Bolwien, R B Gennis, and J Heberle. Direct observation of protonation reactions during the catalytic cycle of cytochrome *c* oxidase. *Proc Natl Acad Sci USA*, 100:8715–8720, 2003.
- [43] F Malatesta, F Nicoletti, V Zickermann, B Ludwig, and M Brunori. Electron entry in a  $\text{Cu}_A$  mutant of cytochrome *c* oxidase from *Paracoccus denitrificans*. conclusive evidence on the initial electron entry metal center. *FEBS Lett*, 434:322–324, 1998.
- [44] H Witt, F Malatesta, F Nicoletti, M Brunori, and B Ludwig. Tryptophan 121 of subunit II is the electron entry site to cytochrome *c* oxidase in *Paracoccus denitrificans*-involvement of a hydrophobic patch in the docking region. *J Biol Chem*, 273:5132–5136, 1998.
- [45] K Wang, Y Zhen, R Sadoski, S Grinnell, L Geren, S Ferguson-Miller, B Durham, and F Millett. Definition of the interaction domain for cytochrome *c* on cytochrome *c* oxidase: II. Rapid kinetic analysis of electron transfer from cytochrome *c* to *Rhodobacter sphaeroides* cytochrome oxidase surface mutants. *J Biol Chem*, 274:38042–38050, 1999.
- [46] Y Zhen, C W Hoganson, G T Babcock, and S Ferguson-Miller. Definition of the interaction domain for cytochrome *c* on cytochrome *c* oxidase. I. Biochemical, spectral and kinetic characterization of surface mutants in subunit II of *Rhodobacter sphaeroides* cytochrome  $aa_3$ . *J Biol Chem*, 274:38032–38041, 1999.
- [47] A Kannt, U Pfitzner, M Ruitenberg, P Hellwig, B Ludwig, W Mantele, K Fendler, and H Michel. Mutation of Arg-54 strongly influences heme composition and rate and directionality of electron transfer in *Paracoccus denitrificans* cytochrome *c* oxidase. *J Biol Chem*, 274:37974–37981, 1999.
- [48] J F Nagle, M Mille, and J Morowitz. Theory of hydrogen bonded chains in bioenergetics. *J Chem Phys*, 72:3959–3971, 1980.
- [49] R B Gennis. Cytochrome *c* oxidase: one enzyme, two mechanisms. *Science*, 280:1712–1729, 1998.
- [50] P Hellwig, J Behr, C Ostermeier, O M H Richter, U Pfitzner, A Odenwald, B Ludwig, H Michel, and W Mantele. Involvement of the glutamic acid 278 in the redox reaction of the cytochrome *c* oxidase from *Paracoccus denitrificans* investigated by FTIR spectroscopy. *Biochemistry*, 37:7390–7399, 1998.

- [51] S Jünemann, B Meunier, N Fisher, and P R Rich. Effects of mutation of the conserved glutamic acid-286 in subunit I of cytochrome *c* oxidase from *Rhodobacter sphaeroides*. *Biochemistry*, 38:5248–5255, 1999.
- [52] M Lübben, A Prutsch, B Mamat, and K Gerwert. Electron transfer induces side-chain conformational changes of glutamate-286 from cytochrome *bo*<sub>3</sub>. *Biochemistry*, 38:2048–2056, 1999.
- [53] H Michel. The mechanism of proton pumping by cytochrome *c* oxidase. *Proc Natl Acad Sci USA*, 95:12819–12824, 1998.
- [54] R Pomès, G Hummer, and M Wikström. Structure and dynamics of a proton shuttle in cytochrome *c* oxidase. *Biochim Biophys Acta*, 1365:255–260, 1998.
- [55] L Weng and G M Baker. Reaction of hydrogen peroxide with the rapid form of resting cytochrome oxidase. *Biochemistry*, 30:5727–5733, 1991.
- [56] D A Proshlyakov, T Ogura, K Shinzawa-Itoh, S Yoshikawa, and T Kitagawa. Resonance raman/absorption characterization of the oxo intermediates of cytochrome *c* oxidase generated in its reaction with hydrogen peroxide: pH and H<sub>2</sub>O<sub>2</sub> concentration dependence. *Biochemistry*, 35:8580–8586, 1996.
- [57] H Michel. Cytochrome *c* oxidase: Catalytic cycle and mechanisms of proton pumping-A discussion. *Biochemistry*, 38:15129–15140, 1999.
- [58] J E Morgan, M I Verkhovsky, G Palmer, and M Wikström. Role of the P<sub>R</sub> intermediate in the reaction of cytochrome *c* oxidase with O<sub>2</sub>. *Biochemistry*, 40:6882–6892, 2001.
- [59] M I Verkhovsky, A Jasaitis, M L Verkhovskaya, J E Morgan, and M Wikström. Proton translocation by cytochrome *c* oxidase. *nature*, 400:480–483, 1999.
- [60] M I Verkhovsky, J E Morgan, M L Verkhovskaya, and M Wikström. Translocation of electrical charge during a single turnover of cytochrome *c* oxidase. *Biochim Biophys Acta*, 1318:6–10, 1997.
- [61] M Wikström, J E Morgan, and M I Verkhovsky. Proton and electrical charge translocation by cytochrome *c* oxidase. *Biochim Biophys Acta*, 1318:299–306, 1997.
- [62] G M Ullmann and E W Knapp. Electrostatic models for computing protonation and redox equilibria in proteins. *Eur Biophys J*, 28:533–551, 1999.
- [63] M D Liptak and G C Shields. Accurate p*K*<sub>a</sub> calculations for carboxylic acids using complete basis set and Gaussian-n models combined with CPCM continuum solvation methods. *J Am Chem Soc*, 123:7314–7319, 2001.
- [64] B Honig and A Nicholls. Classical electrostatics in biology and chemistry. *Science*, 268:1144–1149, 1995.
- [65] D Bashford and M Karplus. p*K*<sub>a</sub>'s of ionizable groups in proteins: Atomic detail from a continuum electrostatic model. *Biochemistry*, 29:10219–10225, 1990.
- [66] M K Gilson and B Honig. Calculation of the total electrostatic energy of a macromolecular system: Solvation energies, binding energies, and conformational analysis. *Proteins*, 4:7–18, 1988.

- [67] M K Gilson. Introduction to continuum electrostatics with molecular applications, 2002.
- [68] A R Leach. *Molecular Modelling*. Pearson Education Limited, 2th edition, 2001.
- [69] D Bashford and K Gewert. Electrostatic calculations of the  $pK_a$  values of ionizable groups in bacteriorhodopsin. *J Mol Biol*, 224:473–486, 1992.
- [70] M Tanokura.  $^1\text{H}$ -NMR study on the tautomerism of the imidazole ring of histidine residues II. Microenvironments of histidine-12 and histidine-119 of bovine pancreatic ribonuclease A. *Biochim Biophys Acta*, 742:586–596, 1983.
- [71] Y Nozaki and C Tanford. Intrinsic dissociation constants of aspartyl and glutamyl carboxyl groups. *J Biol Chem*, 242:4731–4735, 1967.
- [72] J A Schellman. Macromolecular binding. *Biopolymer*, 14:999–1018, 1975.
- [73] G M Ullmann. Relations between protonation constants and titration curves in polyprotic acids: A critical view. *J Phys Chem B*, 107:1263–1271, 2003.
- [74] A R Klingen, E Bombarda, and G M Ullmann. Theoretical investigation of the behavior of titratable groups in proteins. *Photochem Photobiol Sci*, 5:588–596, 2006.
- [75] A R Klingen and G M Ullmann. Negatively charged residues and hydrogen bonds tune the ligand histidine  $pK_a$  values of rieske iron-sulfur protein. *Biochemistry*, 39:12383–12389, 2004.
- [76] P Beroza, D R Fredkin, M Y Okamura, and G Feher. Protonation of interacting residues in a protein by a Monte Carlo method: Application to lysozyme and the photosynthetic reaction center of *Rhodobacter sphaeroides*. *Proc Natl Acad Sci USA*, 88:5804–5808, 1991.
- [77] C Tanford and R Roxby. Interpretation of protein titration curves. Application to lysozyme. *Biochemistry*, 11:2192–2198, 1972.
- [78] N Homeyer, T Essigke, G M Ullmann, and H Sticht. Effects of histidine protonation and phosphorylation on histidine-containing phosphocarrier protein structure, dynamics, and physicochemical properties. *Biochemistry*, 46:12314–12326, 2007.
- [79] C J Kramer. *Essentials of computational chemistry*. Wiley, New York, 2nd edition, 2002.
- [80] F Jensen. *Introduction to computational chemistry*. Wiley, New York, 2nd edition, 1999.
- [81] W Kohn and L J Sham. Self-consistent equations including exchange and correlation. *Phys Rev*, 15:1133–1138, 1965.
- [82] R G Parr and W Yang. *Density-Functional Theory of atoms and Molecules*. Oxford Science, 1st edition, 1989.
- [83] M Springborg. *Methods of Electronic structure calculations*. Wiley, 1st edition, 2000.
- [84] S H Vosko, L Wilk, and M Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: A critical analysis. *Can J Phys*, 58:1200–1211, 1980.

- [85] J P Perdew and W Yue. Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation. *Phys Rev B*, 33:8800–8802, 1986.
- [86] J P Perdew, J A Chevary, S H Vosko, K A Jackson, M R Pederson, D J Singh, and C Fiolhais. Atoms, molecules, solids and surfaces: applications of the generalized gradient approximation for exchange and correlation. *Phys Rev, B* 15:6671–6687, 1992.
- [87] A D Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys Rev A*, 38:3098–3100, 1988.
- [88] J P Perdew, K Burke, and Y Wang. Generalized gradient approximation for the exchange-correlation hole of a many-electron system. *Phys Rev B*, 54:16533–16539, 1996.
- [89] C Adamo and V Barone. Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: The mPW and mPW1PW models. *J Chem Phys*, 108:664–673, 1998.
- [90] C Lee, W Yang, and R G Parr. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys Rev B*, 37:785–789, 1988.
- [91] A D Becke. Density-functional thermochemistry. III. The role of exact exchange. *J Chem Phys*, 98:5648–5652, 1993.
- [92] J P Perdew, J A Chevary, S H Vosko, K A Jackson, M R Pederson, D J Singh, and C Fiolhais. Erratum: Donor transition energy in gas superlattices in a magnetic field along the growth axis. *Phys Rev B*, 48:4978–4978, 1993.
- [93] P J Stephens, F J Devlin, C F Chabalowski, and M J Frisch. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J Phys Chem*, 98:11623–11627, 1994.
- [94] J S Binkley, J A Pople, and W J Hehre. Self-consistent molecular orbital methods. 21. Small split-valence basis sets for first-row elements. *J Am Chem Soc*, 102:939–947, 1980.
- [95] M S Gordon, J S Binkley, J A Pople, W J Pietro, and W J Hehre. Self-consistent molecular-orbital methods. 22. Small split-valence basis sets for second-row elements. *J Am Chem Soc*, 104:2797–2803, 1982.
- [96] J D Dill and J A Pople. Self-consistent molecular orbital methods. XV. Extended Gaussian-type basis sets for lithium, beryllium, and boron. *J Chem Phys*, 62:2921–2923, 1975.
- [97] R Krishnan, J S Binkley, R Seeger, and J A Pople. Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *J Chem Phys*, 72:650–654, 1980.
- [98] M Born. Volumen und hydrationswärme den ionen. *Z Phys*, 1:45–48, 1920.
- [99] J G Kirkwood. The dielectric polarization of polar liquids. *J Chem Phys*, 7:911–919, 1939.

- [100] L Onsager. Electric moments of molecules in liquids. *J Am Chem Soc*, 58:1486–1493, 1936.
- [101] C Lim, D Bashford, and M Karplus. Absolute  $pK_a$  calculations with continuum dielectric methods. *J Phys Chem*, 94:5610–5620, 1991.
- [102] D Sitkoff, N Ben-Tal, and B Honig. Calculation of alkane to water solvation free energies using continuum solvent models. *J Phys Chem*, 100:2744–2752, 1996.
- [103] B Marten, K Kim, C Cortis, R A Friesner, R B Murphy, M N Ringnalda, D Sitkoff, and B Honig. New model for calculation of solvation free energies: Correction of self-consistent reaction field continuum dielectric theory for short-range hydrogen-bonding effects. *J Phys Chem*, 100:11775–11788, 1996.
- [104] J T David, B Marten, R A Friesner R Murphy, D Sitkoff, A Nicholls, B Honig, M Ringnalda, and W A Goddard III. Accurate first principles calculation of molecular charge distributions and solvation energies from ab initio quantum mechanics and continuum dielectric theory. *J Am Chem soc*, 116:11875–11882, 1994.
- [105] S Miertus, E Scrocco, and J Tomasi. Electrostatic interaction of a solute with a continuum. A direct utilizaion of ab initio molecular potentials for the prevision of solvent effects. *Chem Phys*, 55:117–129, 1981.
- [106] S Miertus and J Tomasi. Approximate evaluations of the electrostatic free energy and internal energy changes in solution processes. *Chem Phys*, 65:239–245, 65.
- [107] M Cossi, V Barone, R Cammi, and J Tomasi. Ab initio study of solvated molecules: A new implementation of the polarizable continuum model. *Chem Phys Lett*, 255:327–335, 1996.
- [108] E. Cancés, B Mennucci, and J Tomasi. A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics. *J Chem Phys*, 107:3032–3041, 1997.
- [109] V Barone, M Cossi, and J Tomasi. A new definition of cavities for the computation of solvation free energies by the polarizable continuum model. *J Chem Phys*, 107:3210–3221, 1997.
- [110] M Cossi, V Barone, B Mennucci, and J Tomasi. Ab initio study of ionic solutions by a polarizable continuum dielectric model. *Chem Phys Lett*, 286:253–260, 1998.
- [111] V Barone, M Cossi, and J Tomasi. Geometry optimization of molecular structures in solution by the polarizable continuum model. *J Comp Chem*, 19:404–417, 1998.
- [112] V Barone and M Cossi. Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. *J Phys Chem A*, 102:1995–2001, 1998.
- [113] B Mennucci and J Tomasi. Continuum solvation models: A new approach to the problem of solute's charge distribution and cavity boundaries. *J Chem Phys*, 106:5151–5158, 1997.

- [114] Z Jia, D Du, Z Zhou, A Zhang, and R Hou. Accurate  $pK_a$  determinations for some organic acids using an extended cluster method. *Chem Phys Chem*, 439, 2007. 374–380.
- [115] J R Pliego Jr and J M Riveros. Theoretical calculation of  $pK_a$  using the cluster-continuum model. *J Phys Chem A*, 106:7434–7439, 2002.
- [116] M Namazian and S Halvani. Calculations of  $pK_a$  of carboxylic acids in aqueous solution using density functional theory. *J Chem Thermodynamics*, 38:1495–1502, 2006.
- [117] X J Qi, L Liu, Y Fu, and Q X Guo. Ab initio calculations of  $pK_a$  values of transition-metal hydrides in acetonitrile. *Organometallics*, 25:5879–5886, 2006.
- [118] J B Foresman, T A Keith, K B Wiberg, J Snoonian, and M J Frisch. Solvent effects. 5. Influence of cavity shape, truncation of electrostatics, and electron correlation on ab initio reaction field calculations. *J Phys Chem*, 100:16098–16104, 1996.
- [119] R A Pierotti. A scaled particle theory of aqueous and nonaqueous solutions. *Chem Rev*, 76:717–726, 1976.
- [120] J L Pascual-ahuir, E Silla, and I Tunon. GEPOL: An improved description of molecular surfaces. III. A new algorithm for the computation of a solvent-excluding surface. *J Comp Chem*, 15:1127–1138, 1994.
- [121] J Tomasi and M Persico. Molecular interactions in solution: An overview of methods based on continuous distributions of the solvent. *Chem Rev*, 94:2027–2094, 1994.
- [122] R Cammi and J Tomasi. Remarks on the use of the apparent surface charges (ASC) methods in solvation problems: Iterative versus matrix-inversion procedures and the renormalization of the apparent charges. *J Comp Chem*, 16:1449–1458, 1995.
- [123] J Cioslowski. *Encyclopedia of computational chemistry*. Wiley, New York, p. 892 edition, 1998.
- [124] B Honig and A Nicholls. Classical electrostatics in biology and chemistry. *Science*, 268:1144–1149, 1995.
- [125] Nakamura H. Roles of electrostatic interaction in proteins. *Q Rev Biophys*, 29:1–90, 1996.
- [126] A Warshel and A Papazyan. Electrostatic effects in macromolecules fundamental concepts and practical modelling. *Curr Opin Struct Biol*, 8:211–217, 1998.
- [127] M E Davis and J A McCammon. Electrostatics in biomolecular structure and dynamics. *Chem Rev*, 90:509–521, 1990.
- [128] A S Yang, M R Gunner, R Sampogna, K Sharp, and B Honig. On the calculation of  $pK_a$ s in proteins. *Proteins*, 15:252–265, 1993.
- [129] M R Gunner and B Honig. Electrostatic control of midpoint potentials in the cytochrome subunit of the *Rhodospseudomonas viridis* reaction center. *Proc Natl Acad Sci USA*, 88:9151–9155, 1993.

- [130] G M Ullmann and E W Knapp. Electrostatic computations of protonation and redox equilibria in proteins. *Eur Biophys J*, 28:533–551, 1999.
- [131] F Fogolari, A Brigo, and H Molinari. The Poisson-Boltzmann equation for biomolecular electrostatics: A tool for structural biology. *J Mol Recognit*, 15:377–392, 2000.
- [132] P McCaldon and P Argos. Oligopeptide biases in protein sequences and their use in predicting protein coding regions in nucleotide sequences. *Proteins Struct Funct Gene*, 4:99–122, 1988.
- [133] U Ryde and E Sigfridsson. Comparison of methods for deriving atomic charges from the electrostatic potential and moments. *J Comp Chem*, 19:377–395, 1998.
- [134] F Martin and H Zipse. Charge distribution in the water molecule - A comparison of methods. *J Comp Chem*, 26:97–105, 2004.
- [135] K B Wiberg and P R Rablen. Comparison of atomic charges derived via different procedures. *J Comp Chem*, 14:1504–1518, 1994.
- [136] G S Maciel and E Garcia. Charges derived from electrostatic potentials: Exploring dependence on theory and geometry optimization levels for dipole moments. *Chem Phys Lett*, 409:29–33, 2005.
- [137] E Sigfridsson, U Ryde, and B L Bush. Restrained point-charge models for disaccharides. *J Comp Chem*, 23:351–364, 2002.
- [138] C A Reynolds, J W Essex, and W G Richards. Atomic charges for variable molecular conformations. *J Am Chem Soc*, pages 9075–9079, 1992.
- [139] R S Mulliken. Electronic population analysis on LCAO[Single Bond]MO molecular wave functions. I. *J Chem Phys*, 23:1833–1840, 1955.
- [140] R S Mulliken. Criteria for the construction of good self-consistent-field molecular orbital wave functions, and the significance of LCAO-MO population analysis. *J Chem Phys*, 36:3428–3439, 1962.
- [141] P O Löwdin. Quantum theory of many-particle systems. I. Physical interpretations by means of density matrices, natural spin-orbitals, and convergence problems in the method of configurational interaction. *Phys Rev*, 97:1474–1489, 1955.
- [142] F Weinhold. *Encyclopedia of computational chemistry*. Wiley, New York, p. 1792 edition, 1998.
- [143] A E Reed, R B Weinstock, and F Weinhold. Natural population analysis. *J Chem Phys*, 83:735–746, 1985.
- [144] A E Reed, L A Curtiss, and F Weinhold. Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. *Chem Rev*, 88:899–926, 1988.
- [145] L E Chirlian and M M Francl. Atomic charges derived from electrostatic potentials: A detailed study. *J Comp Chem*, 8:894–905, 1987.

- [146] C M Breneman and K B Wiberg. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J Comp Chem*, 11:361–373, 1990.
- [147] U C Singh and P A Kollman. An approach to computing electrostatic charges for molecules. *J Comp Chem*, 5:129–145, 1984.
- [148] B H Besler, K M Merz Jr., and P A Kollman. Atomic charges derived from semiempirical methods. *J Comp Chem*, 11:431–439, 1990.
- [149] M L Connolly. Analytical molecular surface calculations. *J Appl Cryst*, 16:548–558, 1983.
- [150] C I Bayly, P Cieplak, W Cornell, and P A Kollman. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J Phys Chem*, 97:10269–10280, 1993.
- [151] B T Thole and P Th Duijnen. A general population analysis preserving the dipole moment. *Theor Chim Acta*, 63:209–221, 1983.
- [152] M Swart, P T V Duijnen, and J G Snijders. A charge analysis derived from an atomic multipole expansion. *J Comp Chem*, 22:79–88, 2001.
- [153] A D MacKerell, D Bashford, M Bellott, R L Dunbrack, J D Evanseck, M J Field, S Fischer, J Goa, H Guo, S Ha, D Joseph-McCarthy, L Kuchnir, K Kuczera, F T K Lau, C Mattos, S Michnick, T Ngo, D T Nguyen, B Prodhom, W E Reiher, B Roux, M Schlenkrich, J C Smith, R Stote, J Straub, M Watanabe, J Wiórkiewicz-Kuczera, D Yin, and M Karplus. All-atom empirical potential for molecular modeling and dynamic studies of proteins. *J Phys Chem B*, 102:3586–3616, 1998.
- [154] M J Frisch, G W Trucks, H B Schlegel, G E Scuseria, M A Robb, J R Cheeseman, J A Montgomery, Jr., T Vreven, K N Kudin, J C Burant, J M Millam, S S Iyengar, J Tomasi, V Barone, B Mennucci, M Cossi, G Scalmani, N Rega, G A Petersson, H Nakatsuji, M Hada, M Ehara, K Toyota, R Fukuda, J Hasegawa, M Ishida, T Nakajima, Y Honda, O Kitao, H Nakai, M Klene, X Li, J E Knox, H P Hratchian, 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, P Y Ayala, K Morokuma, G A Voth, P Salvador, J J Dannenberg, V G Zakrzewski, S Dapprich, A D Daniels, M C Strain, O Farkas, D K Malick, A D Rabuck, K Raghavachari, J B Foresman, J V Ortiz, Q Cui, A G Baboul, S Clifford, J Cioslowski, B B Stefanov, G Liu, A Liashenko, P Piskorz, I Komaromi, R L Martin, D J Fox, T Keith, M A Al-Laham, C Y Peng, A Nanayakkara, M Challacombe, P M W Gill, B Johnson, W Chen, M W Wong, C Gonzalez, and J A Pople. Gaussian 03, Revision C.02. Gaussian, Inc., Wallingford, CT, 2004.
- [155] RESP program. <http://www.amber.ucsf.edu/amber>.
- [156] W D Cornell, P Cieplak, C I Bayly, and P A Kollmann. Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. *J Am Chem Soc*, 115:9620–9631, 1993.
- [157] S R Cox and D E Williams. Representation of the molecular electrostatic potential by a net atomic charge model. *J Comp Chem*, 2:304–323, 1981.

- [158] B C Berks, S J Ferguson, J W B Moir, and D J Richardson. Enzymes and associated electron transport systems that catalyse the respiratory reduction of nitrogen oxidase and oxyanions. *Biochim Biophys Acta*, 1232:97–173, 1995.
- [159] D J Richardson. Bacterial respiration: A flexible process for a changing environment. *Microbiol*, 146:551–571, 2000.
- [160] S J Ferguson. Nitrogen cycle enzymology. *Curr Opin Chem Biol*, 2:182–193, 1998.
- [161] I Moura and J J G Moura. Structural aspects of denitrifying enzymes. *Curr Opin Chem Biol*, 5:168–175, 2001.
- [162] O Einsle, A Messerschmidt, P Stach, G P Bourenkov, H D Bartunik, R Huber, and P M H Kroneck. Structure of cytochrome *c* nitrite reductase. *Nature*, 400:476–480, 1999.
- [163] O Einsle, P Stach, A Messerschmidt, J Simon, A Kröger, R Huber, and P M H Kroneck. Cytochrome *c* nitrite reductase from *Wolinella succinogenes*. *J Biol Chem*, 275:39608–39616, 2000.
- [164] V A Bamford, H C Angove, H E Seward, A J Thomson, J A Cole, J N Butt, A M Hemmings, and D J Richardson. Structure and spectroscopy of the periplasmic cytochrome *c* nitrite reductase from *Escherichia coli*. *Biochemistry*, 41:2921–2931, 2002.
- [165] C Costa, J J Moura, I Moura, M Y Liu, H D Peck Jr, J LeGall, Y N Wang, and B H Huynh. Hexaheme nitrite reductase from *Desulfovibrio desulfuricans*. Mossbauer and EPR characterization of the heme groups. *J Biol Chem*, 265:14382–14388, 1990.
- [166] C Costa, J J G Moura, I Moura, Y Wang, and B H Huynh. Redox properties of cytochrome *c* nitrite reductase from *Desulfovibrio desulfuricans* ATCC 27774. *J Biol Chem*, 271:23191–23196, 1996.
- [167] J Simon. Enzymology and bioenergetics of respiratory nitrite ammonification. *FEMS Microbiol Rev*, 26, 2002. 285-309.
- [168] J Simon, R Gross, O Einsle, P M H Kroneck, A Kröger, and O Klimmek. A NapC/NirT-type cytochrome *c* (NrfH) is the mediator between the quinone pool and the cytochrome *c* nitrite reductase of *Wolinella succinogenes*. *Mol Microbiol*, 35:686–696, 2000.
- [169] J D Gwyer, H C Angove, D J Richardson, and J N Butt. Redox-triggered events in cytochrome *c* nitrite reductase. *Bioelectrochem*, 63:43–47, 2004.
- [170] O Einsle, A Messerschmidt, R Huber, P M H, and F Neese. Mechanism of the six-electron reduction of nitrite to ammonia by cytochrome *c* nitrite reductase. *J Am Chem Soc*, 124: 11737–11745, 2002.
- [171] Insight II. <http://www.accelrys.com/products/insight/index.html>.
- [172] P Hohenberg and W Kohn. Inhomogeneous electron gas. *Phys Rev*, 43:864–871, 1964.
- [173] G te Velde, F M Bickelhaupt, E J Baerends, C F Guerra, S J A VanGisbergen, J G Snijders, and T Ziegler. Chemistry with ADF. *J Comp Chem*, 22:931–967, 2001.
- [174] K Hinsien and B Roux. A potential function for computer simulation studies of proton transfer in acetylacetone. *J Comp Chem.*, 18:368–380, 1997.

- [175] A Bondi. Van der waals volumes and radii. *J Phys Chem*, 68:441–451, 1964.
- [176] G S Wilson. Electrochemical studies of porphyrin redox reactions as cytochrome moldes. *Bioelectrochem Bioenerg*, 1:172–179, 1974.
- [177] J D Wyer, D J Richardson, and J N Butt. Resolving complexity in the interactions of redox enzymes and their inhibitors: Contrasting mechanisms for the inhibition of a cytochrome *c* nitrite reductase revealed by protein film voltammetry. *Biochemistry*, 43: 15086–15094, 2002.
- [178] G M Ullmann. The coupling of protonation and reduction in proteins with multiple redox centers. Theory, computational method, and application to cytochrome *c*<sub>3</sub>. *J Phys Chem B*, 104:6293–6301, 2000.
- [179] H Michel, J Behr, A Harrenga, and A Kannt. Cytochrome *c* oxidase: Structure and spectroscopy. *Annu Rev Biophys Biomol Struct*, 27:329–356, 1998.
- [180] M Wikström. Cytochrome *c* oxidase: 25 years of the elusive proton pump. *Biochim Biophys Acta*, 1655:241–247, 2004.
- [181] S F-Miller and G T Babcock. Heme/copper terminal oxidase. *Chem Rev*, 96:2889–2907, 1996.
- [182] E Fadda, N Chakrabarthi, and R Pomès. Acidity of a Cu-bound histidine in the binuclear center of cytochrome *c* oxidase. *J Phys Chem B*, 109:22629–22640, 2005.
- [183] E Fadda, N Chakrabarti, and R Pomès. Reply to “Comment on acidity of a Cu-bound histidine in the binuclear center of cytochrome *c* oxidase”. *J Phys Chem*, 110:17288–17289, 2006.
- [184] P E M Siegbahn, M R A Blomberg, and M L Blomberg. Theoretical study of the energetics of proton pumping and oxygen reduction in cytochrome oxidase. *J Phys Chem B*, 107: 10946–10955, 2003.
- [185] M R A Blomberg and P E M Siegbahn. Quantum chemistry applied to the mechanisms of transition metal containing enzymes - Cytochrome *c* oxidase, a particularly challenging case. *J comp Chem*, 27:1373–1384, 2006.
- [186] P Rydberg, E Sigfridsson, and U Ryde. On role of the axial ligand in heme proteins: A theoretical study. *J Biol Inorg Chem*, 9:203–223, 2004.
- [187] G M Ullmann, L Noodleman, and D A Case. Density functional calculation of  $pK_a$  values and redox potentials in the bovine rieske iron-sulfur protein. *J Biol Inorg Chem*, 7:632–639, 2002.
- [188] J Tomasi, B Mennucci, and R Cammi. Quantum mechanical continuum solvation models. *Chem Rev*, 105:2999–3093, 2005.
- [189] B Ragu and G M Ulmann. Conformational dependence of charge fit. 2003. Unpublished.
- [190] F Floris and J Tomasi. Evaluation of the dispersion contribution to the solvation energy. A simple computational model in the continuum approximation. *J Comp Chem*, 10:616–627, 1989.

- [191] F M Floris, J Tomasi, and J L Pascual Ahuir. Dispersion and repulsion contributions to the solvation energy: Refinements to a simple computational model in the continuum approximation. *J Comp Chem*, 12:784–791, 1991.
- [192] M Tanokura. <sup>1</sup>H-NMR study on the tautomerism of the imidazole ring of histidine residues. I. Microscopic p*K*<sub>a</sub> values and molar ratios of tautomers in histidine-containing peptides. *Biochim Biophys Acta*, 742:576–585, 1983.
- [193] H Walba and R W Isensee. Spectrophotometric study of the hydrolysis constants of the negative ions of some aryl imidazoles. *J Am Chem Soc*, 77:5488–5492, 1955.
- [194] G Yagil. The proton dissociation constant of pyrrole, indole and related compounds. *Tetrahedron*, 23:2855–2861, 1967.
- [195] J R Pliengo and J M Riveros. New values for the absolute solvation free energy of univalent ions in aqueous solution. *Chem Phys Lett*, 332:597–602, 2000.
- [196] H Reiss and A Heller. The absolute potential of the standard hydrogen electrode: A new estimate. *J Phys Chem*, 89:4207–4213, 1985.
- [197] G J Tawa, I A Topol, S K Burt, R A Caldwell, and A A Rashin. Calculation of the aqueous solvation free energy of the proton. *J Chem Phys*, 109:4852–4863, 1998.
- [198] G M Ullmann and E W Knapp. Electrostatic models for computing protonation and redox equilibria in proteins. *Eur Biophys J*, 28:533–551, 1999.
- [199] Tanford C. *Physical Chemistry of Macromolecules*. John Wiley, New York, 1st edition, 1961.
- [200] M R Gunner, A Nicholls, and B Honig. Electrostatic potentials in *Rhodospseudomonas viridis* reaction centers: Implications for the driving force and directionality of electron transfer. *J Phys Chem*, 100:4277–4291, 1996.
- [201] C R Lancaster, H Michel, B Honig, and M R Gunner. Calculated coupling of electron and proton transfer in the photosynthetic reaction center of *Rhodospseudomonas viridis*. *Biophys J*, 70:2469–2492, 1996.
- [202] B Rabenstein, G M Ullmann, and E W Knapp. Energetics of electron-transfer and protonation reactions of the quinones in the photosynthetic reaction center of *Rhodospseudomonas viridis*. *Biochemistry*, 37:2488–2495, 1998.
- [203] V Z Spassov, H Luecke, K Gerwert, and D Bashford. p*K*<sub>a</sub> calculations suggest storage of an excess proton in a hydrogen-bonded water network in bacteriorhodopsin. *J Mol Biol*, 312:203–219, 2001.
- [204] R V Sampogna and B Honig. Environmental effects on the protonation states of active site residues in bacteriorhodopsin. *Biophys J*, 66:1341–1352, 1994.
- [205] Orientations of protein in membranes (OPM) database. <http://opm.phar.umich.edu>.
- [206] M A Lomize, A L Lomize, I D Pogozheva, and H I Mosberg. OPM: Orientations of proteins in membranes database. *Bioinformatics*, 22:623–625, 2006.

- 
- [207] A L Lomize, I D Pogozeva, M A Lomize, and H I Mosberg. Positioning of proteins in membranes: A computational approach. *Protein Science*, 15:1318–1333, 2006.
- [208] M S Till and G M Ulmann. Protein cavities. In preparation.
- [209] J M Mouesca, J L Chen, L Noodleman, D Bashford, and D A Case. Density functional/Poisson-Boltzmann calculations of redox potentials for iron-sulfur clusters. *J Am Chem Soc*, 116:11898–11914, 1994.



Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet habe.

Ferner erkläre ich, dass ich anderweitig mit oder ohne Erfolg nicht versucht habe, diese Dissertation einzureichen. Ich habe keine gleichartige Doktorprüfung an einer anderen Hochschule endgültig nicht bestanden.

Punnagai Munusami

Bayreuth, den 9. Juni 2008