1-1-2014

Computational Studies Of Enzymatic Reactions And Inorganic Compounds

Dong Fang
Wayne State University,

COMPUTATIONAL STUDIES OF ENZYMATIC REACTIONS AND INORGANIC COMPOUNDS

by

DONG FANG

DISSERATAION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2014

MAJOR: CHEMISTRY (Physical)

Approved by:

_____

Advisor Date

_____

_____

_____

_____

DEDICATION

To my parents

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor Professor G. Andrés Cisneros for his guidance, encouragement and support through these five years. I feel I have grown a lot with the group under the care of Professor Cisneros since I joined the group when it started in 2009.

I am also very grateful to the other members in my committee, Professor H. Bernhard Schlegel, Professor Stanislav Groysman and Professor Domenico L. Gatti and former members, Professor Caroline Morgan and Professor Claudio N. Verani. I would also like to thank our other collaborators, Professor Jean-Philip Piquemal, Professor Richard L Lord, Dr Shubin Liu and Dr. Robin Chaudret.

The current and former members of Cisneros group are greatly appreciated for their help and friendship. Schlegel group and Chernyak group are also thanked for their joint group meetings.

I would like to say many thanks to the Chemistry department for giving me the opportunity to study here, Wayne State Computing grid for the computing resources and Wayne State Graduate School for a Rumble fellowship. Thanks are given to some Chemistry department staff, Melissa Barton, Nestor Ocampo, Erin Bachert and Bernadette Miesik for their help.

Last but not the least, I would like to thank my parents, my friends and all other people who have helped me.

TABLE OF CONTENTS

Dedication ii

Acknowledgements iii

List of Tables vii

List of Figures viii

List of Schemes xi

CHAPTER 1 INTRODUCTION AND OVERVIEW 1

1.1 References 4

CHAPTER 2 TOWARDS A DEEPER UNDERSTANDING OF ENZYME REACTIONS USING THE COUPLED ELF/NCI ANALYSIS: APPLICATION TO DNA REPAIR ENZYMES 5

2.1 Introduction 5

2.2 Results and discussion 8

2.2.1 Pol lambda 8

2.2.2 Epsilon subunit 10

2.3 Conclusions 16

2.4 References 17

CHAPTER 3 NOVEL ALKOXIDE CLUSTER TOPOLOGIES FEATURING RARE SEESAW GEOMETRY AT TRANSITION METAL CENTERS 20

3.1 Introduction 20

3.2 Methods 21

3.3 Results and discussion 21

3.4 Conclusions 25

3.5 References 26

CHAPTER 4 *AB INITIO* QM/MM CALCULATIONS SHOW AN INTERSYSTEM CROSSING IN THE HYDROGEN ABSTRACTION STEP IN DEALKYLATION CATALYZED BY ALKB 27

4.1. Introduction 27

4.2. Computational Methods and Theory 30

4.3. Results and Discussion 33

4.3.1. Spin states of the ferryl moiety 33

4.3.2. DFT functional benchmarks, cluster and QM/MM calculations 34

4.3.3 Crossing between the two quintet sub-states 38

4.3.4. Energy decomposition analysis (EDA) 40

4.3.5. Electron Localization Function (ELF) and Non-covalent Interaction (NCI) analyses on critical structures and the MECP 43

4.4. Conclusions 47

4.5 References 48

CHAPTER 5 ALTERNATIVE PATHWAY FOR THE REACTION CATALYZED BY DNA DEALKYLASE ALKB FROM *AB INITIO* QM/MM SIMULATIONS 53

5.1. Introduction 53

5.2. Computational Details and Theory 57

5.3. Results and Discussion 59

5.3.1 Hydrogen abstraction 59

5.3.2 The OH rebound 66

5.3.3 Unbinding from the iron center, proton transfer and the formation of formaldehyde 71

5.3.4 Comparison between the H2O and OH- pathways 76

5.4. Conclusions 77

5.5 References 78

CHAPTER 6 DFT STERIC BASED ENERGY DECOMPOSITION ANALYSIS OF INTERMOLECULAR INTERACTIONS 81

6.1. Introduction 81

6.2. Theory and Methodology 82

6.2.1 DFT Steric analysis 82

6.2.2 Constrained Space Orbital Variation 84

6.2.3 Non-Covalent Interaction (NCI) Analysis 85

6.2.4 Computational details 85

6.3. Results and Discussion 86

6.3.1. Hydrogen bonding dimers: Water-containing dimers and DNA base pair 86

6.3.2. Benzene dimers 97

6.4. Conclusions 101

Abstract 106

Autobiographical Statement 108

LIST OF TABLES

Table 4.1 Relative QM/MM energies in the quintet state calculated by five different DFT functionals (unit: kcal/mol; for each functional, the lowest absolute energy is defined as zero; dash line means data not obtained).....35

Table 4.2 Residues that have more than 1 kcal/mol difference by EDA (all the residues are for Coulomb interaction except that the residue in italic is for Van der Waals interaction) 41

Table 4.3 The change of the population in V(C,H) and V(O,H) basins along the MEP.....46

Table 5.1 Mülliken spin population for Fe and O (oxo) in $^{HS}Fe^{III}$-O_{AF} and $^{IS}Fe^{III}$-O_F states with different distances between Fe and O (d(Fe-O_{oxo})) (Å) for the OH^- pathway. The numbers in italics correspond to the structure where the $^{HS}Fe^{III}$-O_{AF} and $^{IS}Fe^{III}$-O_F states are close in energy (T_{OH}). Except for this point, all other structures are obtained by only moving O with all other atoms fixed starting from the optimized $^{HS}Fe^{III}$-O_{AF} reactant. 63

Table 6.1 CSOV decomposition on water-containing dimers (unit: kcal/mol; B3LYP/6-311++G(2d,2p)).....88

Table 6.2 DFT Steric Analysis on water-containing dimers (unit: kcal/mol; B3LYP/6-311++G(2d,2p)).....91

Table 6.3 CSOV decomposition on DNA base pair. (unit:kcal/mol; B3LYP/6-31G(d)) 96

Table 6.4 DFT Steric Analysis on DNA base pair. (unit:kcal/mol; B3LYP/6-31G(d)) 96

Table 6.5 CSOV decomposition on sandwich benzene dimer. (unit:kcal/mol; B3LYP/6-31G(d)).....98

Table 6.6 DFT Steric Analysis on sandwich benzene dimer. (unit:kcal/mol; B3LYP/6-31G(d)).....98

Table 6.7 CSOV decomposition on T-shaped benzene dimers with B3LYP/6-31G(d) (unit: kcal/mol; the upper benzene is denoted as A.) 100

Table 6.8 DFT Steric Analysis on T-shaped benzene dimers with B3LYP/6-31G(d) (kcal/mol).....101

LIST OF FIGURES

Figure 2.1 NCI analysis for the reactant, TS (state 1), and transphosphorilation structure (state 2) of Polλ with Mg^{2+} and Mn^{2+} (the isovalue is 0.4 au, and the color scale is −0.04 au <sign(λ2)ρ < 0.04 au). 9

Figure 2.2 Combined ELF and NCI surfaces of the critical structures for the reaction catalyzed by the ε subunit with Mg^{2+} and Mn^{2+}. Only atoms subjected to the ELF analysis are shown; the rest of the QM subsystem is omitted for clarity. Panels a, b, and c show the reactant, TS, and product for Mg^{2+}, respectively. Panels d, e, and f correspond to the reactant, TS, and product for Mn2+, respectively. The isovalue for ELF is 0.87, and for NCI it is 0.5 au with the color scale −0.1 au < sign(λ2)ρ < 0.1 au. 12

Figure 2.3 Numbering of the atoms in ε that are used in the ELF calculations. 12

Figure 2.4 Distributed moments, M0 (population), M1 (first moment), and M2 (second moment) of V(O6,P1) and V(O12,P1) and the distances (d(O6,P1) and d(O12,P1), in Å, along the reaction (a, b are for the reactions catalyzed by Mg^{2+} and Mn^{2+}, respectively). 14

Figure 3.1 The structure of the seesaw compounds synthesized by Groysman group 20

Figure 3.2 Superposition of the crystal structures (black) and the calculated structures (red). Only the core part is shown. The letters in panel a are the atomic names; M stands for metals. For (a) Cr; (b) Mn; (c) Fe; (d) Co. 21

Figure 3.3 The ELF (a) and NCI (b) surfaces for the Cr, complex. For ELF, the isovlaue is 0.735. 22

Figure 3.4 NCI surfaces for complex **6**, the isovalue is 0.5. a) The optimized structure with tert-butyl was replaced by methyl (only the core is shown), $-0.07<\text{sign}(\lambda_2)\rho<0.07$ a.u. b) NCI analysis for the complete (unoptimized) crystal structure, $-0.02<\text{sign}(\lambda_2)\rho<0.02$a.u. Regarding the chosen arbitrary color code, red surface indicates strong repulsion; blue surface indicates strong attraction; green surface indicates relatively weak interactions. Different $\text{sign}(\lambda_2)\rho$ ranges may be used to distinguish the strength of the interactions. Note that one can compare the relative strength of the interactions based on the colors only when using the same $\text{sign}(\lambda_2)\rho$ range. 24

Figure 3.5 Optimized structures with the ligands being replaced with H. 24

Figure 3.6 The NCI for complex **3-6** with tert-butyl being replaced with methyl**.** The isovalue is 0.5 and $-0.02<\text{sign}(\lambda_2)\rho<0.02$ au. 25

Figure 4.1 Electron configuration diagram of the iron(IV)-oxo moiety in the quintet and triplet states in an ideal octahedral ligand field. (HS, IS and LS represent high-spin, intermediate-spin and low-spin on Fe, respectively; AF and F stand for antiferromagnetic and ferromagnetic.) 34

Figure 4.2 (a) Relative energies of the critical structures and and spin densities for some key atoms in the quintet and triplet states by ωB97XD with two waters included in QM subsystem (unit: kcal/mol; for the absolute energy of the $^{IS}Fe^{III}$–O_F reactant is defined as zero). (b) Calculated minimum potential and free energy paths (MEP and MFEP respectively). Potential and Helmholtz free energy differences (unit: kcal/mol) are also shown. All the energies are relative to the reactant. ΔE represents the total potential energy difference. ΔE_{QM} is the energy difference for QM subsystem. $\Delta A_{QM/MM}$ is the free energy change in the QM/MM interaction. $\Delta A^{fluct}_{QM/MM}$ is the contribution from the QM subsystem fluctuation. The total free energy difference ΔA is the sum of ΔE_{QM}, $\Delta A_{QM/MM}$ and $\Delta A^{fluct}_{QM/MM}$. 37

Figure 4.3 (a) The minimum $^{HS}Fe^{III}$–O_{AF} reactant. (b) The minimum $^{IS}Fe^{III}$–O_F reactant. (c) The MECP between the $^{HS}Fe^{III}$–O_{AF} state and the $^{IS}Fe^{III}$–O_F sate. (d) The relative single point energies of the $^{HS}Fe^{III}$–O_{AF} state and $^{IS}Fe^{III}$–O_F at different Fe–O bond lengths (unit: kcal/mol; the absolute energy of the $^{HS}Fe^{III}$–O_{AF} reactant is defined as zero).....40

Figure 4.4 (a) EDA results for all the residues. (b) Positions of the residues obtained from EDA (in grey) relative to the active site (ball and stick, black font). (c) Comparison of AlkB (red) and ABH2 (black) showing catalytically important residues that are structurally conserved. (d) Comparison of AlkB (red) and ABH3 (blue) showing catalytically important residues that are structurally conserved.....43

Figure 4.5 ELF and NCI surfaces of the critical structures and the MECP: (a) the $^{IS}Fe^{III}$–O_F reactant; (b) the $^{HS}Fe^{III}$–O_{AF} TS; (c) the $^{HS}Fe^{III}$–O_{AF} intermediate; (d) the $^{HS}Fe^{III}$–O_{AF} MECP; (e) the $^{IS}Fe^{III}$–O_F MECP. (the isovalue for ELF is 0.89; for NCI, the isovalue is 0.5 au, and -0.05 au < sign(λ_2)ρ < 0.05 au; for the NCI surfaces, the color scaleis chosen so that blue indicates relatively strong attraction, green indicates relatively weak interaction, and red indicates relatively strong repulsion in the region of non-covalent interactions.) 45

Figure 5.1 The active site of AlkB with 1-meA in the crystal structure (PDBID: 2FDG). 55

Figure 5.2 Relative energies (in kcal/mol, with $^{IS}Fe^{III}$-O_F as the reference state) of reactant, TS and I1 and Mülliken spin populations of key atoms (Fe, the first O denotes the oxo, the second O denotes the O of OH^- bound to the iron, and C denotes the carbon of methyl group of 1-meA) for the hydrogen abstraction step for OH^- pathway in quintet ($^{IS}Fe^{III}$-O_F and $^{HS}Fe^{III}$-O_{AF}) and triplet states ($^{LS}Fe^{III}$-O_F and $^{IS}Fe^{III}$-O_{AF}).....61

Figure 5.3 α-LUMO and β-LUMO (canonical orbitals, isovalue=0.05 au) of the quintet reactants, MECP and TS structures along the H_2O pathway. 62

Figure 5.4 α-LUMO and β-LUMO (canonical orbitals, isovalue for the surface 0.05 au) of the quintet reactants, T_{OH} (Table 1) and TS structures along the OH^- pathway.....67

Figure 5.5 Relative energies (kcal/mol) of the structures along the minimum energy path (MEP) for the OH rebound step in the quintet state for H_2O pathway. The numbers in the parentheses are reaction barriers, which are the energy differences between intermediates and their corresponding TSs. (a) and OH^- pathway (b). The energy of the corresponding $^{IS}Fe^{III}$-O_F reactant (Figure 5.2) is

taken as zero for each pathway. 69

Figure 5.6 NCI surface (isovalue 0.5 au and a color scale -0.1<sign(λ_2)ρ<0.1 au) of the zwitterion structure for the OH^- pathway. 69

Figure 5.7 (a) the chemical structures of 1-meA and 1-deazameA. (b),(c),(d) are the 1-deazameA-related intermediates. 72

Figure 5.8 Relative energies (in kcal/mol) for the structures along the minimum energy path (MEP) for the detachment of the DNA base from Fe and the formation of formaldehyde in the quintet state for the H_2O pathway. The numbers in the parentheses are reaction barriers, which are the energy differences between intermediates and their corresponding TSs. (a) and OH^- pathway (b). The $^{IS}Fe^{III}$-O_F reactant (Figure 5.2) is taken as the reference for each pathway. 73

Figure 5.9 Relative energies (kcal/mol) for the structures along the minimum energy path (MEP) for the proton transferred to Glu136 in the quintet state for H_2O pathway. (Arg210, Glu136 and a bridging water were added to the QM subsystem for these structures) 75

Figure 6.1 NCI plots and structures of 24 dimers (The molecule on the left is A and the right one is B). The isovalue for NCI surfaces is 0.5 au. The surfaces are colored on a blue-green-red scale with -0.04 au <sing(λ2)ρ< 0.04 au. (dimer 1,2,3: formamide-water; dimer 4,5,6: N-methylformamide-water; dimer 7,8: methanol-water; dimer 9: fluomethane-water; dimer 10: dichloromethane-water; dimer 11: difluoromethane-water; dimer 12: methane-water; dimer 13: ammonia-water; dimer 14: methylamine-water; dimer w1-w10: water-water.) 87

Figure 6.2 Linear fitting of E_e versus E_{Coul} (a), ΔE_e (b) and ΔE_s(c). 91

Figure 6.3 Linear fitting of E_{tot} versus ρ_{BCP}(a), and ΔE_s versus 1/ ρ_{BCP} (b) for water dimers. 93

Figure 6.4 Plots of ∇ρ versus ρ for dimer w1: (a) monomer; (b) dimer.(Cufoff for ∇ρ and ρ are 2 and 0.2 respectively to show the intermolecular points) 93

Figure 6.5 Linear fitting of ΔE_e versus E_{Coul} (a) and $E_{Coul}+E_{CT}+E_{Pol}$ (b) for water dimers. 95

Figure 6.6 Linear fitting of $E_{exch-rep}$ versus ΔE_X (a), $E_{exch-rep}$ versus ΔE_S (b), $E_{Pol}+E_{CT}$ versus ΔE_S (c), and $E_{Pol}+E_{CT}$ versus ΔE_{Pauli} (d) for water dimers. 96

Figure 6.7 Structures of benzene dimers (r represents the distance. For a, it is the distance between the centers of two monomers. For b, it is the distance between one hydrogen atom of the upper benzene and the center of the other monomer. Black circles are critical points for AIM. For BCPs, only one is shown for illustration. BCP stands for bond critical point, and CCP is cage critical point.) 97

Figure 6.8 (a) Linear fitting of ρ_{CCP} versus $\rho_{BCP(C-C)}$. (b) Plot of E_{tot} versus ρ_{CCP}. (c) Linear fitting of ΔE_s versus ρ_{CCP} for the benzene dimer in the sandwich conformation. 98

Figure 6.9 (a) Linear fitting of ρ_{CCP} versus $\rho_{BCP(C-H)}$. (b) Plot of E_{tot} versus ρ_{CCP}. (c) Linear fitting of ΔE_s versus ρ_{CCP} for the T-shaped benzene dimer. 99

LIST OF SCHEMES

Scheme 2.1 Reaction Mechanism for the Reaction Catalyzed by Polλ[a] . [a]The atom labels used in the text are denoted in the reactant structure. 10

Scheme 2.2 Reaction Mechanism for the ε Subunita aIn the first step, a proton is transferred from the ordered water to H162 to activate the nucleophile. Subsequently, the oxygen performs a nucleophilic attack to break the phosphoesther bond coupled with a second proton transfer to E14.(ref 16) 11

Scheme 4.3 Proposed mechanism for the direct dealkylation of the 1-meA catalyzed by AlkB. 28

Scheme 5.1 The proposed mechanism for the steps starting from H abstraction in the dealkylation catalyzed by AlkB based on the TauD mechanism. 54

Scheme 5.2 The detailed H_2O pathway. 56

Scheme 5.3 The OH^- pathway: The newly proposed mechanism with OH^- coordinated to the iron (R_{OH}) 57

CHAPTER 1

INTRODUCTION AND OVERVIEW

Computer simulations have become a powerful tool to study chemical and biological problems. *Ab initio* Quantum Mechanics (QM) methods such as Density Functional Theory (DFT) have been well developed to produce highly accurate results. However, because of the limit on the current computing resources, large-size systems such as enzymes may not be treated purely by these methods. As a result, in order to take advantage of Molecular Mechanics (MM) in computing time, the QM/MM method has been introduced for studying reactions in large-size systems, such as enzymatic reactions.[1,2] This combined method has been widely used and shows good performance. To better understand the reactivity and mechanism for both small-size and large-size systems, it is useful to have knowledge of the intermolecular and intramolecular interactions. The combined electron localization function (ELF) and Non-Covalent Interaction (NCI)[3,4] analyses are powerful tools to meet this demand. The following chapters in this thesis will cover the development and applications of the combined ELF/NCI analysis and QM/MM method.

Chapters 2 and 3, describe the use of the combined ELF/NCI analysis to study the interactions in enzymatic reactions and inorganic metal clusters. ELF can provide information about strong interactions such as covalent interactions, and NCI can provide details on the relatively weak interactions such as hydrogen bonding and van der Waals interactions. The combination of these two methods will cover the whole spectrum of the interactions. The combined technique was originally developed to study small-size organic reactions.[5] This method was later extended by us to study enzymatic reactions.

In Chapter 2, the extension of the combined analysis to study two enzymatic reactions will be presented, the synthesis of DNA by Human DNA polymerase λ and the ε subunit of DNA Polymerase III. For both enzymes, there are two divalent metal cations in the active site for

catalysis. In addition, Both DNA polymerases and exonucleases can employ Mg^{2+} and Mn^{2+} for catalysis. For both reactions, the reaction barrier is higher when Mg^{2+} is in the active site, compared to Mn^{2+}. The barrier difference is explained by the comparison between the interactions for Mg^{2+} and Mn^{2+} revealed by the combined ELF/NCI analysis. In addition, NCI analysis supports the unusual penta-coordination for the catalytic metal for subunit ε. Finally, the combined analysis enables one to predict the location of transition state (TS) from the earlier structures as shown by the evolution of ELF/NCI surfaces along the minimum energy path (MEP).

In Chapter 3, we discuss several important interactions revealed by the combined analysis in detail. These interactions together are responsible for the unusual seesaw shape of several metal clusters synthesized by the Groysman group.

In Chapter 4 and 5, we turn our focus to the *ab initio* QM/MM study of the mechanism for the oxidative repair of alkylated DNA base by AlkB, specifically the steps after the formation of Fe^{IV}-oxo for the repair of 1-methyladinine (1-meA) catalyzed by AlkB.

The hydrogen abstraction step, which has been studied and found to be the rate limiting step[6] is fully discussed in Chapter 4. We found that Fe^{IV}-oxo tends to be Fe^{III}-oxyl where one electron is transferred from O to Fe. This state gives rise to two sub-states for the total quintet and triplet states. The four states have been explored by QM/MM simulations. It is found that the MEP lies on the quintet surface, involving a crossing between the two quintet sub-states. The minimum energy crossing point (MECP) is located. Energy decomposition analysis on the residue level reveals nine residues that may have an impact on the reaction barrier. The conversation of these residues in two Human homologs, ABH2 and ABH3 is obtained by sequence and structure alignments. The combined ELF/NCI analysis shed lights on important interactions and their evolution along the MEP.

Chapter 5 continues the work on the other steps for the traditional H_2O pathway and the study on a new mechanism with OH^- bound to the iron metal center pathway (OH^- pathway) for the repair of 1-meA by AlkB. These steps include an OH rebound from the Fe atom to the methyl moiety and the dissociation of formaldehyde to repair the adenine base. It is shown OH^- facilitates the reorientation of the oxo and has a small impact on the barriers for the hydrogen abstraction and OH rebound steps. The effects of the enzyme and the OH^- ligand on the hydrogen abstraction by the Fe^{IV}-oxo moiety will be discussed in detail. The new OH rebound step is coupled with a proton transfer to the OH^- ligand, and results in a novel zwitterion intermediate. This zwitterion structure can also be characterized as Fe-O-C complex and facilitates the formation of formaldehyde. By contrast, for the pathway with H_2O bound to iron, the hydroxyl product of the OH rebound step first needs to unbind from the metal center before transferring a proton to Glu136 or other residues and substrates. The consistency between our theoretical results and experimental findings are discussed.

A better force field is one way to improve the performance of QM/MM method. Energy decomposition analysis (EDA) methods are useful for the development of force fields. In Chapter 6, a newly proposed EDA method based on DFT-steric analysis[7] is applied on water-containing dimers, DNA base pairs and benzene dimers. Results obtained from our DFT-steric EDA analysis are compared to results from the constrained space orbital variations (CSOV) and Bader's quantum theory of atoms in molecules (QTAIM) topological analysis. The driving forces for the formation of dimers are given by these two EDA methods from different perspectives. We also found DFT steric energy correlates with the electron density at critical points from QTAIM analysis in different patterns for different molecular systems, which qualitatively explains the linear relationships between the steric and total interaction energy. In addition, the DFT steric energy represents effects originating from the spatial arrangement of the electron density when dimers form.

1.1 References

(1) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700.

(2) Warshel, A.; Levitt, M. *J. Mol. Bio.* **1976**, *103*, 227.

(3) Contreras-García, J.; Johnson, E. R.; Keinan, S.; Chaudret, R.; Piquemal, J.-P.; Beratan, D. N.; Yang, W. *J. Chem. Theory Comput.* **2011**, *7*, 625.

(4) Johnson, E. R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W. *J. Am. Chem. Soc.* **2010**, *132*, 6498.

(5) Gillet, N.; Chaudret, R.; Contreras-García, J.; Yang, W.; Silvi, B.; Piquemal, J.-P. *J. Chem. Theory Comput.* **2012**, *8*, 3993.

(6) Liu, H.; Llano, J.; Gauld, J. W. *J. Phys. Chem. B* **2009**, *113*, 4887.

(7) Liu, S. *J. Chem. Phys.* **2007**, *126*.

CHAPTER 2

TOWARDS A DEEPER UNDERSTANDING OF ENZYME REACTIONS USING THE COUPLED ELF/NCI ANALYSIS: APPLICATION TO DNA REPAIR ENZYMES

Reproduced with permission from *J. Chem. Theory Comput.,* **2013,** 9, 2156-2160

2.1 Introduction

DNA polymerases and exonucleases play an important role in DNA replication and repair.[1,2] DNA polymerases catalyze the addition of incoming deoxyribonucleotides (dNTP) to a nascent DNA chain. According to the primary structure of their catalytic subunits, DNA polymerases are categorized into several different families.[3] Human DNA polymerase λ is a member of the X family and has been implicated in the gap filling and end joining processes involved in the repair of DNA double-strand breaks and in base excision repair of damaged bases.[4,5] Exonucleases catalyze the inverse reaction of polymerases, i.e., the excision of DNA bases. As a catalytic subunit of *E. coli* DNA polymerase III, subunit ε accounts for the proofreading exonuclease activity by catalyzing the removal of newly incorporated mispaired nucleotides.[6-8]

Recently, some of us introduced a coupling of two quantum interpretative techniques using both the Electron Localization Function (ELF)[9-11] and the Non-Covalent Interaction (NCI)[12,13] index to study reaction mechanisms through a partition of the electronic density. This combined ELF/NCI analysis was initially applied to investigate in detail prototypical organic reactions[14] since this cross interpretative approach enables the following of the full range of interactions within a complex system. Indeed, ELF/NCI is able to simultaneously unravel regions of strong electron pairing (atoms and bonds) using ELF and regions of low density associated to low density reduced gradient revealing weak interactions using NCI. This approach has been automated and gives chemically intuitive and visual results to describe chemical reactions involving multiple steps.[14] In this contribution, we propose the first application of this technique to large systems of biological

interest using hybrid QM/MM (Quantum Mechanics/Molecular Mechanics) methods. To this end, we have applied ELF/NCI analysis to study the reactions catalyzed by DNA polymerase λ (Pol λ) and the ε subunit of Pol III (ε). The reaction paths associated with the catalytic mechanisms of these two enzymes have been studied in detail by some of us previously using QM/MM simulations.[15,16] In both cases, two divalent metal cations are required for catalysis. Both DNA polymerases and exonucleases can employ different metals for catalysis.[17-19] In our previous studies we have investigated the role of Mg^{2+} and Mn^{2+} for the reaction mechanism catalyzed by both enzymes.[15,16] In addition, we have recently used ELF to investigate metal ion mutagenicity in the synthesis of DNA by Polλ.[17]

Methods

The ELF function was originally proposed to measure the electron localization in atomic and molecular systems based on the Hartree-Fock formalism,[9] and subsequently extended to DFT.[20] Similarly to the Bader's topological analysis of the electron density[21], the ELF values can also be treated as a continuous and differentiable scalar field in 3D space. The maxima of ELF (critical points) are named attractors and are located on atoms, bonds and lone pairs following chemical intuition. Consequently, the molecular space can be divided into regions, named basins that respectively contain all points whose ELF gradient field converges toward the same attractor. The different basins do not overlap and the surface separating two basins is called a separatrix. Depending on the electrons they refer to, the different basins can be denoted as C(), for core electrons or V() for valence electrons. Core basins are defined as belonging to a single nucleus. Valence basins may belong to a single atom and are termed monosynaptic (e.g. lone pairs), or may belong to two (or three) atoms and are termed disynaptic (e.g. bonds).

Recently, a third category has been proposed, subvalence electrons have been proposed to be the outer-shell "core" electrons of specific metal cations.[18] Basins associated with chemical bonds between atoms X and Y are denoted as V(X,Y). Furthermore, the ability to divide the molecular

space into regions allows not only the computation of integrated properties, such as charge and volume, but also multipoles or Fukui functions over these regions.[18,22]

The NCI[12,13] index enables the study of the domains of the electronic density associated with weak interactions that exhibit both low electron density and low reduced density gradient(s):

$$s(\rho)=\frac{|\nabla\rho|}{2(3\pi^2)^{1/3}\rho^{4/3}} \quad (1)$$

where ρ is the electron density. By multiplying the density by the sign of the second eigenvalue of the density Hessian (λ_2), one can distinguish the strength and the attractive or repulsive nature of the interactions.[23] In the non-covalent region, an arbitrary color code was chosen so that, for NCI surfaces, the red color is used for regions with a positive λ_2 that are associated with steric repulsion; the blue color is used for domains with a negative λ_2 and therefore indicating relatively strong attraction (for example, H-bond regime); the green color is used for regions of very weak reduced density gradient, associated with van der Waals interactions. The ELF calculations were carried out using the TopMod software[24]. The NCI [12,13] results were calculated with the NCIPLOT program.[12,13]

The details for about the QM/MM computation on both reactions are discussed in detail in reference 15 and reference 16.The molecular orbitals obtained from the QM/MM calculations are used for the ELF and NCI calculations. For the ELF calculations, the grid size is 200*200*200 with a step size of 0.1 au. For the supplementary animations, the isovalue for ELF is 0.87, and for NCI is 0.5 with -0.5 au<ρ<0.5 au. For the reaction containing Mn2+ catalyzed by ε subunit, instead of the LANL2DZ pseudopotential and corresponding basis set, a full-electron basis set 6-31G* is adopted to get the molecular orbitals used in the ELF calculations.

2.2 Results and discussion

2.2.1 Pol lambda

The ELF analysis of the critical structures for the reaction catalyzed by Polλ have been presented in a previous study.[17] Herein, only their NCI surfaces are discussed. The NCI analysis shows the evolution of the different interactions along the reaction path (see Figure 2.1). Interestingly, the interactions related to the two legs of the path are already visible at the reactant state. As expected, the strong interaction between HO and OD2 from Asp490 (circled in purple) clearly appears at the reactant state and evolves along the first part of the reaction (see reactant structure in Scheme 2.1 for atom labels). After the proton transfer, a strong hydrogen bond between O3' and HO remains clearly visible. Concerning the second part of the reaction (transphosphorylation), the interaction between O3' and O2α (circled in black), that can be seen from the reactant state, will evolve along the reaction path to give rise to the Pα – O3' bond at the product state. The strength of this basin is characteristic of the van der Waals interactions at the reactant state, it will however evolve into a delocalized double basin whose attractive part is characteristic of the Pα – O3' attraction and repulsive part of the O3' – O2α electrostatic repulsion. This is similar to the evolution of the C – C interaction into a C – C bond and the development of ring tension during the electrocyclization of butadiene as previously described.[14]

The comparison between the Mg^{2+} and Mn^{2+} NCI topologies shows some important differences. The O3'–Mg interaction is weaker that the O3'–Mn one. Conversely, the O3'–O2α interaction is weaker (and so is O_3'–Pα) in the Mn^{2+} structure than in the Mg^{2+} case. Indeed, since Mg^{2+} is a hard cation and Mn^{2+} is a soft one, the O3'–Mg interaction involves less electrons than the O3'–Mn. Therefore more electrons are accessible to the O–H bond and the other lone pair in the Mg case suggesting that (i) the OD2–HO interaction is weaker because HO is less acid and (ii) the O3'–O2α and Pα interaction is stronger. Finally, this is also consistent with our previous QM/MM

energetic results that showed that the barrier for proton transfer is higher when Mg^{2+} is present in the active site compared to Mn^{2+}.[17]

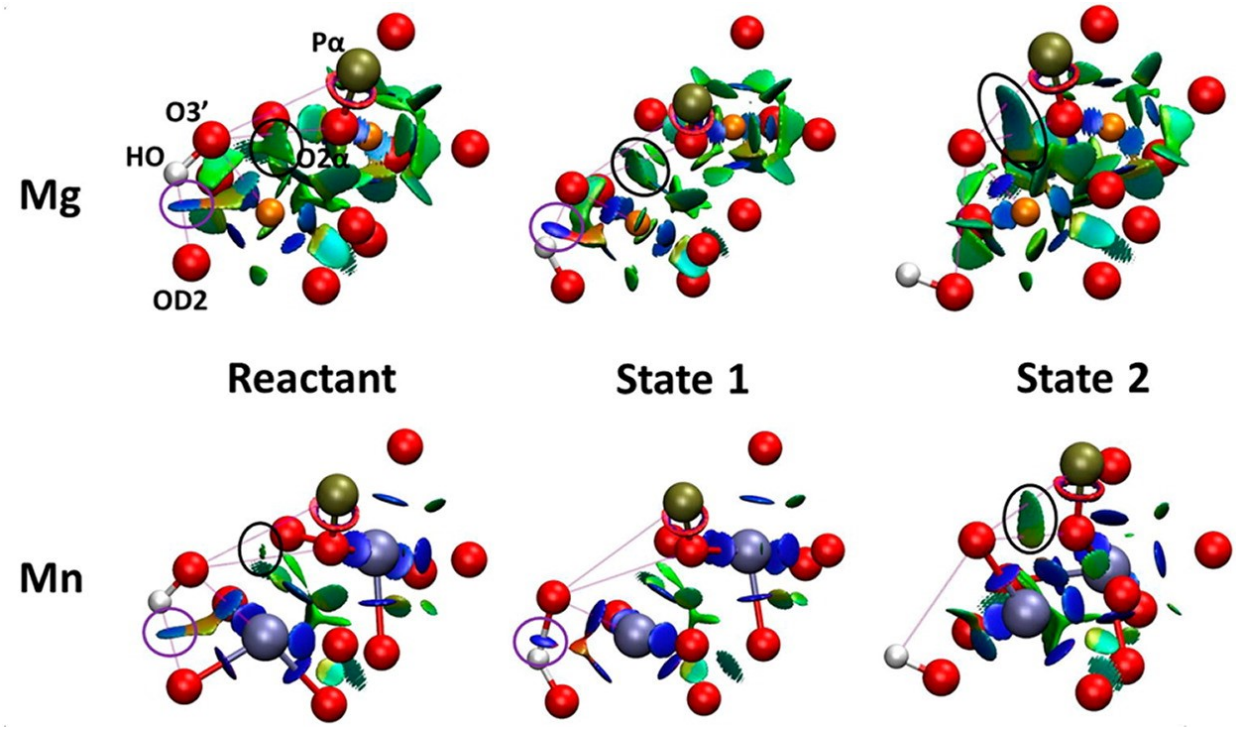
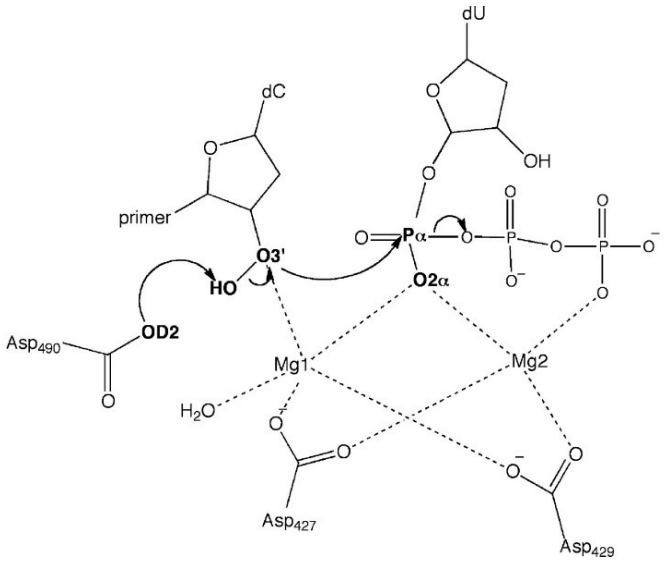


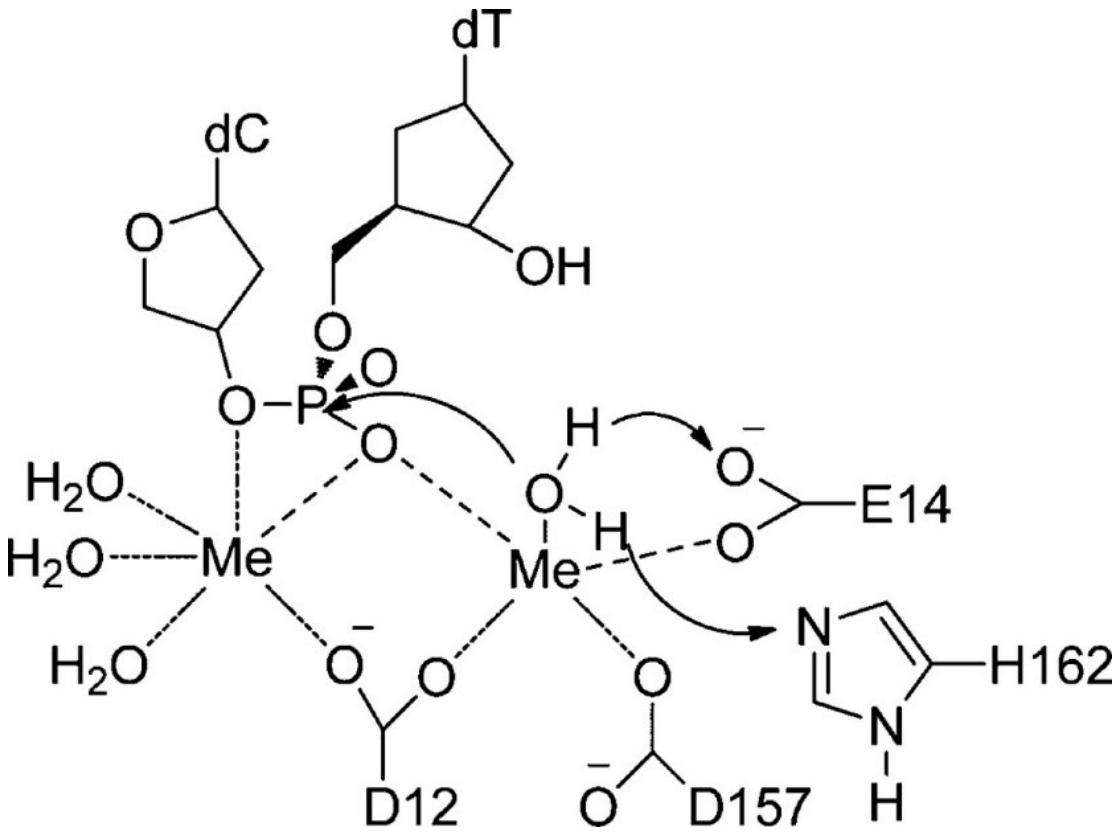
Figure 2.1 NCI analysis for the reactant, TS (state 1), and transphosphorilation structure (state 2) of Polλ with Mg^{2+} and Mn^{2+} (the isovalue is 0.4 au, and the color scale is −0.04 au < sign(λ2)ρ < 0.04 au).



Scheme 2.1 Reaction Mechanism for the Reaction Catalyzed by Polλ[a] . [a]The atom labels used in the text are denoted in the reactant structure.

2.2.2 Epsilon subunit

As described above, ε catalyzes the excision of a nucleotide from DNA (see Scheme 2.2). Figure 2.2 shows the ELF and NCI surfaces of the critical points on the path for the reaction catalyzed by the ε subunit with Mg^{2+} and Mn^{2+}. As shown in Figure 2.3, the basins of the Mg^{2+} ions are spherical as expected. Conversely, there are several basins around each of the Mn^{2+}. This has been observed previously in the context of other metalloproteins and has been termed "subvalence splitting".[17,18] It has been shown that the splitting of the subvalence in metal cations is due to the partial covalency between the metal cation and the ligands around it [18] and may be related to some biological features of the metal cations.[17,18]



Scheme 2.2 Reaction Mechanism for the ε Subunita aIn the first step, a proton is transferred from the ordered water to H162 to activate the nucleophile. Subsequently, the oxygen performs a nucleophilic attack to break the phosphoesther bond coupled with a second proton transfer to E14.(ref 16)

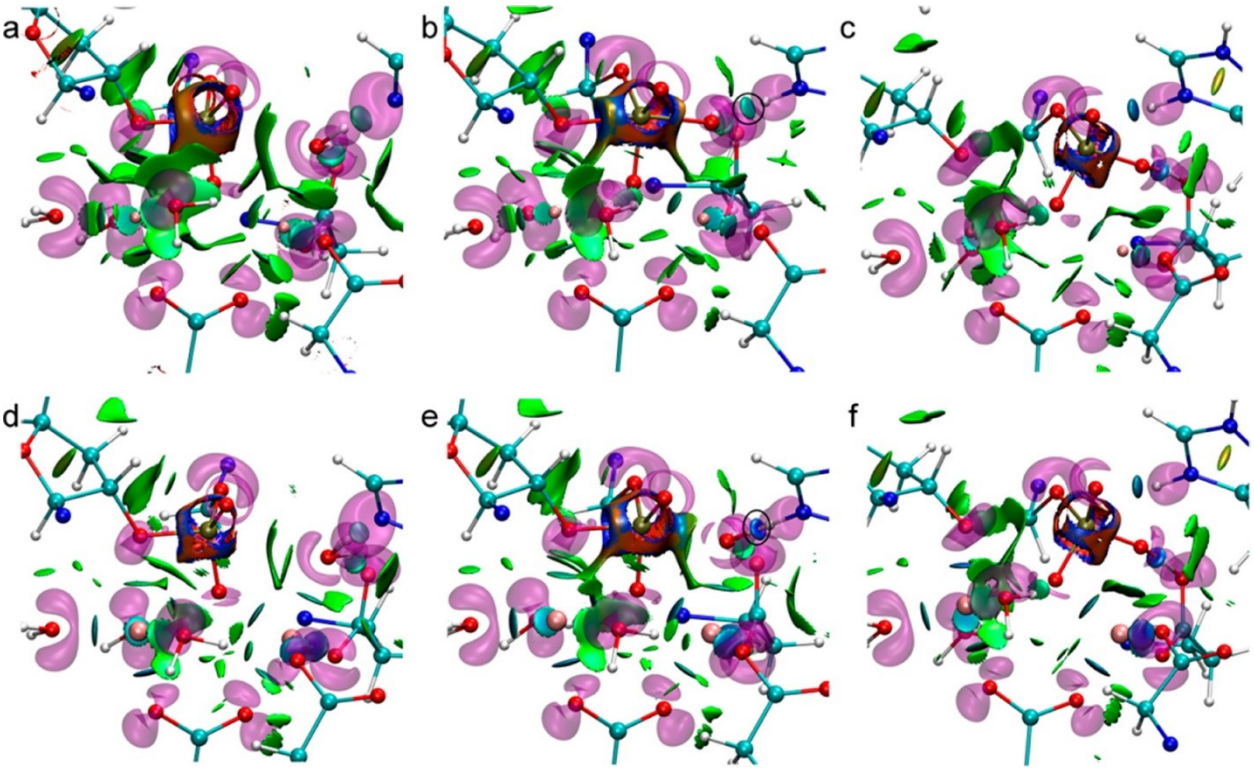


Figure 2.2 Combined ELF and NCI surfaces of the critical structures for the reaction catalyzed by the ε subunit with Mg^{2+} and Mn^{2+}. Only atoms subjected to the ELF analysis are shown; the rest of the QM subsystem is omitted for clarity. Panels a, b, and c show the reactant, TS, and product for Mg^{2+}, respectively. Panels d, e, and f correspond to the reactant, TS, and product for Mn2+, respectively. The isovalue for ELF is 0.87, and for NCI it is 0.5 au with the color scale −0.1 au < sign(λ2)ρ < 0.1 au.

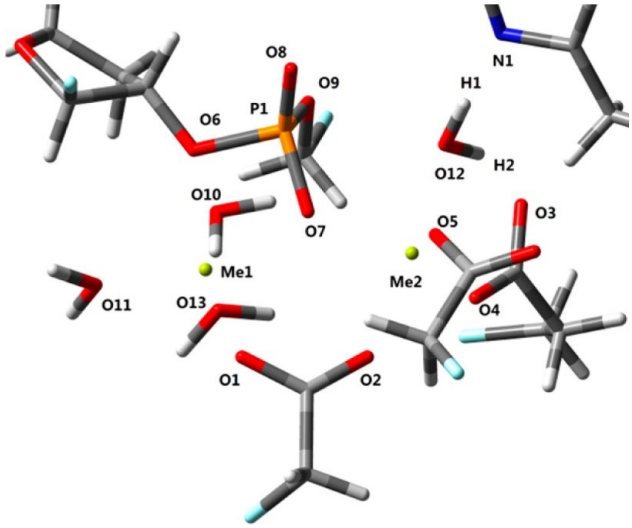


Figure 2.3 Numbering of the atoms in ε that are used in the ELF calculations.

The electronic density of the two Mn^{2+} splits into several distinct basins, while that of the Mg^{2+} cations show only one basin. This suggests that the interactions between the Mn^{2+} and the ligands in their first coordination shell are stronger than those of the Mg^{2+} and its corresponding ligands. For the Mn^{2+} cation, the catalytic metal (Me2) has five subvalence basins while the binding metal (Me1) has six subvalence basins. In this special case, the number of the splitting subvalence basins for each metal ion is equal to the number of ligands in the first coordination shell for Mn^{2+}. In addition, the number of basins in the subvalence for the metals remains constant during the progress of the reaction.

A multipolar expansion of the electron density of each ELF basin can provide further insight into the changes in the basins along the reaction. In this case, the first and second moments of each basin for both metal cations show only slight changes. However, ELF can also provide insights on the bonds that break (between O6 and P1) and form (between O12 and P1) during the reaction by analyzing the evolution of the basins associated with these bonds. Figure 2.4 shows the distance evolution O6-P1 and O12-P1, and the multipolar decomposition evolution (population, first and second moments, M_0, M_1 and M_2 respectively) of two bond basins, V(O6,P1) and V(O12,P1) for the Mn^{2+} and Mg^{2+} catalyzed reactions. In the case of the Mg^{2+} catalyzed reactions, the TS is late in the reaction coordinate, with a longer distance for the breaking bond, d(O6-P1), than for the forming bond, d(O12-P1). When Mn^{2+} is in the active site the TS is located earlier in the path with a significantly shorter distance for the breaking bond and a longer distance for the forming bond. In particular, the two distances and the populations of these two basins change dramatically around the TS.

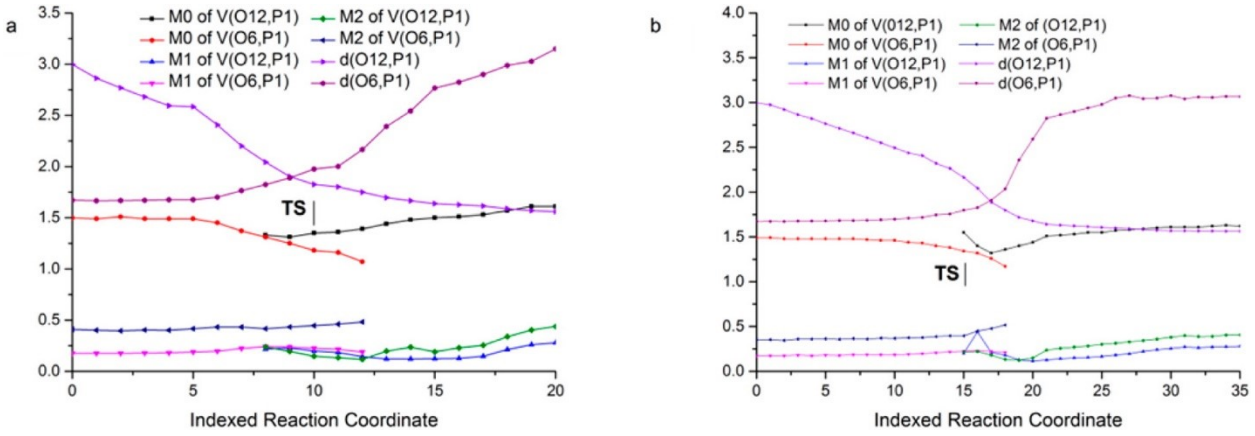


Figure 2.4 Distributed moments, M0 (population), M1 (first moment), and M2 (second moment) of V(O6,P1) and V(O12,P1) and the distances (d(O6,P1) and d(O12,P1), in Å, along the reaction (a, b are for the reactions catalyzed by Mg^{2+} and Mn^{2+}, respectively).

The most significant change is in the population of the basins for the forming and breaking bonds. Interestingly, the appearance of the V(O12,P1) basin (corresponding to the forming bond) shows significant differences between the two cations. In the case of Mg^{2+}, the basin appears before the system reaches the TS with a small population of 1.25, which increases continuously as the reaction reaches the TS and plateaus as the reaction continues to the product. For the Mn^{2+} case, the basin only arises at the TS with a much larger population of around 1.7, which shows a marked decrease and then an increase as the reaction proceeds. This population change indicates that the Mn^{2+} cation induces a larger charge transfer on the three atoms involved in these two basins. The earlier TS and larger population change in the Mn^{2+} catalyzed reaction help explain the lower energy barrier compared to the Mg^{2+} catalyzed reaction. The first and second moments (M_1 and M_2 respectively) for both basins show only modest variations along the path, indicating only a small change in polarization on these basins.

Since most of the time the interactions between the metal and its ligands are not traditional covalent bonding, no typical ELF basins for covalent bonds can be found. In this case NCI plays an important role in the deep understanding of this type of intermolecular interactions. Indeed, the

number of NCI surfaces is able to indicate the coordination number of Mg^{2+} and Mn^{2+} as well. For ε, the catalytic metal (Me2) is surrounded by five NCI surfaces for both cations, which suggests a penta-coordination in both cases. This is consistent with our previous results for the reaction path calculation, which showed that this penta-coordination on the catalytic metal polarizes the water to facilitate the first de-protonation.[16] In contrast, the nucleotide binding metal (Me1) is surrounded by six NCI surfaces for both Mg^{2+} and Mn^{2+}.

For the same metal center, different ligands produce NCI surfaces with different color depths, which can differentiate the strength of the interactions. Generally speaking, the NCI surfaces between ligands and the metal centers for Mn^{2+} are bluer than the ones in Mg^{2+}. This means that the interactions for Mn^{2+} are stronger than Mg^{2+}, which is in accord with the lower energy barrier for Mn^{2+}. For the metal in the reactant structure, the surfaces for O2, O7, O5, and O4, are bluer than the one for the oxygen of the nucleophilic water (O12). For Me1, O6 has the weakest interaction compared to the other ligands coordinated to Me1. Compared to the reactant, the surface between P1 and O12 in the TS becomes bluer, indicating the increasing attraction between these atoms as the bond forms. In addition, the blue surfaces between the transferred hydrogen atoms (H1, H2) show strong hydrogen bonding (circled in black, Figs. 2b/2e) with the corresponding oxygen atoms (O8, O12).

Thus, these two analyses are consistent with each other. Furthermore, regarding the proton H1, which is transferred from O12 to N1, in both TS structures shows a shared basin with the lone electron pair of N_1, which means the proton transfer from the nucleophilic water to N1 has already started. The distance between H1 and N1 is 1.06 Å for the Mg^{2+} catalyzed structure. For Mn^{2+}, the distance is 1.14 Å, close to the value for Mg^{2+}. However, for Mg^{2+}, the population of this shared basin is about 2.72. In comparison, for Mn^{2+}, H1 forms a “bond-like” basin with N with a population of 0.86. The rest of the lone pair of N1 forms an extra basin with a population is 1.07. The NCI surfaces between H1 and O12 (circled in black in Figure 2.2) are much bluer for Mn^{2+} than for

Mg^{2+}, indicating a stronger interaction between these atoms for Mn^{2+} than for Mg^{2+}. This is in good agreement with the smaller basin population for the Mn^{2+} structure.

In general, in terms of NCI surfaces, Mn^{2+} cations have bluer NCI surfaces between the ligands than their corresponding Mg^{2+} cations. With regards to ELF, Mn^{2+} possesses more splitting basins. Both of them are suggestive of the stronger ligand interactions for Mn^{2+}, which is in agreement with its lower reaction barrier than Mg^{2+}.[15,16]

As shown in the movie (http://pubs.acs.org/doi/suppl/10.1021/ct400130b), the NCI surfaces are able to help predict the forthcoming forming or breaking of bonds, and the ELF basins provide the details of existing bonds. When the bonds (for example the bonds between O12 and P1, H1 and N1, H2 and O3) begin to form as the reaction progresses, NCI surfaces appear and become bluer, then they turn into rings (due to the cutoff for the electron density) and enlarge gradually, concurrent with the appearance of ELF basins. When the bonds begin to break (for instance, O6-P1, H1-O12), the ELF basins become smaller, coupled to the emergence of NCI surfaces. Similarly to the findings of Gillet et al.[14], the ELF attractors and the NCI interaction critical points (the holes) overlap. That is, ELF basins appear where NCI surfaces disappear, and viceversa providing ways to anticipate transition states.

2.3 Conclusions

In summary, we have demonstrated the applicability of the recently developed ELF/NCI topological approach to complex enzyme reactions involving large biological systems treated at the QM/MM level. Once again ELF and NCI have been found to be highly complementary, giving information on bond formation and breaking, and on weak interactions thanks to NCI, and on the reorganization of existing bonds and lone pairs with ELF. In addition NCI was shown to be able to provide further insights into the factors governing the lowering of the reaction barrier in Polλ when Mn^{2+} is in the active site compared to Mg^{2+}. In the case of the ε subunit, both ELF and NCI proved to be useful to understand the reasons for the lower barrier in the Mn^{2+} catalyzed reaction

compared to Mg^{2+}. Moreover, the NCI analysis provided further support for the unusual penta-coordination observed in ε for the cation in the metal 2 position, which enables the activation of the nucleophile. Thus, the combination of these two techniques can be a useful and powerful tool to study the interaction changes along the enzyme reactions involved with metals. Finally, extending on the findings of Gillet et al.[14], the ELF attractors and the NCI interaction critical points overlap as NCI critical points appear to be precursors of ELF attractors, providing new strategies to understand and predict the location of transition states in enzymes.

2.4 References

(1) Loeb, L. A.; Monnat, R. J. *Nat. Rev. Genet.* **2008**, 9, 594−604.

(2) Thomas, K. R.; Olivera, B. M. *J. Biol. Chem.* **1978**, 253, 424−429.

(3) Bebenek, K.; Kunkel, T. A. *Adv. Protein Chem.* **2004**, 69, 137−165.

(4) Garcia-Diaz, M.; Bebenek, K.; Gao, G.; Pedersen, L. C.; London, R. E.; Kunkel, T. A. *DNA Repair* **2005**, 4, 1358−1367.

(5) Garcia-Diaz, M.; Bebenek, K.; Krahn, J. M.; Pedersen, L. C.; Kunkel, T. A. *Cell* **2006**, 124, 331−342.

(6) Brenowitz, S.; Kwack, S.; Goodman, M. F.; O'Donnell, M.; Echols, H. *J. Biol. Chem.* **1991,** 266, 7888−7892.

(7) McHenry, C. S. *J. Biol. Chem.* **1991,** 266, 19127−19130.

(8) Derbyshire, V.; Pinsonneault, J. K.; Joyce, C. M. *Methods Enzymol.* **1995**, 262, 363−385.

(9) Becke, A. D.; Edgecombe, K. E. *J. Chem. Phys.* **1990,** 92, 5397−5403.

(10) Piquemal, J. P.; Pilme, J.; Parisel, O.; Ge´ rard, H.; Fourre´ , I.;´Berges, J.; Gourlaouen, C.; De La Lande, A.; Van Severen, M. C.; Silvi,`B. *Int. J. Quantum Chem.* **2008**, 108, 1951−1969.

(11) Silvi, B.; Savin, A. *Nature* **1994**, 371, 683−686.

(12) Contreras-García, J.; Johnson, E. R.; Keinan, S.; Chaudret, R.;Piquemal, J.-P.; Beratan, D. N.; Yang, W. *J. Chem. Theory Comput.* **2011**, 7, 625−632.

(13) Johnson, E. R.; Keinan, S.; Mori-Sanchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W. *J. Am. Chem. Soc.* **2010**, 132, 6498−6506.

(14) Gillet, N.; Chaudret, R.; Contreras-García, J.; Yang, W.; Silvi, B.; Piquemal, J.-P. *J. Chem. Theory Comput.* **2012**, 8, 3993−3997.

(15) Cisneros, G. A.; Perera, L.; García-Díaz, M.; Bebenek, K.; Kunkel, T. A.; Pedersen, L. G. *DNA Repair* **2008**, 7, 1824−1834.

(16) Cisneros, G. A.; Perera, L.; Schaaper, R. M.; Pedersen, L. C.; London, R. E.; Pedersen, L. G.; Darden, T. A. *J. Am. Chem. Soc.* **2009**, 131, 1550−1556.

(17) Chaudret, R.; Piquemal, J.-P.; Cisneros, G. A. *Phys. Chem. Chem. Phys.* **2011**, 13, 11239−11247.

(18) de Courcy, B.; Pedersen, L. G.; Parisel, O.; Gresh, N.; Silvi, B.; Pilme, J.; Piquemal, J. P.´ *J. Chem. Theory Comput.* **2010**, 6, 1048−1063.

(19) Sirover, M. A.; Loeb, L. A. *Science* **1976**, 194, 1434−1436.

(20) Savin, A.; Jepsen, O.; Flad, J.; Andersen, O.; Preuss, H.; von Schnering, H. *Angew. Chem., Int. Ed.* **1992**, 31, 187−188.

(21) Bader, R. F. W.; Austen, M. A. *J. Chem. Phys.* **1997,** 107, 4271−4285.

(22) Pilme, J.; Piquemal, J.-P.´ *J. Comput. Chem*. **2008**, 29, 1440−1449.

(23) Contreras-Garcia, J.; Yang, W.; Johnson, E. R. *J. Phys. Chem. A* **2011,** 115, 12983−12990.

(24) Noury, S.; Krokidis, X.; Fuster, F.; Silvi, B. *Comput. Chem. 1999,* 23, 597−604

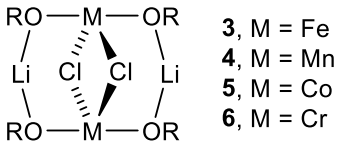
CHAPTER 3

NOVEL ALKOXIDE CLUSTER TOPOLOGIES FEATURING RARE SEESAW GEOMETRY AT TRANSITION METAL CENTERS

Adapted with permission from *Chem. Eur. J.* **2013,** 19, 12225-12228

3.1 Introduction

While most of the four-coordinate compound are either tetrahedral or square planar, the seesaw shape is also possible according to valence shell electron pair repulsion (VSEPR) theory. Some seesaw compounds can be found for main-group elements. However, it is quite rare for transition metal to have a seesaw geometry. It attracts a lot of interest because the vacancy around the metal are plausible for substrate coordination and activation.[1] Previous study has proposed that only the metal centers with d^6 and closed-shell d^{10} possess the unusual seesaw shape. James et al. in Groysman group have synthesized same seesaw compounds with a series of transition metals (Cr^{II}—Co^{II}) featuring varying electron configurations (d^4–d^7). [1,2] The structure of compound is shown in Figure 3.1. It has the steric bulk of the ancillary ligands, alkoxides, combined with the inclusion of lithium chloride in the structure. We carry out Density function theory (DFT) colocations combined with ELF/NCI analysis to understand the bonding information and provide insights into their electronic structures.



3, M = Fe
4, M = Mn
5, M = Co
6, M = Cr

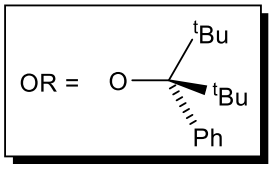


Figure 3.1 The structure of the seesaw compounds synthesized by Groysman group

3.2 Methods

The geometry optimization is performed with B3LYP/SDD using Gaussian09 program[3]. The wavefunction for ELF and NCI calculations is obtained from the single point calculation using B3LYP/6-31(d). ELF was calculated with a modified TopMod package[4,5]. A size of 200^3 au with a step size of 0.1 au grid is used for the ELF calculations. The NCI calculation was performed the NCIPLOT program[6,7]. *t*Bu groups are replaced by Me groups for calculations.

3.3 Results and discussion

Continuous symmetry measurements provide a useful handle to evaluate the geometry of the metal center.[8,9] The continuous symmetry values have been calculated using the method detailed by Cicera, Alemany, and Alvarez in the Shape2 software.[1] The continuous symmetry values for the crystal (computationally optimized) structures (**3**–**6**) are 2.2 (2.6), 2.2 (2.4), 2.0 (2.9) and 2.1 (2.5), respectively. As the continuous symmetry values range from 0 to 100 (0 means perfect seesaw, whereas 100 means totally distorted), both crystal and computational structures present geometry close to that of a perfect seesaw.

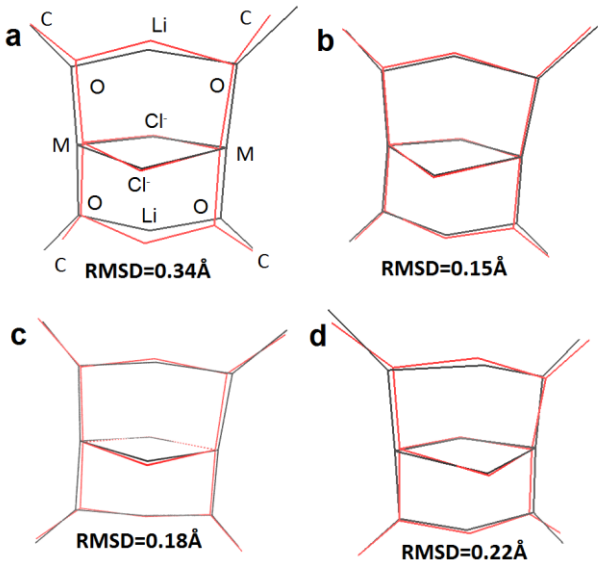


Figure 3.2 Superposition of the crystal structures (black) and the calculated structures (red). Only the core part is shown. The letters in panel a are the atomic names; M stands for metals. For (a) Cr; (b) Mn; (c) Fe; (d) Co.

The superposition of the calculated and crystal structures shows good agreement between them as shown in Figure 3.2. Combined electron localization function (ELF)/noncovalent interaction (NCI) analysis was performed to gain a better understanding of interatomic interactions. As shown in Figure 3.3, the ELF basins for the outer shell electrons of the transition metals are split into several small basins (in circle). This is due to the metal interactions with its surrounding ligands.[10]Different metals have different orientations for the splitting basins. As ELF values inside are larger than the space outside the black surface, it means the electron is more localized inside than outside the surface. The orientations of these splitting surfaces should give the information about positions where are better for an electrophilic attack.

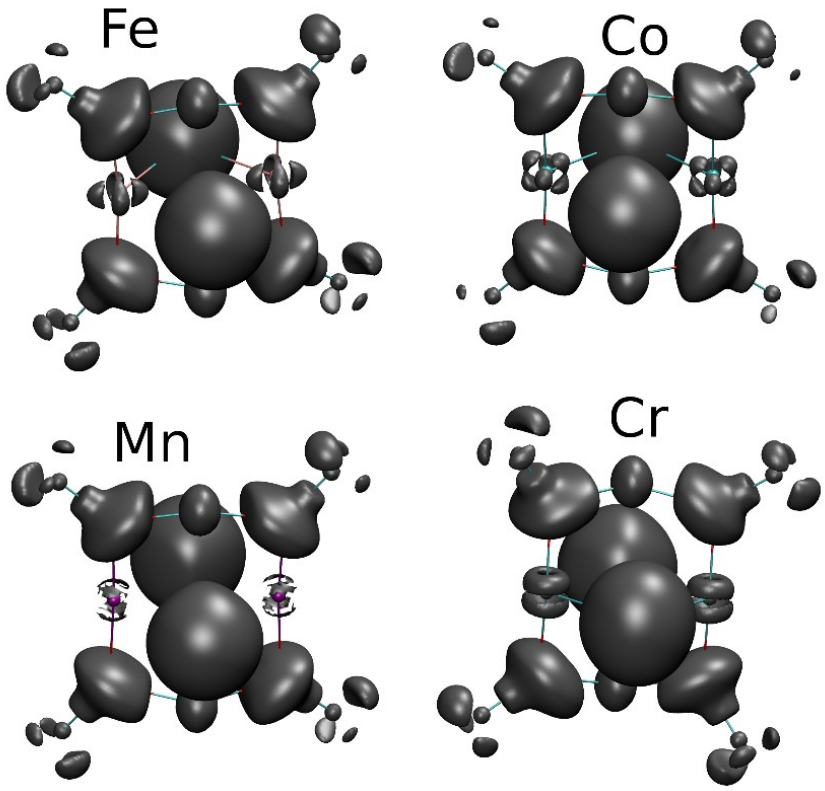


Figure 3.3 The ELF (a) and NCI (b) surfaces for the Cr, complex. For ELF, the isovlaue is 0.735.

As mentioned in previous chapter, NCI describes interatomic interactions by peaks of the reduced electron density gradient at low electron densities. The NCI surfaces are shown in Figure 3.4, 3.5 and 3.6. Taking **6** as an example, the strong metal—ligand (Cr—O and Cr—Cl) interaction is evidenced by the blue NCI surfaces in Figure 3.4a. In addition, the surface for Cr—O (circled in black) has a darker blue color than the surface for Cr—Cl (circled in red), indicating that the interaction between Cr and Cl is slightly weaker than between Cr and O. Moreover, the interaction between Li and Cl is much weaker than the one between Cr and Cl as demonstrated by the green surfaces (circled in yellow). Next, we replaced the alkyl substituents R of the OR groups by hydrogen atoms to investigate their effect on the stability of the seesaw geometry in the core. Although the optimized Cr complex did not retain the seesaw geometry, the other (Mn—Co) complexes were found to only undergo slight distortion (Figure 3.5). These results underscore the importance of lithium—oxygen interactions (circled in blue) in these structures for the stabilization of the seesaw core. To determine the steric effect of the [*t*Bu$_2$(Ph)CO] ligand, we carried out the NCI analysis on the full crystal structure of **6** (unoptimized, Figure 4.4b). Some of the notable interactions include a weak attraction between Cr and *ortho*-H (circled in red), and a weak attraction between Cl and H (*t*Bu) atoms (circled in blue). Similar interactions are also present in the optimized simplified structures bearing Me substituents (Figure 3.6). However, the full structure also demonstrates a weak attraction between the *t*Bu groups and neighboring phenyl group (circled in violet). This interaction may be responsible for the formation of the cage-like structure.

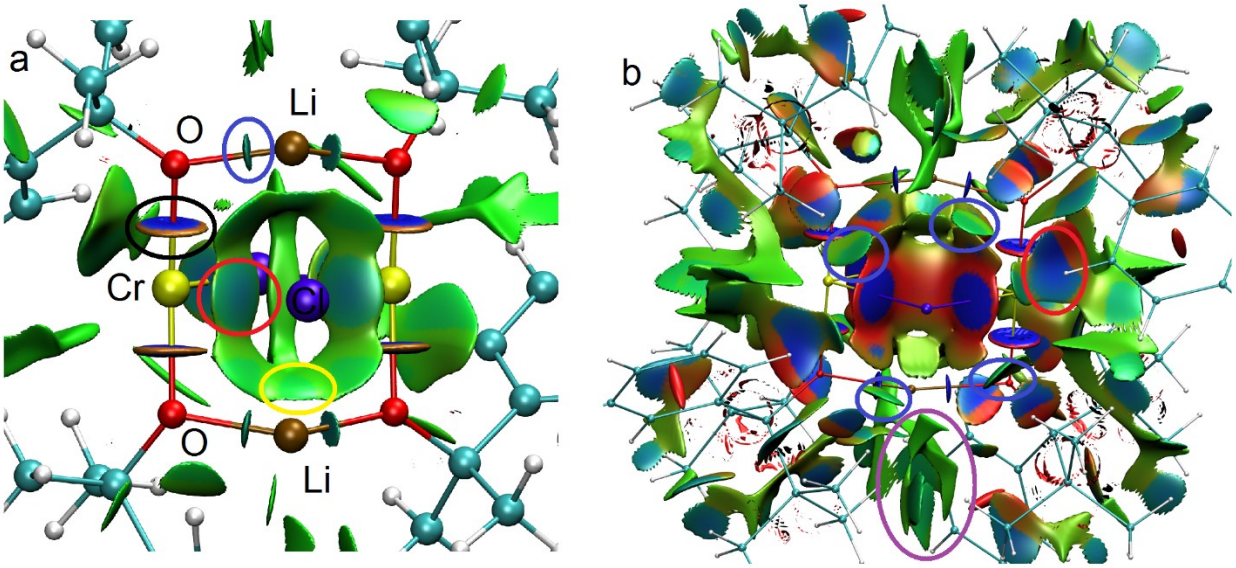


Figure 3.4 NCI surfaces for complex **6**, the isovalue is 0.5. a) The optimized structure with *tert*-butyl was replaced by methyl (only the core is shown), $-0.07<\text{sign}(\lambda_2)\rho<0.07$ a.u. b) NCI analysis for the complete (unoptimized) crystal structure, $-0.02<\text{sign}(\lambda_2)\rho<0.02$ a.u. Regarding the chosen arbitrary color code, red surface indicates strong repulsion; blue surface indicates strong attraction; green surface indicates relatively weak interactions. Different $\text{sign}(\lambda_2)\rho$ ranges may be used to distinguish the strength of the interactions. Note that one can compare the relative strength of the interactions based on the colors only when using the same $\text{sign}(\lambda_2)\rho$ range.

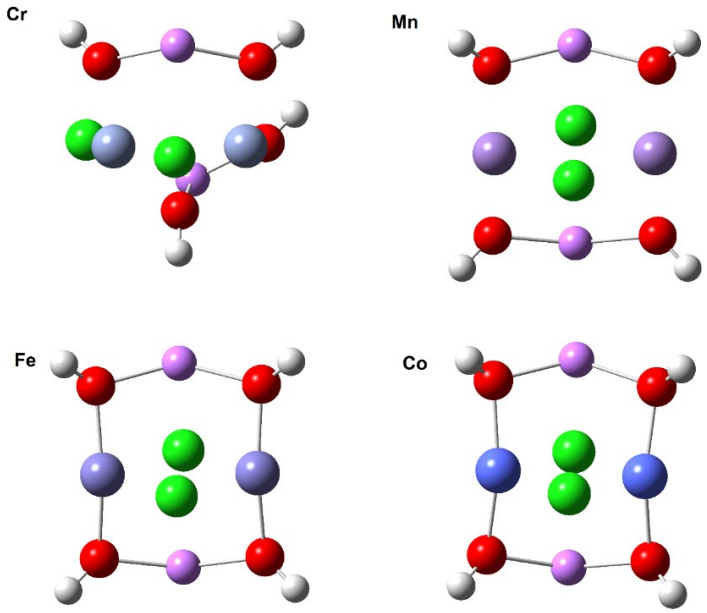


Figure 3.5 Optimized structures with the ligands being replaced with H.

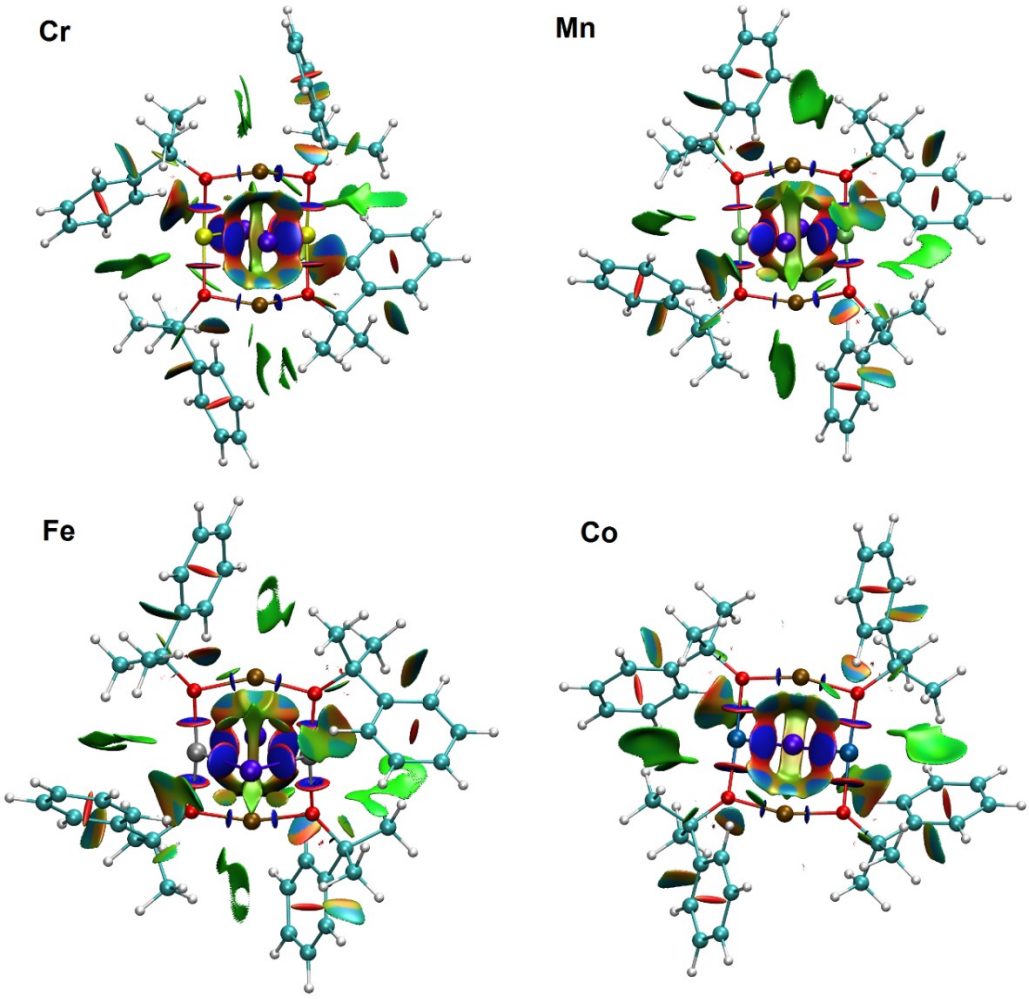


Figure 3.6 The NCI for complex **3-6** with tert-butyl being replaced with methyl**.** The isovalue is 0.5 and -0.02<sign(λ_2)ρ<0.02 au.

3.4 Conclusions

In summary, we have theoretically studied the seesaw metal clusters synthesized by Groysman group. DFT calculations have confirmed their unusual seesaw shape. Moreover, the combined ELF/NCI analysis have revealed the important intermolecular interactions that may play an important role in maintaining the unusual seesaw shape.

3.5 References

(1) Cirera, J.; Alemany, P.; Alvarez, S. *Chem. Eur. J.* **2004**, *10*, 190.

(2) Oliván, M.; Clot, E.; Eisenstein, O.; Caulton, K. G. *Organometallics* **1998**, *17*, 3091.

(3) Frisch, M. J.; Trucks, et al.; Gaussian, Inc.: Wallingford CT, 2009.

(4) Noury, S.; Krokidis, X.; Fuster, F.; Silvi, B. *Computers & Chemistry* **1999**, *23*, 597.

(5) Pilmé, J.; Piquemal, J.-P. *Journal of Computational Chemistry* **2008**, *29*, 1440.

(6) Johnson, E. R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W. *Journal of the American Chemical Society* **2010**, *132*, 6498.

(7) Contreras-García, J.; Johnson, E. R.; Keinan, S.; Chaudret, R.; Piquemal, J.-P.; Beratan, D. N.; Yang, W. *Journal of Chemical Theory and Computation* **2011**, *7*, 625.

(8) Pinsky, M.; Avnir, D. *Inorg. Chem.* **1998**, *37*, 5575.

(9) Zabrodsky, H.; Peleg, S.; Avnir, D. *J. Am. Chem. Soc.* **1992**, *114*, 7843.

(10) de Courcy, B.; Pedersen, L. G.; Parisel, O.; Gresh, N.; Silvi, B.; Pilmé, J.; Piquemal, J. P. *J. Chem. Theory Comput.* **2010**, *6*, 1048.

CHAPTER 4

AB INITIO QM/MM CALCULATIONS SHOW AN INTERSYSTEM CROSSING IN THE HYDROGEN ABSTRACTION STEP IN DEALKYLATION CATALYZED BY ALKB

Reproduced with permission from *J. Phys. Chem. B,* **2013,** 117, 6410-6420

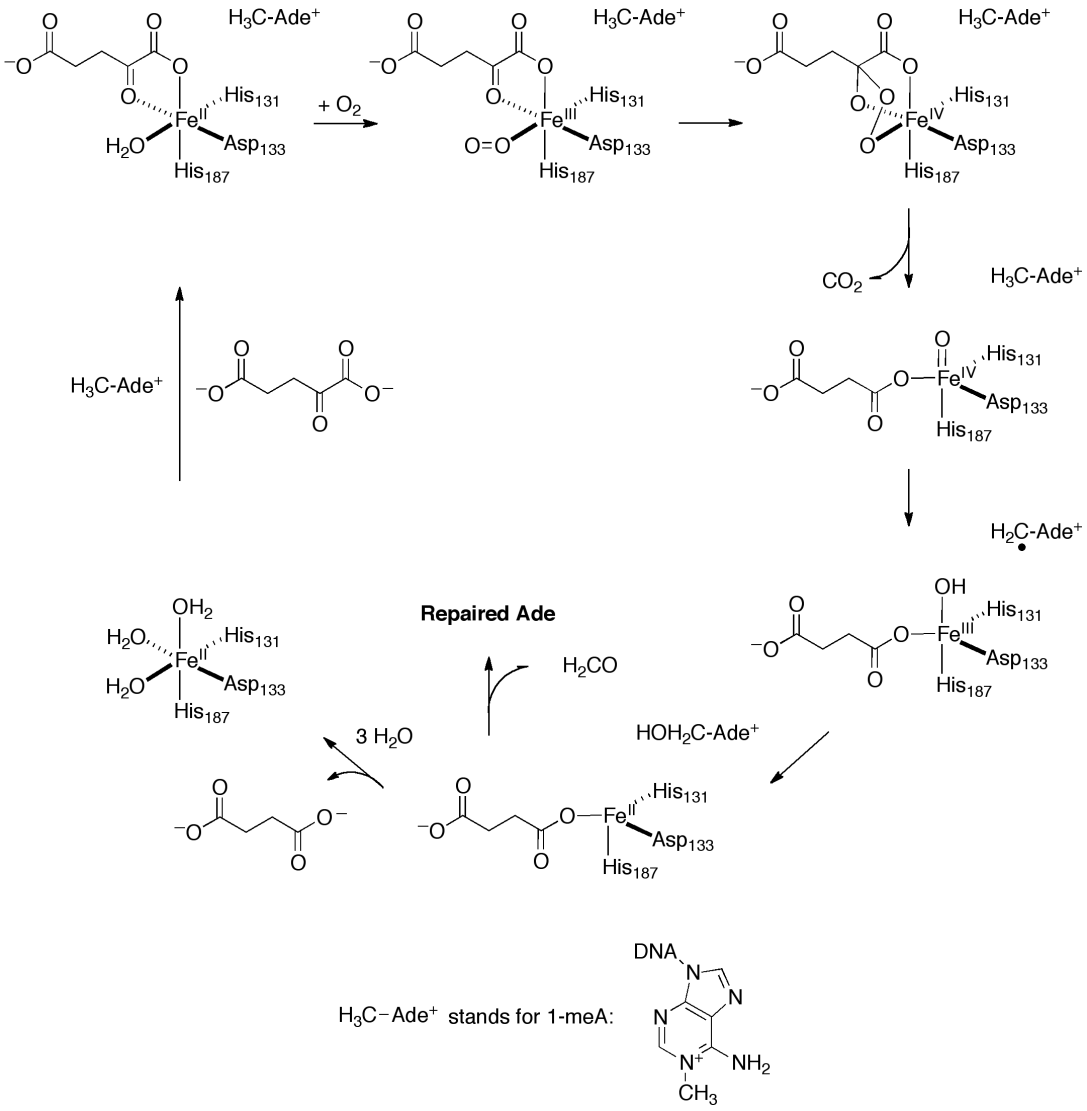
4.1. Introduction

DNA is subject to damage by various environmental or chemical agents. One of the possible damages is the alkylation of DNA bases. Alkylation lesions to the DNA bases can be caused by endogenous or exogenous substrates and can lead to mutagenic and cytotoxic consequences. These effects can have negative consequences, but have also been used advantageously in cancer therapy with a proper amount of alkylating agents.[1-3] Three main chemical mechanisms that can remove such lesions have been discovered so far[1]: base excision repair, direct dealkylation by an alkyltransferase, or oxidative dealkylation. The last mechanism involves the oxidation of the alkyl group with the repaired base and oxidized alkyl moiety as products.[4,5]

A family of proteins called AlkB has been recently found to use this last mechanism to directly repair alkylated bases such as 1-methyladenine (1-meA) and 3-methylcytosine (3-meC).[6-8] The AlkB family is comprised of the title AlkB enzyme from *E. coli* and nine human homologues, ABH(1-9). These enzymes are part of the non-heme iron and α-ketoglutarate (α-KG) dependent super family of enzymes.

The proposed reaction mechanism is shown in Scheme 4.1. Briefly, these enzymes bind the alkylated base, α-KG and molecular oxygen. This is followed by the oxidation of α-KG, the release of CO_2 and formation of a high spin ferryl (Fe^{IV}-oxo) intermediate. This intermediate abstracts a

hydrogen from the alkyl group, followed by an OH rebound to the alkyl radical. Finally, this results in the oxidative dealkylation of the base and release of the products.



Scheme 4.1 Proposed mechanism for the direct dealkylation of the 1-meA catalyzed by AlkB.

This mechanism has been assumed to be general for the α-KG dependent super family. However, there are issues still to be resolved. For example, recently an alternative mechanism for Fe/α-KG dependent enzymes, which avoids the OH rebound step, has been proposed.[9] In the specific case of the AlkB family, the crystal structure of a zwitterionic intermediate after the transfer of the

hydrogen atom from the hydroxymethyl group to a surrounding hydroxide ion for 3-meC has been captured crystallographically and verified by QM/MM calculations, which raises the possibility of an alternative pathway for the last steps before dealkylation of the base.[10] The verifications of these alternative mechanisms, which will be our future work, require a better understanding of the reactivity of this enzyme family.

AlkB relies on a ferryl (Fe^{IV}-oxo) intermediate for the hydrogen abstraction process. This is an important active moiety prevalent in many Fe- and α-ketoglutarate- dependent dioxygenase catalyzed reactions. For example, taurine dioxygenase (TauD), which catalyzes the release of sulfite, involves the same intermediate and has been extensively studied theoretically and experimentally.[9,11-18] All theoretical work modeling reactivity of AlkB related enzymes has focused on model systems to our knowledge.[15-17] For AlkB, Liu et al. used DFT calculations combined with the self-consistent reaction field to study a model system of the AlkB active site.[19] The triplet, quintet, and septet potential energy surfaces were explored and the rate-determining step was found to be hydrogen abstraction. This finding agrees with rapid freeze-quench Mössbauer spectroscopy results for TauD[20] Later, similar results were reported by means of another DFT method (BP86+10%HF) on a smaller model system.[21]

One crucial issue in the study of the hydrogen abstraction by Fe^{IV}-oxo is its complicated electronic structure, especially its various spin states.[19,21] An Fe^{III}–oxyl radical has been proposed to form when approaching the activated C–H bond for a TauD model system.[22] Furthermore, two distinct channels with the same overall spin state (triplet and quintet) have been investigated for some models.[23,24] These pathways have been termed π and σ depending on the orbital occupied upon hydrogen atom transfer. If a spin-down (relative to iron's unpaired electrons) electron is transferred into the Fe–O π^* orbital, this defines a π trajectory. Conversely, if the electron is spin-up and is transferred into the Fe–O σ^* orbital, a σ trajectory is obtained. Although these two pathways start from the same reactant, the transition states have different Fe–O–H angles in

order to obtain the largest orbital overlap between the Fe–O and the C–H (120° for π, 180° for σ) in model systems. In biological systems the group donating the abstracted hydrogen is not free to move, and the plausibility of both pathways may be greatly influenced by the enzyme active site.

In this chapter, we present computational reactivity results for the σ and π channels for the hydrogen abstraction step in the full AlkB enzyme by employing QM/MM methods, especially highlighting a comprehensive understanding of the electronic structure of Fe^{IV}-oxo and its evolution along the hydrogen abstraction process under enzyme evironment. According to previous experimental and theoretical cluster model results on AlkB and TauD, the ground states correspond to open shell systems [12,19,25,26]; therefore, we will focus on the quintet and triplet surfaces. As described in detail below, our results show that the reactant is an Fe^{III}-oxyl with an intermediate-spin Fe (*S*=3/2) ferromagnetically coupled to the oxyl radical. As the reaction proceeds, an intersystem crossing is observed between the quintet $^{IS}Fe^{III}$-oxyl and the quintet state with a high-spin Fe (*S*=5/2) antiferromagnetically coupled to the oxyl radical. Moreover, energy decomposition analysis reveals nine catalytically important residues around the active site, several of which are conserved on two human homologues. It is anticipated that this study will provide helpful guides not only on the role of catalytically important residues and the chemistry of non-heme complexes but also to heme complexes containing the Fe^{IV}-oxo moiety by highlighting how the enzyme active site influences local electronic and geometric structures.

4.2. Computational Methods and Theory

After the addition of hydrogen atoms, solvation and charge neutralization, the crystal structure of AlkB (PDBID: 2FDG[8]) was subjected to 10ns Molecular Dynamics (MD) simulations in the NVT (Canonical) ensemble at 300K with the Amberff99 force field using a 1fs step size. All MD simulations were performed with the PMEMD program in AMBER11[27]. Ten stable structures were extracted from the MD simulation to perform QM/MM optimizations on the reactant state. The

snapshot with the lowest energy was chosen for further QM/MM calculations. All the QM/MM calculations were performed by means of an in-house program that links modified versions of Gaussian09[28] and TINKER[29] to perform additive QM/MM with electrostatic embedding. An iterative method was adopted in the QM/MM optimization.[30] The boundary atoms were modeled by the pseudobond approach[31], which provides a smooth connection at the QM/MM interface. Frequency calculations were performed for all critical point structures and it was found that reactants and intermediates have no imaginary frequencies while TSs have only one imaginary frequency. After obtaining the reactants and intermediates, the Quadratic String Method (QSM) [32] was used for the pathway optimization between critical structures. In all QSM optimizations, a constrained MM path optimization was performed.[33] For the initial guess to the the path the MM environment from one of the end points (reactant) was employed for all the images along the path. The path was optimized in 9 steps by applying constraints on the MM atoms during the MM part of the optimization and these constraints were reduced gradually after each QM/MM path optimization cycle. The TS structures were optimized with the QST3 method taking the highest energy point from QSM results as an initial guess. The minimum energy crossing point between the two subsidiary states of the quintet was optimized with the Lagrangian multiplier approach[34]. For the free energy perturbation[30,33], each point on the minimum energy path was first equilibrated for 20ps, followed by 20ps of production calculations in both the forward and backward direction. The free energy corrections to the QM were obtained by performing harmonic vibrational frequency calculations at 300.15 K with a scaling factor 0.975 [35](the factor for ωB97XD/def2-TZVP). For the energy decomposition analysis, Coulomb and Van der Waals interactions between the QM subsystem (approximated by ESP fitted charges) and each residue in the MM environment are calculated for an ensemble of structures (10 ns run with snapshots every 100 ps) and the values are averaged. The difference between the average interaction energies in the reactant and the TS structures can qualitatively indicate the effect the residue has on the reaction barrier.

The electron localization function (ELF), which was originally proposed based on the HF approach[36] and was extended for DFT[37], can measure the electron localization in molecular systems. The topological analysis of ELF, a continuous and differentiable scalar field in three-dimensional space, is performed by calculating its gradient vector field and associated Hessian matrix. The points where the gradient is zero and the three eigenvalue of the Hessian are all negative are defined as attractors. The points whose gradient paths end at the same attractor belong to the same basin. The ELF basins resemble the domains of the models from the valence shell electron pair repulsion (VSEPR) theory.[38] According to the types of the electrons located in the basins, the basins can be divided into two types: core (C()) and valence (V()). For the valence basins, they can be lone pairs or chemical bonds (V(A,B), where A, B are the atoms included the basin). The first moments and second moments are the dipolar and quadrupolar polarization components of the charge distribution. ELF was calculated with a modified TopMod package[39,40]. A size of 200^3 au with a step size of 0.1 au grid is used for the ELF calculations. For the ELF calculation, the wavefunction are truncated only including the first layer atoms around the iron. The isovalue for the visualization is 0.89.

Non-Covalent Interaction (NCI) has been developed to plot the reduced density gradient versus the product of the sign of the second eigenvalue (λ_2) of the electron-density Hessian matrix and the electron density in order to visualize the non-covalent interaction. The desired non-covalent interaction is characterized by the peaks at low electron density. For NCI surfaces, according the color scale we choose, red color is a sign of relatively strong repulsion; blue color indicates relatively strong attraction; green color means relatively weak interaction. For NCI calculations, because of NCIPlot's limit on the orbital type, the wavefunction from the single point calculation with the basis set LANL2DZ and its corresponding pseudopotential was used instead of the one from 6-31G(d,p). The NCI calculation was performed the NCI Plot program[41]. The isovalue for the plot is 0.5 au.

Corresponding orbital analysis[42] maximally matches up the α and β orbitals by means of biorthogonalizing the α and β molecular orbitals. It allows us to unambiguously assign magnetically coupled orbitals for the broken symmetry states.

4.3. Results and Discussion

The organization of the chapter is as follows: We begin with discussion on the spin states of the ferryl moiety, followed by tests on different functionals to determine the most appropriate one for this system. This is followed by a discussion of the QM/MM potential and free energy paths. Subsequently, the calculation of an inter-system crossing is presented. Energy decomposition analysis is performed to investigate the role of individual residues. Finally, topological analysis of the electronic localization function (ELF) and non-covalent interaction (NCI) analyses on the QM/MM wavefunction are discussed.

4.3.1. Spin states of the ferryl moiety

According to the proposed reaction mechanism, the iron-oxo moiety bears two positive charges. Ideally, as shown in Figure 4.1, the charge of iron should be +4, and the charge of O should be -2. However, electron transfer between Fe and O is possible and both Fe^{IV}–O^{2-} and Fe^{III}–O^{-} should be considered. In O^{-} the electron can be spin-down or spin-up, so there are two sub-states for the quintet and triplet surfaces corresponding, respectively, to antiferro- or ferromagnetic coupling to iron (see Figure 4.1). We will use the shorthand notations in Figure 4.1 for these sub-states in the remainder of this work. Note that for the quintet surface, the two sub-states correspond to σ (antiferromagnetic) and π (ferromagnetic) reactivity channels. One of our key new findings is that these two channels begin from two reactants with different electronic structures and lie on two surfaces in the enzyme system. In addition, the Fe^{III}-oxyl has already formed at the reactant for our system and will be verified in the following discussion.

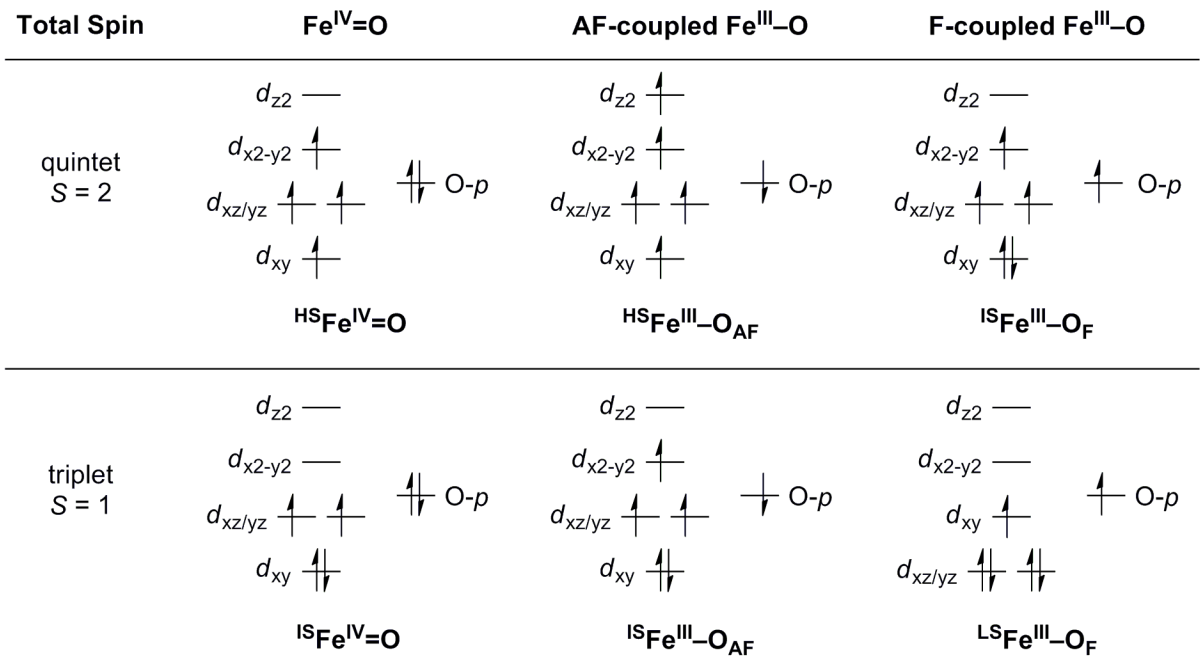


Figure 4.1 Electron configuration diagram of the iron(IV)-oxo moiety in the quintet and triplet states in an ideal octahedral ligand field. (HS, IS and LS represent high-spin, intermediate-spin and low-spin on Fe, respectively; AF and F stand for antiferromagnetic and ferromagnetic.)

4.3.2. DFT functional benchmarks, cluster and QM/MM calculations

Since previous studies on models have shown the quintet state is lower than the triplet state, we have performed tests with five different DFT functionals only on the quintet surface. The functionals tested for the QM subsystem are: B3LYP*[43], B3LYP[44,45], ωB97XD[46,47], M06, and M06-2X[48,49]. It can be seen from Table 4.1 that only two methods (M06-2X and ωB97XD) are able to capture both sub-states for all critical points (i.e. Reactant, TS and Intermediate). In addition, B3LYP*, B3LYP and ωB97XD predict the reactant is in $^{IS}Fe^{III}$-O_F state while M06 and M06-2X predict the state for the reactant is $^{HS}Fe^{III}$-O_{AF}. These results can be qualitatively explained by the competition between the exchange stabilization arising from the unpaired electron in the iron orbitals and the orbital energy difference between the oxygen and iron [24]. In terms of the HF exchange in each functional, B3LYP* has 15%, B3LYP has 20%, ωB97XD has 22%, M06 has 27%, and M06-2X has 54%. More HF exchange in the functional will introduce more exchange

stabilization effects. For B3LYP*, B3LYP and ωB97XD, the stabilization from the electron being transferred to a lower iron's orbital in $^{IS}Fe^{III}$-O_F than $^{HS}Fe^{III}$-O_{AF} dominates. However, for M06 and M062X, the increased exchange stabilization takes the dominant role.

Table 4.1 Relative QM/MM energies in the quintet state calculated by five different DFT functionals (unit: kcal/mol; for each functional, the lowest absolute energy is defined as zero; dash line means data not obtained).

Funtional	State	Reactant	TS	Intermediate
B3LYP*	$^{HS}Fe^{III}$-O_{AF}	----	29.0	11.0
	$^{IS}Fe^{III}$-O_F	0.0	----	14.1
B3LYP	$^{HS}Fe^{III}$-O_{AF}	----	27.4	4.6
	$^{IS}Fe^{III}$-O_F	0.0	----	10.9
ωB97XD	$^{HS}Fe^{III}$-O_{AF}	10.6	25.9	-0.8
	$^{IS}Fe^{III}$-O_F	0.0	46.5	6.9
M06	$^{HS}Fe^{III}$-O_{AF}	0.0	14.0	-13.0
	$^{IS}Fe^{III}$-O_F	----	----	6.9
M062X	$^{HS}Fe^{III}$-O_{AF}	0.0	18.3	9.2
	$^{IS}Fe^{III}$-O_F	16.6	40.0	-11.7

Our results also suggest that the magnitude of the barrier is highly related to the Hartree-Fock (HF) exchange percentage in the hybrid functional. The reaction barrier is observed to decrease with the increase of HF exchange as observed for the barrier calculated with B3LYP*, B3LYP and ωB97XD. This is because the TS is in $^{HS}Fe^{III}$-O_{AF} state in which the exchange stabilization dominates, compared to the $^{IS}Fe^{III}$–O_F reactant. In contrast, for M06 and M062X, the reaction barrier increases with the increase of HF exchange in the functional. This is because both the

reactant and the TS are $^{HS}Fe^{III}$-O_{AF}, and the exchange stabilization effects on the reactant are larger than the TS when the HF exchange in the functional increases.

The free energy barrier estimated from experimental kinetic results corresponds to 19.8 kcal/mol[50]. The values of the barrier predicted by B3LYP*, B3LYP, and ωB97XD are higher than this experimental value while M06-2X and M06 underestimate the barrier. Although M06-2X gives the closet potential energy barrier to the experimental value, its HF exchange percentage is too high, and it is not recommended to calculate properties for compounds that contain transition metals[48,49]. In addition, when considering the free energy correction to the potential energy, M06-2X should give a lower value than the experimental one. Furthermore, it has been discovered that the dispersion correction plays an important role in evaluating the barrier for the hydrogen abstraction catalyzed by Cytochrome P450[51,52], a heme enzyme that also utilizes the active iron-oxo species for the hydrogen abstraction. The only functional that includes dispersion corrections in our test is ωB97XD. Therefore, based on these factors we decided to use ωB97XD for the remainder of our QM/MM calculations.

Cluster calculations in the gas phase were also carried out to investigate the effects of the MM environment. After examining the optimized structures of the critical points obtained using ωB97XD, two waters in the MM environment were always found to be within 6 Å of the iron. Consequently, aiming to study the effects of these two water molecules on the reaction, we performed QM/MM calculations including these two waters in the QM region.

Our results indicate that these waters have a large impact on relative energies but a small effect on geometries. As shown in Figure 4.2a, the calculated minimum energy barrier corresponds to the difference between the $^{IS}Fe^{III}$–O_F reactant and the $^{HS}Fe^{III}$–O_{AF} TS. The calculated reaction barrier is 22.4 kcal/mol, which is slightly higher than the approximate free energy barrier of 19.8 kcal/mol obtained using transition state theory from the experimental results[50]. The spin densities verify the existence of the sub-states discussed above in the quintet and triplet states (see Figure

4.2a). According to a corresponding orbital analysis on the $^{HS}Fe^{III}$–O_{AF} TS, the spin-up electron in the d_{z2} orbital of Fe is antiferromagnetically coupled to the spin-down electron in the forming O–H / breaking C–H σ orbital.

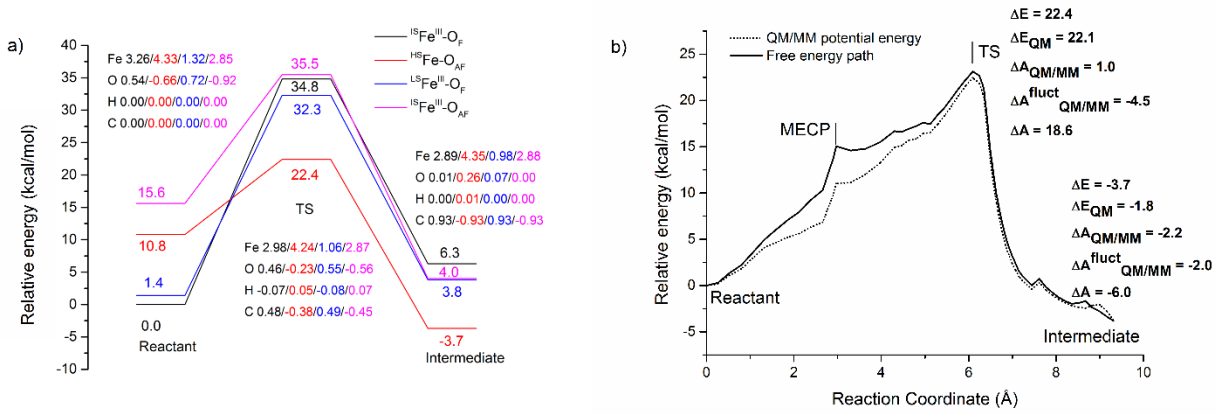


Figure 4.2 (a) Relative energies of the critical structures and and spin densities for some key atoms in the quintet and triplet states by ωB97XD with two waters included in QM subsystem (unit: kcal/mol; for the absolute energy of the $^{IS}Fe^{III}$–O_F reactant is defined as zero). (b) Calculated minimum potential and free energy paths (MEP and MFEP respectively). Potential and Helmholtz free energy differences (unit: kcal/mol) are also shown. All the energies are relative to the reactant. ΔE represents the total potential energy difference. ΔE_{QM} is the energy difference for QM subsystem. $\Delta A_{QM/MM}$ is the free energy change in the QM/MM interaction. $\Delta A^{fluct}{}_{QM/MM}$ is the contribution from the QM subsystem fluctuation. The total free energy difference ΔA is the sum of ΔE_{QM}, $\Delta A_{QM/MM}$ and $\Delta A^{fluct}{}_{QM/MM}$.

For the minimum energy reactant, the calculated spin density of the iron is 3.26, and the oxygen is 0.56. This result is consistent with the nonheme and heme iron^{IV}-oxo in previous experimental and theoretical studies[12,13,53,54]. In addition, as shown in Figure 4.1, in spite of more exchange stabilization for the $^{HS}Fe^{III}$–O_{AF} reactant since it has more single spin electrons[24], the electron from the oxo occupies a high-energy antibonding orbital making this state disfavored. In contrast, at the TS, exchange stabilization dominates and leads to the $^{HS}Fe^{III}$–O_{AF} TS because of Fe–O elongation along the reaction coordinate that stabilizes the antibonding orbital. The Fe–O–H angle in this TS, which corresponds to the σ channel, is 136.2 degrees. This angle is significantly smaller than the 180 degrees observed in model systems and is due to geometric constraints from 1-meA, which cannot move freely in the protein environment. In the $^{IS}Fe^{III}$–O_F TS, which corresponds to

the π channel, the angle is 129.5 degrees. This value is closer to the ideal 120 degrees observed in the model compound pathways. Additionally, based on these results, alternative pathways involving a water in the vacancy between the ferryl and the methyl group acting as a hydrogen shuttle were tested and deemed unfavorable.

To obtain the Helmholtz free energy barrier for the reaction, free energy perturbation (FEP) and inclusion of the QM effects to the free energy[30,33] were carried out after the minimum energy crossing point was obtained (see below). As shown in Fig. 4.2b, these contributions lower the barrier to 18.6 kcal/mol, which is slightly lower than the estimated experimental value of 19.8 kcal/mol. This suggests the hydrogen abstraction should be the rate-limiting step in agreement with previous theoretical studies.[19,21] The intermediate is 6.0 kcal/mol lower than the reactant, indicating that the hydrogen abstraction step is thermodynamically favored, in agreement with model studies.[19,21] The overall free energy surface is similar to the respective potential energy surface, suggesting that there is no significant contribution from the MM fluctuations to the overall reaction. The free energy perturbation surface suggests that the enzyme environment provides the correct geometric orientation for the TS but there is only minimal electrostatic stabilization for this step, which is consistent with previous QM/MM calculations for TauD.[14] This is possibly due to the reference point, which in the present study corresponds to the ferryl intermediate instead of the apo enzyme, since we are only concentrating on the rate-limiting step and not the full catalytic cycle. This point will be investigated further in future studies we have planned for the full catalytic cycle.

4.3.3 Crossing between the two quintet sub-states

As shown in Fig. 4.2a, the minimum energy pathway (MEP) starts from the $^{IS}Fe^{III}$–O_F reactant and continues to the $^{HS}Fe^{III}$–O_{AF} surface for the TS and intermediate structures. A crossing point between these surfaces should exist between the reactant and the TS geometries. By definition, this crossing point corresponds to an intersection where both surfaces have the same energy,

and is sometimes referred to as the minimum energy crossing point (MECP). In order to determine the position of the MECP, we proceeded to generate an initial structure for the optimization on the crossing seam.

It can be seen from Figure 4.3a and 4.3b that the main difference between the $^{IS}Fe^{III}–O_F$ and $^{HS}Fe^{III}–O_{AF}$ reactant structures is the Fe–O distance. After performing a single point calculation on these two structures using both spin states for each, we found that the crossing point likely exists on the reaction coordinate connecting 4.3a and 4.3b. Therefore, structures were generated by varying the Fe–O bond length, while keeping all other QM atoms and the MM environment from structure 4.4a frozen.

Single point energies (without optimization) of this scan at different Fe–O distances are shown in Figure 4.3d. In general, the change of the energy correlates closely with the Fe–O distance. Near the MECP the energy of the $^{HS}Fe^{III}–O_{AF}$ state varies slightly while the $^{IS}Fe^{III}–O_F$ state increases. Near the $^{IS}Fe^{III}–O_F$ reactant, the energy variation for $^{HS}Fe^{III}–O_{AF}$ is much larger than for $^{IS}Fe^{III}–O_F$. These results implicate the vital role of the Fe–O bond length for controlling the relative energies of the iron electron configurations during the activation of the C–H bonds.

When the distance is 1.77 Å, the energy difference between the sub-states is only 0.2 kcal/mol. Using this structure as the initial guess, optimization on the crossing seam results in the MECP structure (Figure 4.3c). The main difference between the MECP and $^{HS}Fe^{III}–O_{AF}$ reactant is the Fe–O distance, which only decreases slightly from 1.81 Å to 1.78 Å. Thus, it is not surprising that the energy difference between them is very small, indicating that the MECP is quite close to the high-spin reactant.

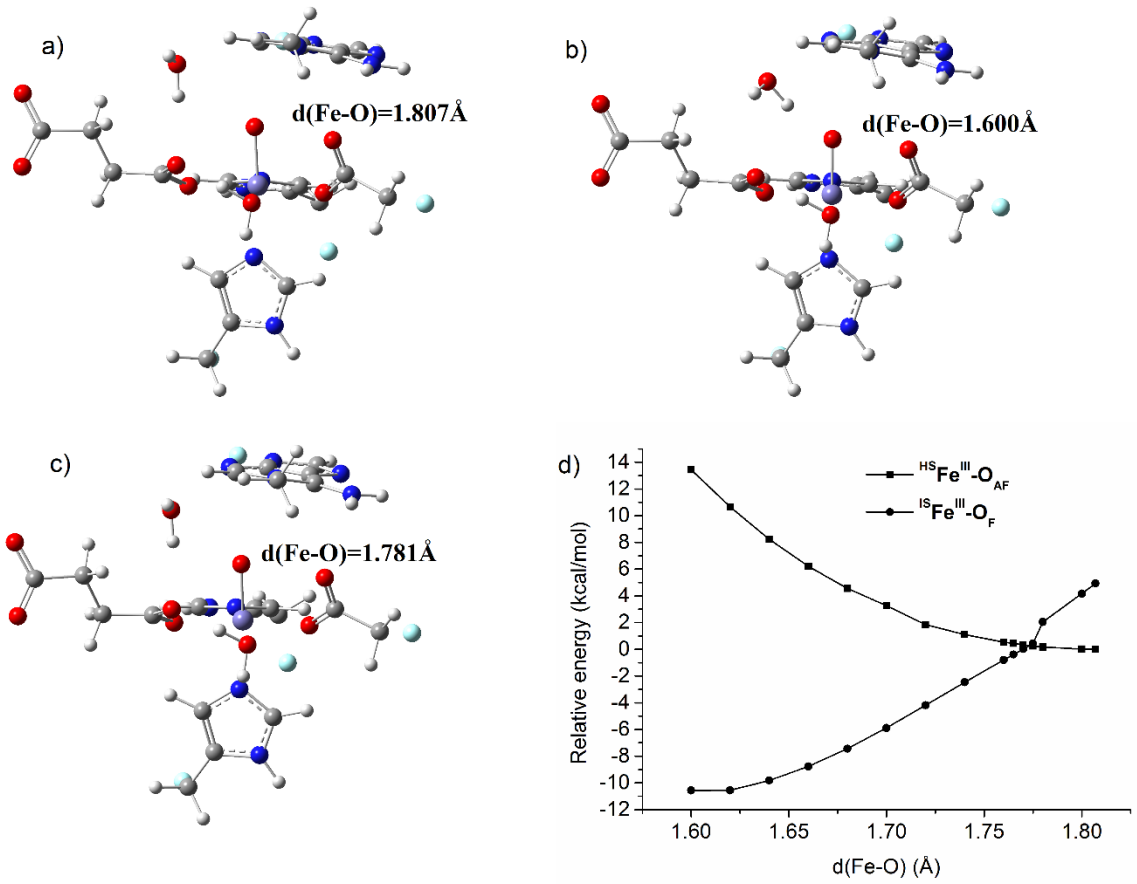


Figure 4.3 (a) The minimum $^{HS}Fe^{III}$–O_{AF} reactant. (b) The minimum $^{IS}Fe^{III}$–O_F reactant. (c) The MECP between the $^{HS}Fe^{III}$–O_{AF} state and the $^{IS}Fe^{III}$–O_F sate. (d) The relative single point energies of the $^{HS}Fe^{III}$–O_{AF} state and $^{IS}Fe^{III}$–O_F at different Fe–O bond lengths (unit: kcal/mol; the absolute energy of the $^{HS}Fe^{III}$–O_{AF} reactant is defined as zero).

4.3.4. Energy decomposition analysis (EDA)

With the optimized structures of the reactant and the TS in hand, we can gain a qualitative understanding of the effects each residue in the MM environment have on the QM subsystem by performing an energy decomposition analysis (EDA) at the residue level (see Figure 4.4a). This analysis is based on the subtraction of the average intermolecular interaction energy (ΔE) between each residue and the QM subsystem at both the reactant and TS. [25,55-59] For residues that have a |ΔE| with magnitude larger than 1 kcal/mol the sign of this energy difference may be used to qualitatively determine the catalytic role of the residue. A positive ΔE indicates that this

residue should increase the reaction rate, while the reaction rate should be decreased with residues that have a negative ΔE. One of us has previously employed this method on other enzymes, and residues predicted from our calculations have been confirmed experimentally subsequently.[60,61] EDA based on the optimized structures reveals nine residues around the active site that have a |ΔE| ≥ 1 kcal/mol. These residues include T51, R73, Y76, K127, Q132, Q155, R161, F185 and R210. The location of these residues in the protein are shown in Figure 4.4b (see Table 4.2 for ΔE values).

Table 4.2 Residues that have more than 1 kcal/mol difference by EDA (all the residues are for Coulomb interaction except that the residue in italic is for Van der Waals interaction).

Residue	**ΔE(kcal/mol)**
T51	-1.0
R73	2.6
Y76	-2.6
K127	-3.0
Q155	1.2
R161	-2.1
F185	-1.2
R210	3.5
Q132	*-1.3*

Experimental mutagenesis results have been reported for only three residues around the active site of AlkB.[62] Experimentally, the mutation of R161 to alanine does not affect the reaction rate, According to our results, R161 provides a stabilizing effect to the TS. This disagreement may be due to the qualitative nature of the EDA method. In particular, R161 is close to the DNA backbone of 1meA, which results in the stabilizing Coulomb interactions in our calculations. Interestingly, although R161 does not directly impact the reaction rate, it does have a big role in the overall reaction. Indeed, this residue is key for the correct discrimination of the damaged base by AlkB.[62]

The other two residues previously studied experimentally correspond to Y76 and T51.[62]. In the case of T51, the mutation of this residue to alanine results in a decrease in the experimental catalytic rate that corresponds to an increase in the reaction barrier of 0.5 kcal/mol. For the Y76 mutation, the result is more dramatic, with a decrease in the experimental rate that corresponds to an increase of 1.0 kcal/mol in the reaction barrier. These results are in good agreement with our EDA results, which suggest that T51 and Y76 help reduce the barrier by -1.0 kcal/mol and -2.6 kcal/mol respectively. Thus, our results not only agree on these residues being important for catalysis, but also in the fact that Y76 has a larger impact on catalysis. Moreover, the experimental results from Ref. 60 do not provide an explanation for how these two mutations affect the reaction rate. Therefore, our calculated results provide a way to fill this gap. In particular, our simulations show that both Y76 and T51 play a role in this reaction step by providing electrostatic stabilization to the TS.

To gain further understanding of the possible conservation of the catalytically important residues obtained by EDA, structural alignments of AlkB with ABH2 [63] and ABH3 [63] have been performed (see Fig. S13 for sequence alignments). As can be seen from panels c and d in Figure 4.4, several residues predicted by the EDA analysis to be important for enzymatic activity are structurally conserved in ABH2 and ABH3 (see Fig. S13 for sequence alignment). In particular, the residues homologous to R210 in AlkB, ABH2-R254 and ABH3-R275, are observed to be in almost the same conformation for all three structures. In addition, the R161 residue in AlkB mentioned above, is also structurally conserved in both ABH2 (corresponding to R203) and ABH3 (corresponding to R218). Thus, these residues are important for the catalytic function of this family of enzymes, and could be interesting targets for experimental mutagenesis studies.

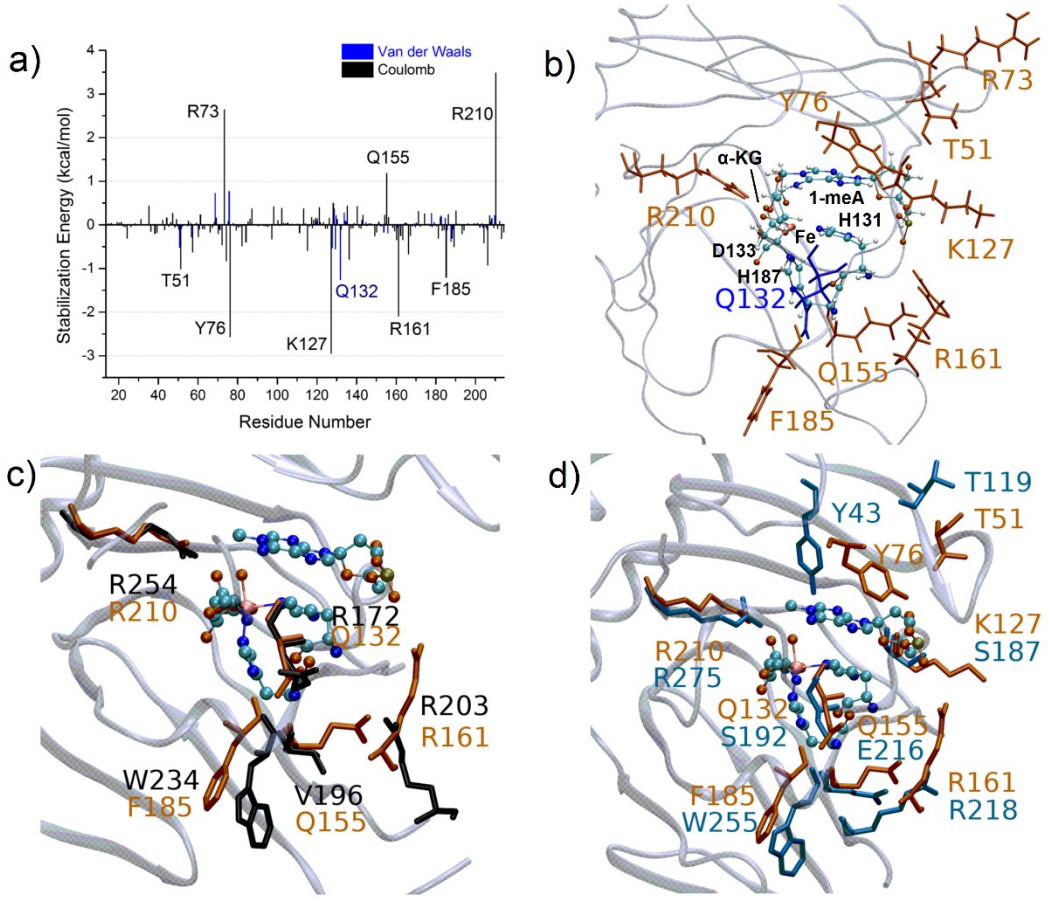


Figure 4.4 (a) EDA results for all the residues. (b) Positions of the residues obtained from EDA (in grey) relative to the active site (ball and stick, black font). (c) Comparison of AlkB (red) and ABH2 (black) showing catalytically important residues that are structurally conserved. (d) Comparison of AlkB (red) and ABH3 (blue) showing catalytically important residues that are structurally conserved.

4.3.5. Electron Localization Function (ELF) and Non-covalent Interaction (NCI) analyses on critical structures and the MECP

ELF[36,37] and NCI[41] can be used to investigate the localization of electrons and the interactions between molecules. We have applied both of these analyses to investigate the QM region of our model using the QM/MM polarized wavefunctions. ELF can indicate the positions of the electrons in the system, e.g., atomic centers, lone pairs, and covalent bonds, thus revealing strong interactions. Conversely, NCI highlights weaker, non-covalent interactions such as hydrogen

bonding and Van der Waals forces. As shown below, the combination of these two analyses allows one to view the evolution of the electron structure elaborated above without the need for complicated orbital analyses.

The changes of the basin populations of the bonds (V(C,H) and V(O,H)) involved with the hydrogen transfer along the pathway are shown in Table 4.3. When the system is closer to the TS, the basin population of C–H for the H atom that will be abstracted declines. At the TS, the basin is shared by the C, H and the oxo O (see Figure 4.5b and Table 4.3), suggesting a three-center two-electron bond between the transferred H atom and the C and O atoms. After the TS, this basin begins to split, and the O–H basin starts to emerge and its population increases. Ultimately, the picture of hydrogen atom transfer is captured without the need for interpretation of an entire orbital manifold. This process is also pictured in the animation in the supplementary information (http://pubs.acs.org/doi/abs/10.1021/ct400130b), which dynamically portrays the ELF and NCI surfaces along the pathway.

According to the results of the ELF population analysis on the ferryl, the total electron population of all the basins of the iron is about 23, and the population of the oxygen basins is about 7, which confirms that the iron indeed withdraws an electron from oxygen. Additionally, the NCI surfaces between the ligands and the iron appear to reflect the coordination number of the metal atom (see Figure 4.5). Given the positions of NCI rings around the metal, the iron is considered to be in an approximate octahedral field.

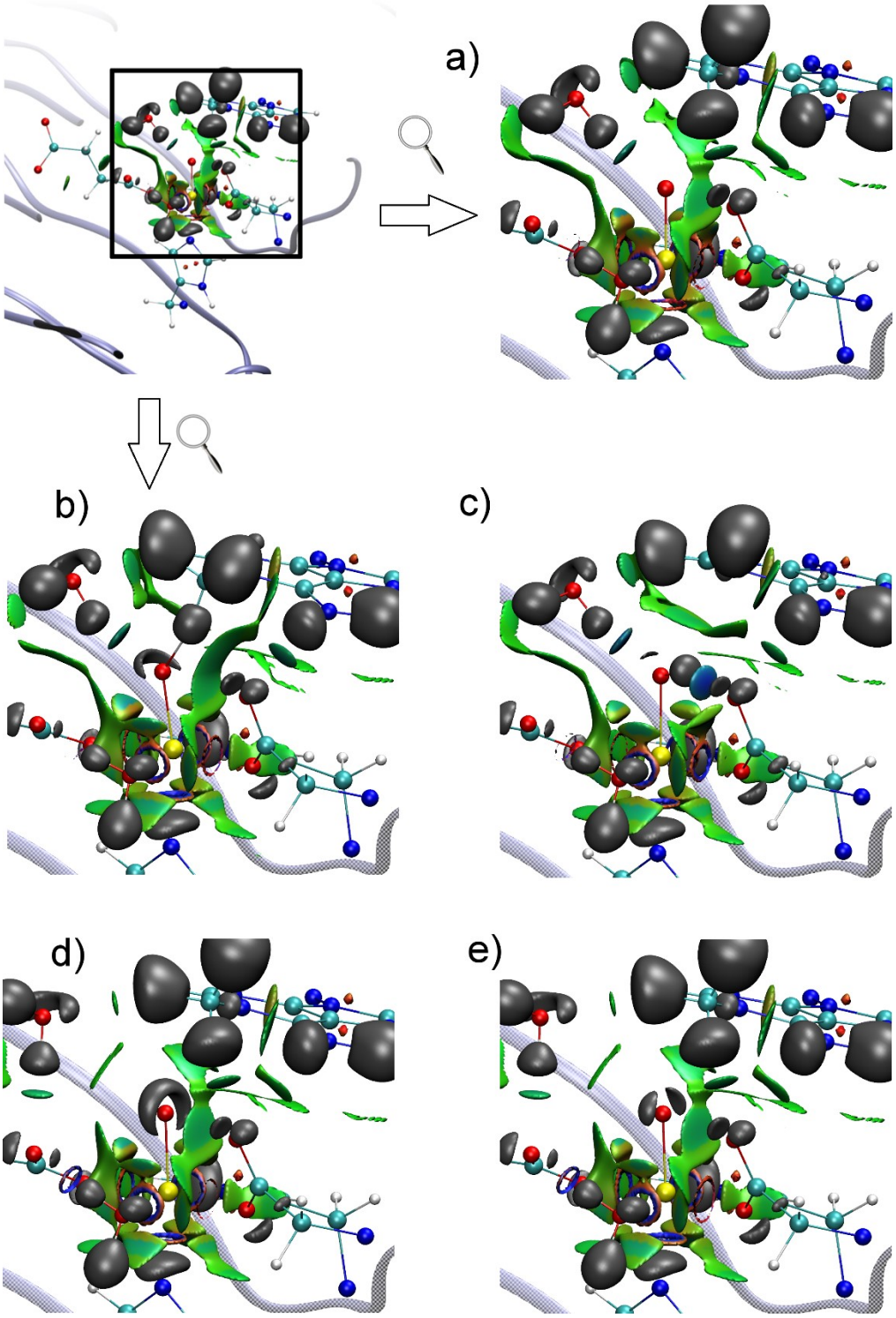


Figure 4.5 ELF and NCI surfaces of the critical structures and the MECP: (a) the $^{IS}Fe^{III}$–O_F reactant; (b) the $^{HS}Fe^{III}$–O_{AF} TS; (c) the $^{HS}Fe^{III}$–O_{AF} intermediate; (d) the $^{HS}Fe^{III}$–O_{AF} MECP; (e) the $^{IS}Fe^{III}$–O_F MECP. (the isovalue for ELF is 0.89; for NCI, the isovalue is 0.5 au, and -0.05 au < sign(λ_2)ρ < 0.05 au; for the NCI surfaces, the color scale is chosen so that blue indicates relatively strong attraction, green indicates relatively weak interaction, and red indicates relatively strong repulsion in the region of non-covalent interactions.)

Table 4.3 The change of the population in V(C,H) and V(O,H) basins along the MEP.

	V(C,H)	V(O,H)
Reactant ($^{IS}Fe^{III}$–O_F)	2.02	0
	2.02	0
	2.02	0
	2.02	0
	2.01	0
	2.02	0
	2.01	0
	2.01	0
	2.00	0
	1.95	0
	1.85	0
TS ($^{HS}Fe^{III}$–O_{AF})		1.73* (V(C,H,O))
	0.90	0.67
	0.70	1.39
	0	1.60
	0	1.67
	0	1.67
	0	1.72
	0	1.72
	0	1.75
Intermediate ($^{HS}Fe^{III}$–O_{AF})	0	1.74

*For TS, the basin is shared by these three atoms.

As shown in Figure 4.5d and 4.5e, NCI and ELF surfaces for both states at the MECP are very similar. However, the ELF basins of the oxygen show a marked difference. For the $^{HS}Fe^{III}$–O_{AF} state, oxygen only has one basin. In contrast, this basin is split into two domains in the $^{IS}Fe^{III}$–O_F

state. When the hydrogen approaches, the interaction between the iron and the oxygen should weaken. This weaker interaction is demonstrated by the NCI surface between the iron and oxygen, which is less blue in the $^{HS}Fe^{III}–O_{AF}$ state than the $^{IS}Fe^{III}–O_{F}$ state. Additionally, after performing a multipolar expansion on the populations of all the basins (see Methods), the first moment of the oxygen is found to be much larger in the $^{HS}Fe^{III}–O_{AF}$ state than in $^{IS}Fe^{III}–O_{F}$ at the MECP. The main contributions to the increment come from the Fe–O–H plane. This suggests that the oxo is highly polarized after the transition in preparation for the hydrogen abstraction.

A recent study using the combined technique of ELF and NCI reveals that NCI can help predict the future reaction steps, and ELF attractors and NCI interaction critical points overlap for organic reactions.[64] We have also found that this technique can be applied to enzyme reactions[65]. As indicated by the supplementary animation (http://pubs.acs.org/doi/full/10.1021/jp403116e), an NCI surface is apparent between the oxo oxygen and the hydrogen being transferred. This surface shows that the attraction increases as the reaction progresses (bluer) before giving rise to the shared ELF basin between C, H and the oxo O. Moreover, ELF basins form at almost the same positions where the NCI surfaces disappear, or NCI surfaces appear where ELF surfaces vanish. As the reaction continues past the TS the basin between the O and H is formed. In addition, a new hydrogen bond is formed between the transferred H and the free oxygen on D133. This additional hydrogen bond is very strong (as judged by the color of the NCI surface) and helps explain the additional stability of the intermediate compared to the reactant.

4.4. Conclusions

We have investigated the rate-determining step in the dealkylation catalyzed by AlkB using QM/MM and DFT calculations. All possible quintet and triplet states have been explored for the first time using ωB97XD. Waters near the reaction center were seen to have a significant effect on the relative energies despite a small geometric impact. A combined ELF and NCI analysis on the QM/MM polarized wavefunction provides similar insights to detailed molecular orbital analyses

and helps understand the evolution of the interactions during the H abstraction without the need for cumbersome, and often subjective, interpretations. We believe these results will encourage more researchers to employ these simple tools to aid in the understanding of complicated electronic structure analysis. The minimum crossing point between the two sub-states on the quintet surface has been obtained, and shown to be near the minimum high-spin reactant. The results provide clarification regarding the electronic structure of Fe^{IV}–oxo and its evolution during the C–H activation. Deviations from the ideal angles of Fe–O–H for the π and σ channels are observed due to geometric hindrance of the enzyme active site. EDA qualitatively predicts nine residues, which may (de)stabilize the TS via Coulomb or Van der Waals interactions. Some of these residues are shown to be structurally conserved in ABH2 and ABH3 and provide new targets for mutagenesis studies in AlkB and the human homologues.

4.5 References

(1) Drabløs, F.; Feyzi, E.; Aas, P. A.; Vaagbø, C. B.; Kavli, B.; Bratlie, M. S.; Peña-Diaz, J.; Otterlei, M.; Slupphaug, G.; Krokan, H. E. *DNA Repair* **2004**, *3*, 1389.

(2) Fu, D.; Calvo, J. A.; Samson, L. D. *Nat Rev Cancer* **2012**, *12*, 104.

(3) Hurley, L. H. *Nat Rev Cancer* **2002**, *2*, 188.

(4) Mishina, Y.; Duguid, E. M.; He, C. *Chem. Rev.* **2006**, *106*, 215.

(5) Sedgwick, B.; Bates, P. A.; Paik, J.; Jacobs, S. C.; Lindahl, T. *DNA Repair* **2007**, *6*, 429.

(6) Falnes, P. O.; Johansen, R. F.; Seeberg, E. *Nature* **2002**, *419*, 178.

(7) Trewick, S. C.; Henshaw, T. F.; Hausinger, R. P.; Lindahl, T.; Sedgwick, B. *Nature* **2002**, *419*, 174.

(8) Yu, B.; Edstrom, W. C.; Benach, J.; Hamuro, Y.; Weber, P. C.; Gibney, B. R.; Hunt, J. F. *Nature* **2006**, *439*, 879.

(9) Grzyska, P. K.; Appelman, E. H.; Hausinger, R. P.; Proshlyakov, D. A. *Proc. Natl. Acad. Sci.* **2010**, *107*, 3982.

(10) Yi, C.; Jia, G.; Hou, G.; Dai, Q.; Zhang, W.; Zheng, G.; Jian, X.; Yang, C.-G.; Cui, Q.; He, C. *Nature* **2010**, *468*, 330.

(11) Eichhorn, E.; van der Ploeg, J. R.; Kertesz, M. A.; Leisinger, T. *Journal of Biological Chemistry* **1997**, *272*, 23031.

(12) Price, J. C.; Barr, E. W.; Tirupati, B.; Bollinger, J. M.; Krebs, C. *Biochemistry* **2003**, *42*, 7497.

(13) Sinnecker, S.; Svensen, N.; Barr, E. W.; Ye, S.; Bollinger, J. M.; Neese, F.; Krebs, C. *J. Am. Chem. Soc.* **2007**, *129*, 6168.

(14) Godfrey, E.; Porro, C. S.; de Visser, S. P. *The Journal of Physical Chemistry A* **2008**, *112*, 2464.

(15) de Visser, S. P. *J. Am. Chem. Soc.* **2006**, *128*, 9813.

(16) Usharani, D.; Janardanan, D.; Shaik, S. *J. Am. Chem. Soc.* **2010**, *133*, 176.

(17) Ye, S.; Price, J. C.; Barr, E. W.; Green, M. T.; Bollinger, J. M.; Krebs, C.; Neese, F. *J. Am. Chem. Soc.* **2010**, *132*, 4739.

(18) de Visser, S. P. *Angew. Chem., Int. Ed.* **2006**, *45*, 1790.

(19) Liu, H.; Llano, J.; Gauld, J. W. *The Journal of Physical Chemistry B* **2009**, *113*, 4887.

(20) Krebs, C.; Price, J. C.; Baldwin, J.; Saleh, L.; Green, M. T.; Bollinger, J. M. *Inorg. Chem.* **2005**, *44*, 742.

(21) Cisneros, G. A. *Interdiscip Sci Comput Life Sci* **2010**, *2*, 70.

(22) Ye, S.; Neese, F. *Proceedings of the National Academy of Sciences* **2011**, *108*, 1228.

(23) Geng, C.; Ye, S.; Neese, F. *Angew. Chem., Int. Ed.* **2010**, *49*, 5717.

(24) Shaik, S.; Chen, H.; Janardanan, D. *Nat Chem* **2011**, *3*, 19.

(25) Cisneros, G. A.; Perera, L.; García-Díaz, M.; Bebenek, K.; Kunkel, T. A.; Pedersen, L. G. *DNA Repair* **2008**, *7*, 1824.

(26) Bollinger, J. M.; Price, J. C.; Hoffart, L. M.; Barr, E. W.; Krebs, C. *Eur. J. Inorg. Chem.* **2005**, *2005*, 4245.

(27) Case, D.; Darden, T. A.; Cheatham, T. E.; Simmerling, C.; Wang, J.; Duke, R.; Luo, R.; Crowley, M.; Walker, R.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P.; University of California, San Francisco.: 2010.

(28) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.; Gaussian, Inc.: Wallingford CT, 2009.

(29) Ponder, J. W.; Washington University in Saint Louis: Saint Louis, MO, 2008.

(30) Zhang, Y.; Liu, H.; Yang, W. *J. Chem. Phys.* **2000**, *112*, 3483.

(31) Parks, J. M.; Hu, H.; Cohen, A. J.; Yang, W. *J. Chem. Phys.* **2008**, *129*, 154106.

(32) Burger, S. K.; Yang, W. *J. Chem. Phys.* **2006**, *124*, 054109.

(33) Cisneros, G. A.; Yang, W. In *Multi-scale Quantum Models for Biocatalysis*; York, D., Lee, T.-S., Eds.; Springer Netherlands: 2009; Vol. 7, p 57.

(34) Manaa, M. R.; Yarkony, D. R. *J. Chem. Phys.* **1993**, *99*, 5251.

(35) Alecu, I. M.; Zheng, J.; Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2010**, *6*, 2872.

(36) Becke, A. D.; Edgecombe, K. E. *J. Chem. Phys.* **1990**, *92*, 5397.

(37) Silvi, B.; Savin, A. *Nature* **1994**, *371*, 683.

(38) Gillespie, R. J.; Robinson, E. A. *Chem. Soc. Rev.* **2005**, *34*, 396.

(39) Noury, S.; Krokidis, X.; Fuster, F.; Silvi, B. *Comput. Chem.* **1999**, *23*, 597.

(40) Pilmé, J.; Piquemal, J.-P. *J. Comput. Chem.* **2008**, *29*, 1440.

(41) Contreras-García, J.; Johnson, E. R.; Keinan, S.; Chaudret, R.; Piquemal, J.-P.; Beratan, D. N.; Yang, W. *J. Chem. Theory Comput.* **2011**, *7*, 625.

(42) Neese, F. *Journal of Physics and Chemistry of Solids* **2004**, *65*, 781.

(43) Reiher, M.; Salomon, O.; Artur Hess, B. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)* **2001**, *107*, 48.

(44) Becke, A. D. *Physical Review A* **1988**, *38*, 3098.

(45) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(46) Chai, J.-D.; Head-Gordon, M. *J. Chem. Phys.* **2008**, *128*, 084106.

(47) Chai, J.-D.; Head-Gordon, M. *Physical Chemistry Chemical Physics* **2008**, *10*, 6615.

(48) Zhao, Y.; Truhlar, D. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)* **2008**, *120*, 215.

(49) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157.

(50) Koivisto, P.; Duncan, T.; Lindahl, T.; Sedgwick, B. *Journal of Biological Chemistry* **2003**, *278*, 44348.

(51) Lonsdale, R.; Harvey, J. N.; Mulholland, A. J. *The Journal of Physical Chemistry Letters* **2010**, *1*, 3232.

(52) Lonsdale, R.; Harvey, J. N.; Mulholland, A. J. *J. Chem. Theory Comput.* **2012**.

(53) England, J.; Martinho, M.; Farquhar, E. R.; Frisch, J. R.; Bominaar, E. L.; Münck, E.; Que, L. *Angew. Chem., Int. Ed.* **2009**, *48*, 3622.

(54) Gupta, R.; Lacy, D. C.; Bominaar, E. L.; Borovik, A. S.; Hendrich, M. P. *J. Am. Chem. Soc.* **2012**, *134*, 9775.

(55) Cisneros, G. A.; Liu, H.; Zhang, Y.; Yang, W. *J. Am. Chem. Soc.* **2003**, *125*, 10384.

(56) Cisneros, G. A.; Wang, M.; Silinski, P.; Fitzgerald, M. C.; Yang, W. *Biochemistry* **2004**, *43*, 6885.

(57) Cisneros, G. A. s.; Perera, L.; Schaaper, R. M.; Pedersen, L. C.; London, R. E.; Pedersen, L. G.; Darden, T. A. *J. Am. Chem. Soc.* **2009**, *131*, 1550.

(58) Li, G.; Cui, Q. *J. Am. Chem. Soc.* **2003**, *125*, 15028.

(59) Senn, H. M.; O'Hagan, D.; Thiel, W. *J. Am. Chem. Soc.* **2005**, *127*, 13643.

(60) Metanis, N.; Keinan, E.; Dawson, P. E. *J. Am. Chem. Soc.* **2005**, *127*, 5862.

(61) Bebenek, K.; Garcia-Diaz, M.; Zhou, R.-Z.; Povirk, L. F.; Kunkel, T. A. *Nucleic Acids Research* **2010**, *38*, 5419.

(62) Holland, P. J.; Hollis, T. *PLoS ONE* **2010**, *5*, e8680.

(63) Duncan, T.; Trewick, S. C.; Koivisto, P.; Bates, P. A.; Lindahl, T.; Sedgwick, B. *Proceedings of the National Academy of Sciences* **2002**, *99*, 16660.

(64) Gillet, N.; Chaudret, R.; Contreras-García, J.; Yang, W.; Silvi, B.; Piquemal, J.-P. *J. Chem. Theory Comput.* **2012**. *8*, 3993.

(65) Fang, D.; Chaudret , R.; Piquemal , J.-P.; Cisneros, G. A. *J. Chem. Theory Comput.* **2013**. *9,* 2156

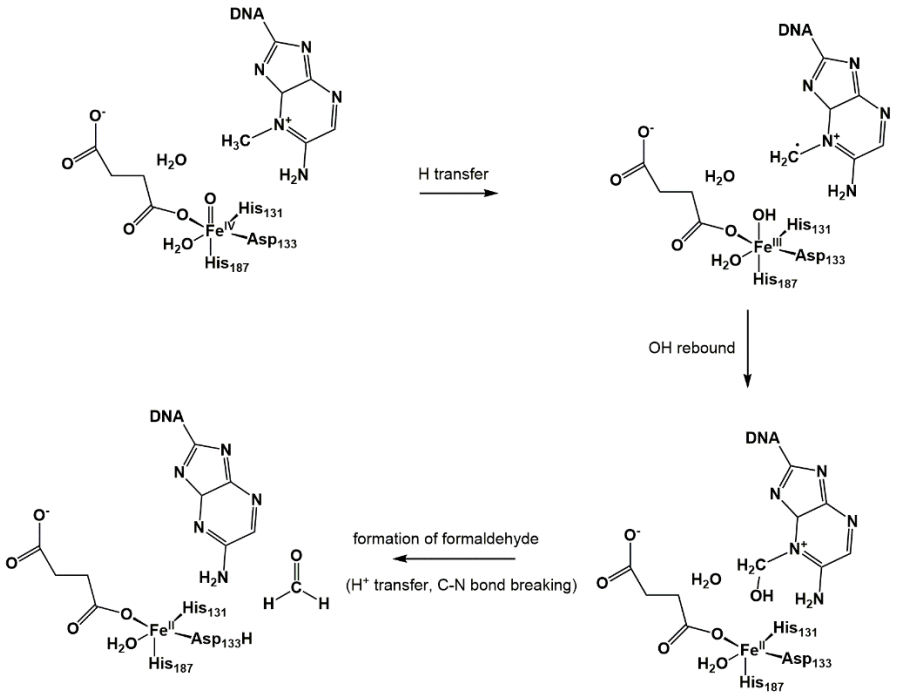
CHAPTER 5

ALTERNATIVE PATHWAY FOR THE REACTION CATALYZED BY DNA DEALKYLASE ALKB FROM *AB INITIO* QM/MM SIMULATIONS

5.1. Introduction

E. coli AlkB is a member of the Fe^{II} and α-KG dependent dioxygenase superfamily of enzymes. AlkB can repair alkylated bases such as 1-meA and N^3-methyl cytosine (3-meC) via an oxidative dealkylation.[1] The proposed mechanism, based on the mechanism of the related enzyme TauD,[2] involves a series of steps that can be separated in two parts. The first part is composed of the formation of an Fe^{IV}=O (ferryl) intermediate along with the release of CO_2 and formation of succinate. After the formation of the iron(IV)-oxo, the oxo moiety undergoes a reorientation from an axial to equatorial position. The subsequent steps comprise the second part, which involve the oxidation of the methyl moiety on the base as shown in Scheme 5.1 (see Scheme 4.1 for the full mechanism).

After the reorientation of the oxo, the Fe^{IV}=O moiety abstracts a hydrogen atom from the methyl group of 1-meA, followed by the OH rebound to the carbon radical. Subsequently, the proton on the recently added OH is transferred and the C-N bond breaks, resulting in the formation of formaldehyde. However, the details of the formation of formaldehyde such as where the proton is transferred and when the C-N bond breaks, are still not clear. Moreover, recent experimental discoveries suggest a possible alternative pathway. The crystal structure of an intermediate in the dealkylation of 3-meC has been recently reported. Based on the crystal data and QM/MM calculations, a zwitterion structure was proposed.[3] In addition, time-resolved Raman spectra reveal the possible existence of a metal-coordinated oxygenated intermediate, such as Fe^{II}-O-C for TauD, another dioxygenase undergoing a similar mechanism to AlkB.[4]



Scheme 5.1 The proposed mechanism for the steps starting from H abstraction in the dealkylation catalyzed by AlkB based on the TauD mechanism.

In the case of 1-meA as substrate, the zwitterion structure and Fe^{II}-O-C complex may form after the deprotonation of the product of the OH rebound process. The crystal structures of AlkB with succinate and different substrates (PDBID: 2FDG, 2FDJ, 3OIS, 3OIU, 3OIT, 3OIV) show a vacancy between the succinate and the aspartate residue (Asp133) bound to iron (see Figure 5.1 for 2FDG[5]). This vacancy is a result of the reorientation of the oxo moiety from an equatorial to an axial position.[6,7] This vacancy can be occupied by a water molecule, which results in the traditional pathway as shown in Scheme 5.2 (H_2O pathway). On the other hand, the physiological pH is slightly basic, and the optimal pH values for repair of 1-meA are 7.5-8. They both indicate the possibility of hydroxide in the environment.[8] Moreover, hydroxide carries a negative charge and is a stronger iron-binding ligand than water. Hence, an alternative pathway (OH^- pathway) with the participation of hydroxide is possible (see Scheme 3).

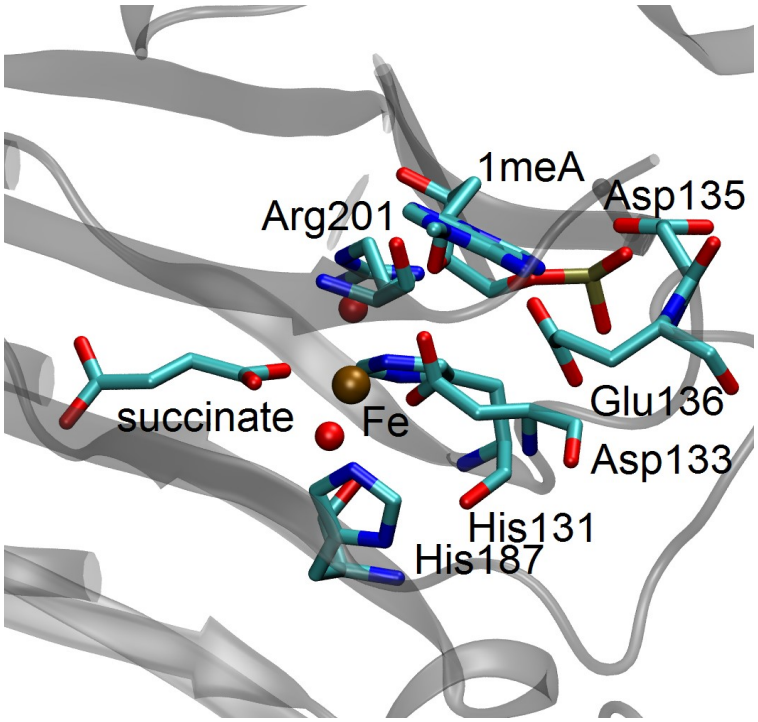
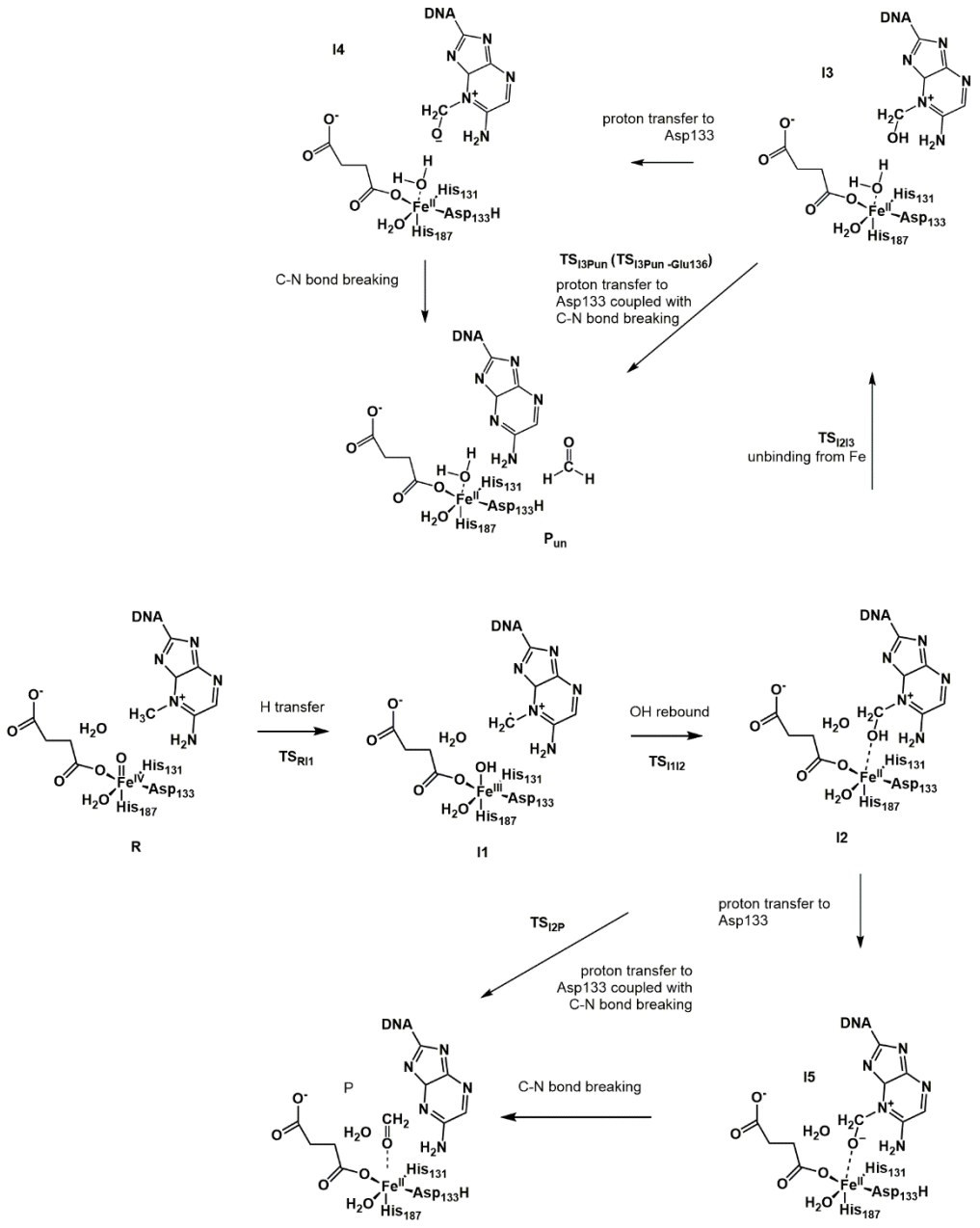


Figure 5.1 The active site of AlkB with 1-meA in the crystal structure (PDBID: 2FDG).

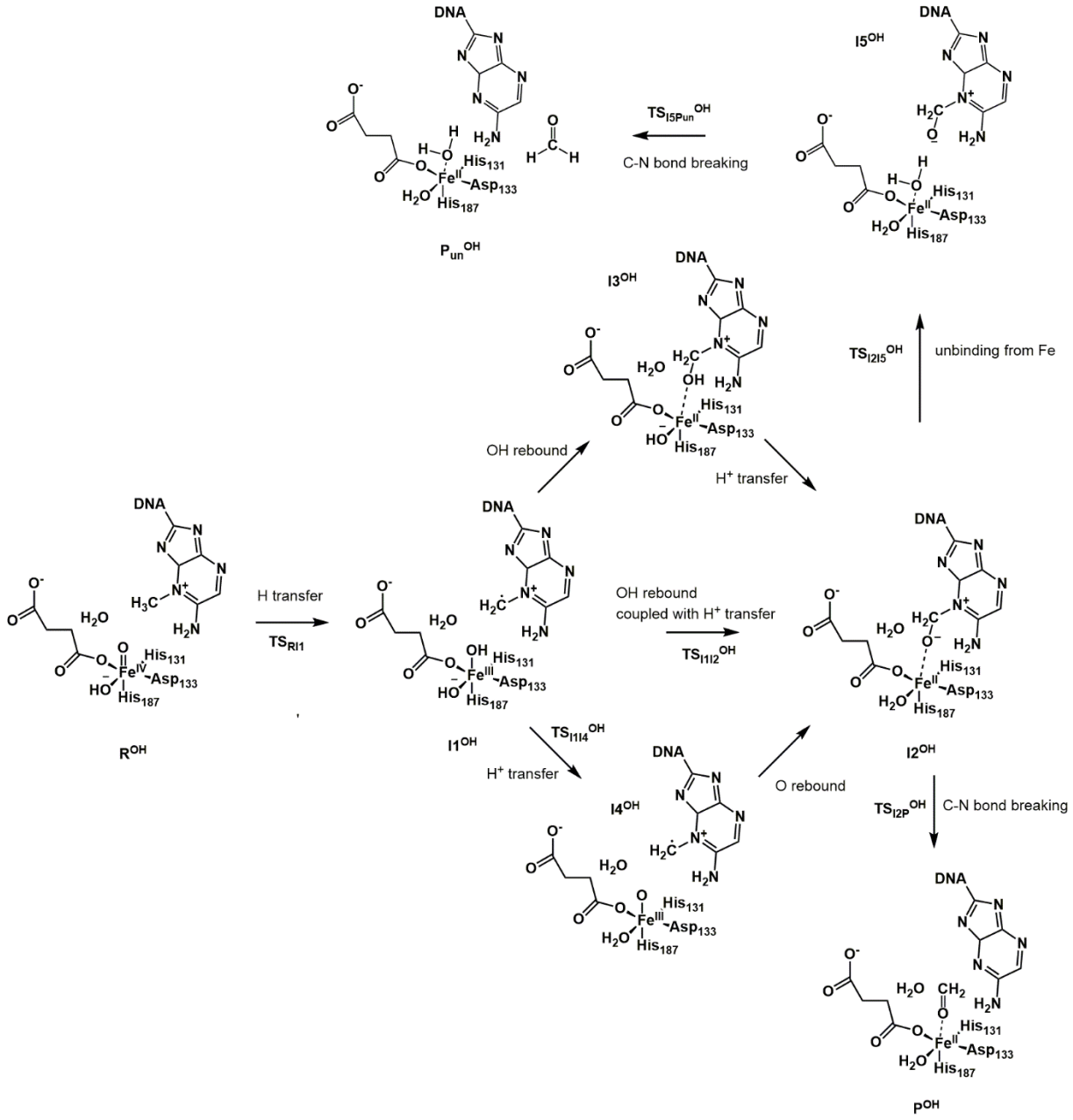
Furthermore, the relative positions of the $–CH_2OH$ (or CH_2O^-) moiety connected to the DNA substrate with respect to the iron atom are not the same for different DNA bases in their corresponding crystal structures. In some crystal structures (3OIU), the substrates are bound to the iron center while for others (3OIS, 3OIV) the oxidized methyls are unbound, and in another case, the moiety is located in an intermediate position (3OIT). In the case of the substrates being unbound to the metal center, the vacancy around the iron left by the substrates may be occupied by a water molecule. This raises the question of whether the unbinding process is always required.

We have previously used *ab initio* QM/MM to study the rate-limiting H atom abstraction step for the H_2O pathway in detail.[9] In this contribution, we elucidate the H_2O pathway (Scheme 5.2) after the hydrogen abstraction step, which includes the OH rebound step and the formation of formaldehyde based on the results from QM/MM simulations. In addition, we report results from QM/MM simulations of the new OH^- pathway (Scheme 5.3) and its comparison with the H_2O pathway to provide new theoretical insights to experimental findings. In section 2 we present the

details for the setup of the systems including the different structures for the different steps and the computational methods. Subsequently, a detailed analysis of the results for the two different pathways explored, H_2O and OH^-, is presented.



Scheme 5.2 The detailed H_2O pathway.



Scheme 5.3 The OH^- pathway: The newly proposed mechanism with OH^- coordinated to the iron (R_{OH})

5.2. Computational Details and Theory

The setup of the system for the simulation and determination of appropriate simulation parameters and levels of theory has been described in detail previously.[9] Briefly after adding hydrogen atoms, water box, and counterions to a crystal structure of AlkB (PDBID: 2FDG[5]), we carried out

Molecular Dynamics (MD) simulations in the NVT (Canonical) ensemble at 300 K using the Amberff99 force field with a 1fs step size. The MD simulations are done with the PMEMD program in AMBER11. The snapshot with the lowest QM/MM energy among the selected ten snapshots was chosen for the further optimization on all reactants, intermediates and products. The QM/MM calculations were performed with an in-house program that links a modified version of Gaussian09[10] with a modified version of TINKER[11]. Electrostatic embedding and an iterative method was adopted for all QM/MM optimizations.[12] The pseudobond approach was used to model the boundary atoms at the QM/MM interface.[13] The TSs were optimized using the QST3 method starting with the structure that has the highest energy obtained from the optimizaed paths calculated with the Quadratic String Method[14]. For OH^- pathway, the water coordinated to the iron was replaced by OH^-. Following the results from our previous simulations, we have employed the ωB97XD[15] functional coupled to the 6-31G(d,p) basis set for the QM part. The structures of reactants, intermediates, and products were confirmed to have no imaginary frequency and all transition states (TS) only have one imaginary frequency corresponding to the vibration along the reaction coordinate connecting the two minima for that step.

To understand the interactions between Fe and its surrounding ligands, Non-covalent interaction (NCI)[16] analysis was performed. Non-covalent interaction (NCI) plots the reduced density gradient versus the product of the sign of the second eigenvalue (λ2) of the electron-density Hessian matrix and the electron density. For the NCI surfaces for visualization, a chosen small reduced electron gradient is used as the isovalue for the NCI surfaces. The types of the interactions can be distinguished by the sign λ2 and represented by the different colors. Positive λ2 means repulsion while negative λ2 means attraction. For the color scale (Blue-Green-Red) we choose, red color represents repulsion and blue color represents attraction. The value of the electron density is represented by the depth of the color. Deeper color means larger electron density, and small electron density is green. Therefore, red color indicates relatively strong repulsion; blue color

represents relatively strong attraction; green color is a sign of relatively weak interaction. This analysis has been proven to be a powerful tool to probe the interactions in small-size molecules[17-19] and large-size systems such as enzymes[20]. The NCI calculations were performed with the NCIPlot program[21].

5.3. Results and Discussion

In this section we present the results and discussion for the steps following H abstraction for the water and hydroxyl pathways. Section 5.3.1 presents the comparison of the hydrogen abstraction step between OH^- and H_2O pathway regarding the energetics and the electronic structure of Fe^{IV}-oxo. The differences between these two pathways on the OH rebound step and the unbinding of the DNA base from the metal and the formation of formaldehyde step will be discussed in section 5.3.2 and 5.3.3, respectively. Finally, in section 5.3.4, we present the experimental findings based on the whole picture of energetics for these two pathways.

5.3.1 Hydrogen abstraction

The oxo O is required to be reoriented from a cis to a trans configuration with respect to the His187 residue (equatorial to axial orientation) after the formation of Fe^{IV}-oxo, and before the abstraction of a hydrogen atom from 1-meA. The calculated barrier for this step when water is bound to the Fe atom has been reported previously to be 7.9-12.0 kcal/mol.[6,7] That is, a significant barrier needs to be overcome for the oxo O reorientation. Conversely, our geometry optimization starting from the structure with the oxo in the cis postion with respect to His187 and a hydroxide near the oxo spontaneously leads to a structure (R^{OH}) with OH^- pushing the oxo to the trans position and the OH^- migrating to an equatorial position. That is, the axial to equatorial transition of the oxo O is barrierless in the presence of OH^-. This suggests that OH^- facilitates the reorientation of the oxo.

Our previous study on the hydrogen abstraction step of the H_2O pathway shows there are two sub-states for each spin state as Fe^{IV}-oxo tends to be Fe^{III}-oxyl where one electron is transferred

from Fe to oxo. For the total quintet spin state, there are two sub-states: The first corresponds to $^{HS}Fe^{III}$-O_{AF}, where s=5/2 (high spin, HS) Fe is antiferromagnetically (AF) coupled with s=-1/2 O. The second state is $^{IS}Fe^{III}$-O_F where s=3/2 (intermediate spin, IS) Fe is ferromagnetically coupled with s=1/2 O. Similarly, for the total triplet spin state, there are also two sub-states: $^{IS}Fe^{III}$-O_{AF} where s=3/2 (intermediate spin, IS) Fe is antiferromagnetically (AF) coupled with s=-1/2 O and $^{LS}Fe^{III}$-O_F where s=1/2 (low spin, LS) Fe is ferromagnetically coupled with s=1/2 O (see Figure S2 for the electronic configuration diagram)[9]. Figure 5.2 depicts the relative energies for the hydrogen abstraction step of the OH^- pathway along with the Mülliken spin populations on selected atoms for each of the critical points on the path (see Figure 4.2a for the energy profile for the H_2O pathway[9]).

Similar to the H_2O pathway, the lowest energy state in the hydroxyl pathway is the $^{IS}Fe^{III}$-O_F sub-state for the reactant R^{OH}. Conversely the $^{HS}Fe^{III}$-O_{AF} sub-state has the lowest energy for the transition state (TS) $TS_{RI1}{}^{OH}$ and intermediate $I1^{OH}$ structures. This indicates an inter-system crossing from $^{IS}Fe^{III}$-O_F R^{OH} to $^{HS}Fe^{III}$-O_{AF} before the $TS_{RI1}{}^{OH}$. As expected for the $^{IS}Fe^{III}$-O_F reactant, the distance between OH^- and Fe (1.78 Å) for the OH^- pathway is shorter than the distance between H_2O and Fe (2.09Å) for the H_2O pathway, and it is a sign of a stronger interaction between OH^- and Fe.

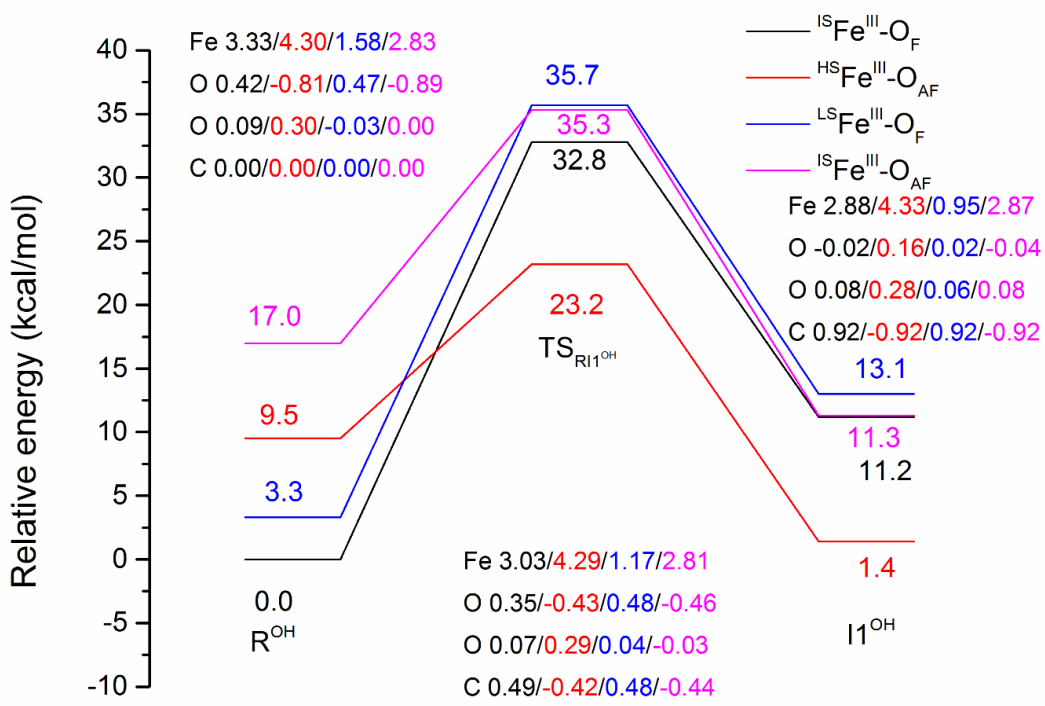


Figure 5.2 Relative energies (in kcal/mol, with $^{IS}Fe^{III}$-O_F as the reference state) of reactant, TS and I1 and Mülliken spin populations of key atoms (Fe, the first O denotes the oxo, the second O denotes the O of OH^- bound to the iron, and C denotes the carbon of methyl group of 1-meA) for the hydrogen abstraction step for OH^- pathway in quintet ($^{IS}Fe^{III}$-O_F and $^{HS}Fe^{III}$-O_{AF}) and triplet states ($^{LS}Fe^{III}$-O_F and $^{IS}Fe^{III}$-O_{AF}).

The calculated energy barrier for the OH^- pathway (23.2 kcal/mol) is 0.8 kcal/mol higher than the barrier calculated for the H_2O pathway (22.4 kcal/mol).[9] The slight energy difference indicates that the H_2O and OH^- pathways are almost equally preferred for the hydrogen abstraction step. The intermediate $I1^{OH}$ is 1.4 kcal/mol higher than the reactant R^{OH} for the OH^- pathway, compared to 3.7 kcal/mol lower than R for the H_2O pathway (Figure 5.3). This result suggests OH^- stabilizes the reactant more than the intermediate, whereas in the presence of water the opposite is observed. In the present studies we have not performed free energy calculations based on the minimum energy paths (MEPs) since the potential energy barrier for the rate limiting step for the

OH^- pathway is very close to the H_2O pathway one. Thus, we expect the associated free energies to show a similar trend as observed before.[9]

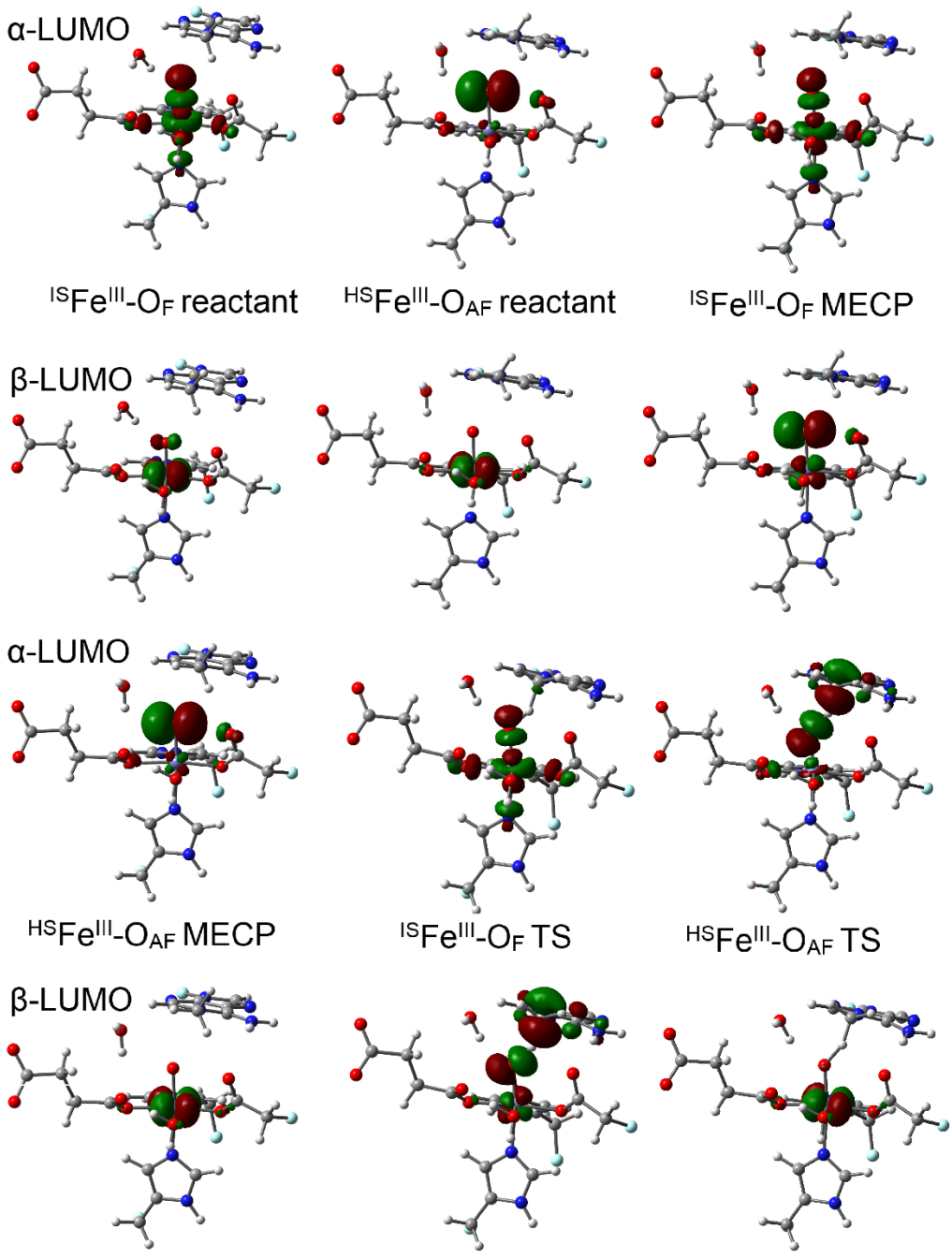


Figure 5.3 α-LUMO and β-LUMO (canonical orbitals, isovalue=0.05 au) of the quintet reactants, MECP and TS structures along the H_2O pathway.

Despite the slight difference in the barrier, the electronic structure of the Fe^{IV}-oxo moiety exhibits different features for the OH^- pathway because of the equatorial OH^- ligand. For the H_2O pathway, the number of unpaired electrons of the iron and oxo O atoms increase during the elongation of the Fe-O distance (d(Fe-O_{oxo})) for both $^{IS}Fe^{III}$-O_F and $^{HS}Fe^{III}$-O_{AF}. For the OH^- pathway the trend is the same for $^{HS}Fe^{III}$-O_{AF} (see Table 5.1). However, for $^{IS}Fe^{III}$-O_F, the elongation of d(Fe-O_{oxo}) causes an increase in the number of unpaired electrons on the iron and a decrease on the oxo, which suggests one electron is transferred back to the oxo from the iron. Note that we modified the d(Fe-O_{oxo}) by only moving the oxo and fixing all other atoms from $^{HS}Fe^{III}$-O_{AF} reactant instead of relaxing the whole complex with frozen d(Fe-O_{oxo})s for both pathways.

Table 5.1 Mülliken spin population for Fe and O (oxo) in $^{HS}Fe^{III}$-O_{AF} and $^{IS}Fe^{III}$-O_F states with different distances between Fe and O (d(Fe-O_{oxo})) (Å) for the OH^- pathway. The numbers in italics correspond to the structure where the $^{HS}Fe^{III}$-O_{AF} and $^{IS}Fe^{III}$-O_F states are close in energy (T_{OH}). Except for this point, all other structures are obtained by only moving O with all other atoms fixed starting from the optimized $^{HS}Fe^{III}$-O_{AF} reactant.

d(Fe-O_{oxo})	Spin population (Fe, $^{HS}Fe^{III}$-O_{AF})	Spin population (O_{oxo}, $^{HS}Fe^{III}$-O_{AF})	Spin population (O_{OH}, $^{HS}Fe^{III}$-O_{AF})	Spin population (Fe, $^{IS}Fe^{III}$-O_F)	Spin population (O_{oxo}, $^{IS}Fe^{III}$-O_F)	Spin population (O_{OH}, $^{HS}Fe^{III}$-O_{AF})	E($^{HS}Fe^{III}$-O_{AF})-E($^{IS}Fe^{III}$-O_F)
1.79	*3.96*	*-0.34*	*0.17*	*4.24*	*-0.72*	*0.26*	*-0.3*
1.75	3.80	-0.13	0.15	4.22	0.68	0.25	0.6
1.70	3.60	0.10	0.13	4.16	-0.63	0.30	7.0
1.68	3.54	0.17	0.12	4.16	-0.62	0.29	9.6
1.65	3.47	0.25	0.11	4.13	-0.59	0.28	13.7
1.63	3.43	0.39	0.12	4.10	-0.55	0.27	16.5

To understand these differences we turn to the analysis of the canonical molecular orbitals for both pathways. During the H atom abstraction, an electron will be transferred to the oxo O. Figure 5.3 shows the α-LUMO (lowest unoccupied molecular orbital) and β-LUMO for the reactant structures for the H_2O pathway. For the $^{IS}Fe^{III}$-O_F, the α-LUMO (σ* orbital) only changes slightly

during the elongation of the d(Fe-oxo) from the reactant (1.60 Å) to the MECP (1.78 Å). In contrast, the percentage of O in β-LUMO (π* orbital) increases and starts to dominate.

According to previous studies[22-35] on the reactivity of Fe^{IV}-oxo, the path for an α electron being transferred from the substrate to the α-LUMO has been referred to as the σ channel, and the path for a β electron being transferred from the substrate to the β-LUMO has been proposed as the π channel. [22-25] For the σ channel, the substrate approaches the Fe^{IV}-oxo in a collinear fashion from the top to maximize the overlap between their orbitals. For the π channel, the substrate is supposed to approach the Fe^{IV}-oxo horizontally. However, the Pauli repulsion between them will make the angle close to 120° instead of 90°.The two channels were proposed to arise from the same $^{IS}Fe^{III}$-O_F reactant and form a Fe^{III}-oxyl radical on the way to the TS [22-25].

As shown in Figure 5.3, the $^{HS}Fe^{III}$-O_{AF} and $^{IS}Fe^{III}$-O_F transition state (TS) structures for the H_2O pathway correspond to the TSs for the proposed σ and π channel respectively. However, the $^{HS}Fe^{III}$-O_{AF} state, which may become the ground state at a relatively long d(Fe-O_{oxo}) was not considered in previous studies as it might be in a relatively high energy level for the complexes. The α-LUMOs of both reactant and MECP in $^{HS}Fe^{III}$-O_{AF} are mostly comprised of oxygen's p orbital with little covalency with the iron's d orbital. Therefore, the Fe(IV)-oxo in $^{HS}Fe^{III}$-O_{AF} can be characterized as an oxyl weakly coupled with Fe(III). The oxyl should be a good α electron acceptor from the substrate. Moreover, the $^{HS}Fe^{III}$-O_{AF} is more stable than $^{IS}Fe^{III}$-O_F at a relatively long d(Fe-oxo) when approaching the TS. However, one disadvantage for $^{HS}Fe^{III}$-O_{AF} is that its α-LUMO is perpendicular to the iron-oxo. As a result, similar to the π channel mentioned above, the substrate tends to approach the iron-oxo horizontally to maximize the orbital overlap but causes large Pauli repulsion, this interplay yields a bent Fe-O-H angle. For the model systems where the substrate can freely move in previous studies[22-25], the σ channel for $^{IS}Fe^{III}$-O_F, which involves less Pauli repulsion is a better choice than the $^{HS}Fe^{III}$-O_{AF} state.

Considering all the factors discussed above, the channel involving $^{HS}Fe^{III}$-O_{AF} may compete with the σ channel arising from the $^{IS}Fe^{III}$-O_F reactant to accept the α electron from the substrate, especially in the enzyme environment, where the $^{HS}Fe^{III}$-O_{AF} state may be stabilized by surrounding ligands or the substrate cannot approach Fe-O collinearly. To access this channel, the complex needs to transit from $^{IS}Fe^{III}$-O_F to $^{HS}Fe^{III}$-O_{AF}. For the H_2O pathway, the crossing between these two states happens via a minimum energy crossing point (MECP).[9] The angles formed by the Fe-O-H atoms for these two quintet states at the TS are around 130°, and the d(Fe-oxo) for the $^{HS}Fe^{III}$-O_{AF} (1.80 Å) reactant is close to the one for the lowest TS ($^{HS}Fe^{III}$-O_{AF}, 1.77 Å) and larger than the value for the $^{IS}Fe^{III}$-O_F reactant (1.60 Å). [9] It can be seen from Figure 5.3 that the LUMOs of $^{HS}Fe^{III}$-O_{AF} resembles those of $^{HS}Fe^{III}$-O_{AF} MECP and the LUMOs of $^{IS}Fe^{III}$-O_F TS resemble those of $^{IS}Fe^{III}$-O_F MECP. The transition from the $^{IS}Fe^{III}$-O_F reactant to a low-lying state at a relatively long d(Fe-O), the description of which matches the features of $^{HS}Fe^{III}$-O_{AF} also occurs and results in the non-collinear Fe-O-C in the $^{HS}Fe^{III}$-O_{AF} TS for some simple model complexes with free substrates and small-size negative charged ligands such as $[Fe^{IV}(O)(F)_5]^{3-}$.[36]

Regarding the two quintet reactants for the OH^- pathway (see Figure 5.4 for their LUMOs), the α-LUMO of the $^{IS}Fe^{III}$-O_F reactant contains a large component from 1-meA, and the orbital of Fe-oxo is not in a good orientation for overlapping with the orbitals of the methyl group of 1-meA. In contrast, the α-LUMO of $^{HS}Fe^{III}$-O_{AF} is still mainly comprised of oxygen's p orbital and should be more likely to accept a α electron from the substrate. It was expected that the elongation of d(Fe-O) would lead to more electrons being transferred from the oxo to the iron for both sub-states and complete the state transition via a MECP like the H_2O pathway. However, as mentioned before, the change of Mülliken spin population shows the elongation of d(Fe-oxo) leads to a β electron being transferred from the iron to the oxo for $^{IS}Fe^{III}$-O_F, and smoothly change into $^{HS}Fe^{III}$-O_{AF}. In practice, the optimization after the structure ($T_{OH,}$ in Table 1 d(Fe-oxo)=1.79 Å) that has a small energy difference for $^{HS}Fe^{III}$-O_{AF} (and $^{IS}Fe^{III}$-O_F) gives the result that both states converge to $^{HS}Fe^{III}$-

O_{AF}. In terms of d(Fe-O), the value is 1.62 Å in $^{IS}Fe^{III}$-O_F reactant and 1.88 Å in $^{HS}Fe^{III}$-O_{AF} reactant and TS, which suggests that the $^{HS}Fe^{III}$-O_{AF} TS resembles the $^{HS}Fe^{III}$-O_{AF} reactant.

Figure 5.4 shows the LUMOs (α and β) of the $^{HS}Fe^{III}$-O_{AF} TS also resemble the respective LUMOs of the $^{HS}Fe^{III}$-O_{AF} reactant. In addition, different from the the $^{IS}Fe^{III}$-O_F reactant where its β-LUMO is mainly comprised of the orbitals for the substrate, the β-LUMO in $^{IS}Fe^{III}$-O_F T_{OH} is mainly a Fe-O π^* orbital, which is more similar to the β-LUMO of the $^{IS}Fe^{III}$-O_F TS. These results underscore the role of the iron-oxo(oxyl) distance, d(Fe-O_{oxo}), in tuning the reactivity of the Fe^{IV}-oxo moiety. In order to obtain the precursor for the $^{IS}Fe^{III}$-O_F TS, the OH^- also needs to be considered to investigate the electronic structure change from the $^{IS}Fe^{III}$-O_F reactant to the $^{IS}Fe^{III}$-O_F TS. Since this pathway is on a higher energy level than the MEP (minimum energy pathway), it will be not discussed here. When studying the hydrogen abstraction by Fe^{IV}-oxo, the $^{HS}Fe^{III}$-O_{AF} state should be paid attention to, as it may become the ground state on the way to the TS, especially for equatorial electron donating ligands or enzymes with constrained substrates.

5.3.2 The OH rebound

Figure 5.5 shows the relative energy for this step for both pathways. The lowest energy states for I1 and $I1^{OH}$ are both $^{HS}Fe^{III}$-O_{AF}, and the two sub-states for the quintet and triplet merge into one after the OH rebound step at I2 and $I2^{OH}$. Therefore, only the quintet state is considered for the calculations of this and subsequent steps. The next step for the H_2O pathway is a typical OH rebound process. The calculated barrier, TS_{I1I2}, is 12.2 kcal/mol, which is in agreement with a previous QM/MM study[6]. For the OH^- pathway, the corresponding intermediate $I3^{OH}$ is not stable. During the optimization, the proton is transferred to the iron-bound OH^- spontaneously, and the zwitterion structure $I2^{OH}$ forms.

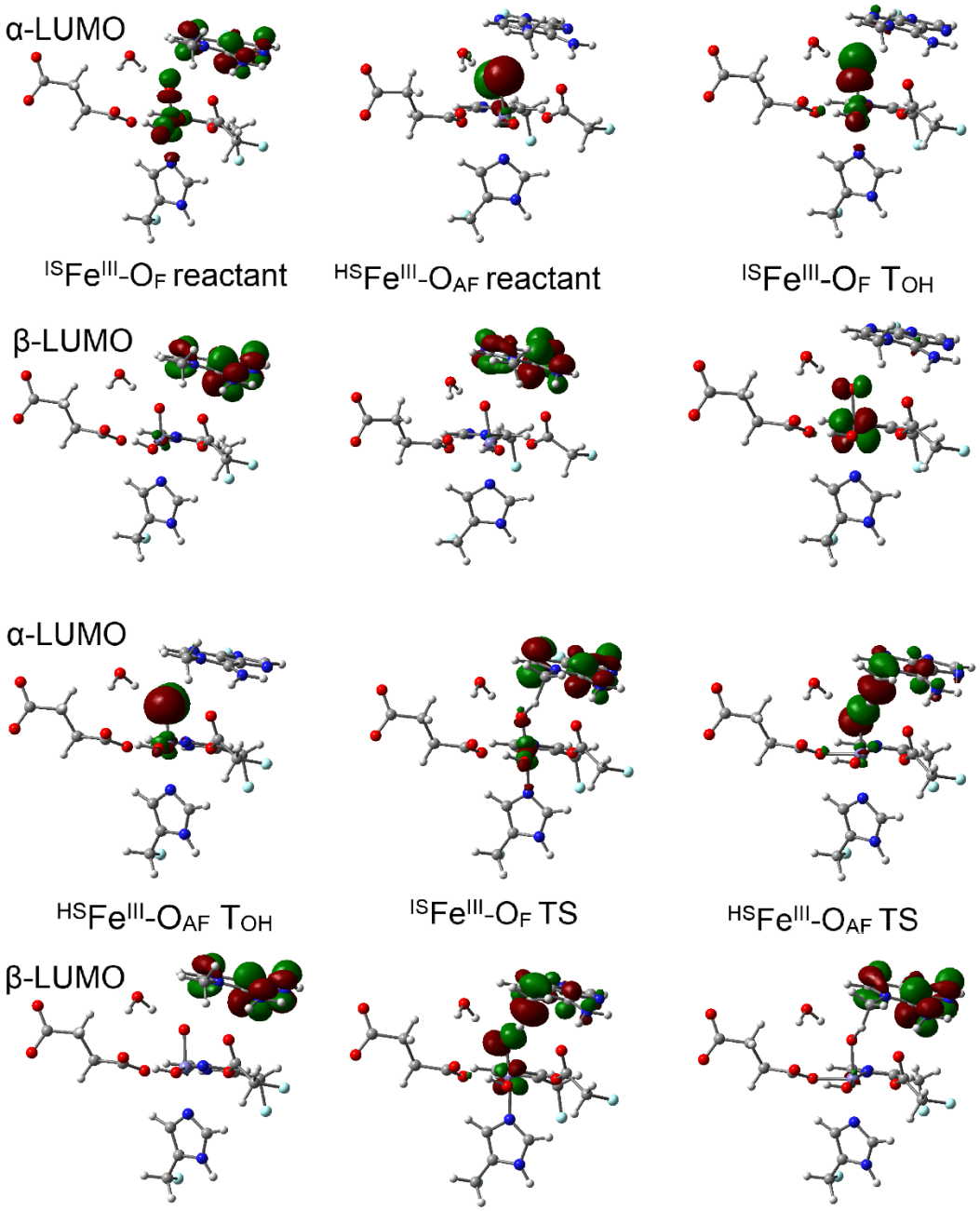


Figure 5.4 α-LUMO and β-LUMO (canonical orbitals, isovalue for the surface 0.05 au) of the quintet reactants, T_{OH} (Table 1) and TS structures along the OH^- pathway.

Two possible pathways from $I1^{OH}$ to $I2^{OH}$ can be proposed. One is a concerted pathway in which the OH rebound process is coupled to the proton transfer via $TS_{I1I2}{}^{OH}$. The other is a stepwise

pathway where the proton is first transferred to form an intermediate $I4^{OH}$, followed by an oxygen transfer. The calculated TS_{I1I2}^{OH} for the concerted pathway is 11.9 kcal/mol. For the stepwise pathway, TS_{I1I4}^{OH} between $I1^{OH}$ and $I4^{OH}$ is 24.7 kcal/mol, and the subsequent oxygen transfer is a barrierless downhill process. Based on these barriers, the concerted pathway is favored over the stepwise pathway. Regarding the zwitterion structure $I2^{OH}$, the distance between Fe and O of $–CH_2O^-$ (on the adenine base) is short (2.05Å). As shown in Figure 5.6, among the ligands coordinated to the iron, the NCI surface between the $–CH_2O^-$ moiety and the Fe atom has the deepest blue color. This means that this group has the strongest attraction to the Fe. The NCI result suggests that the zwitterion structure also exhibits the characteristic of the Fe-O-C bond proposed for the TauD mechanism from time resolved RAMAN studies.[4] Similarly, one may expect a concerted pathway to be preferred over a stepwise pathway as proposed for TauD from the radical intermediate to the Fe-O-C intermediate.[4]

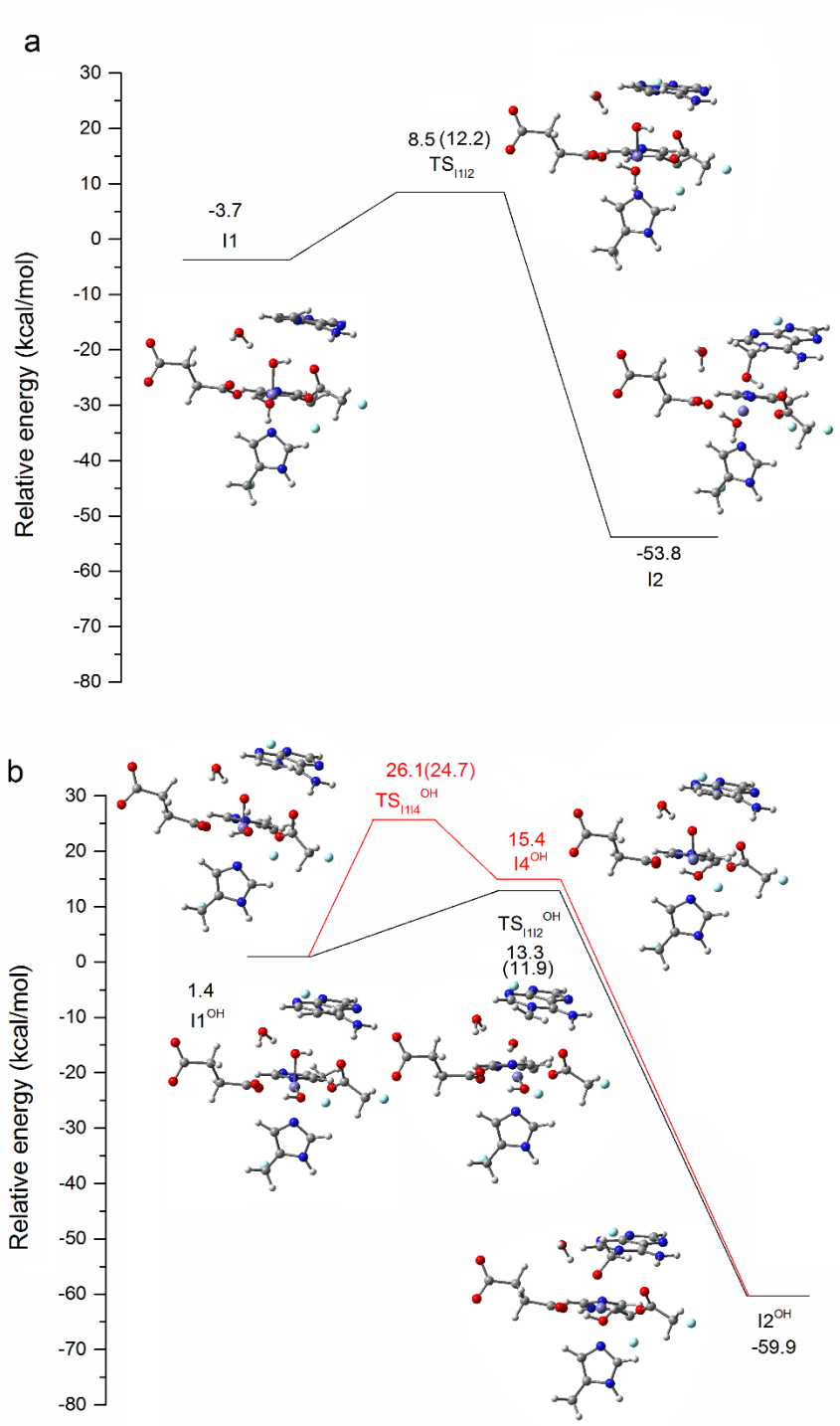


Figure 5.5 Relative energies (kcal/mol) of the structures along the minimum energy path (MEP) for the OH rebound step in the quintet state for H_2O pathway. The numbers in the parentheses are reaction barriers, which are the energy differences between intermediates and their corresponding TSs. (a) and OH^- pathway (b). The energy of the corresponding $^{IS}Fe^{III}$-O_F reactant (Figure 5.2) is taken as zero for each pathway.

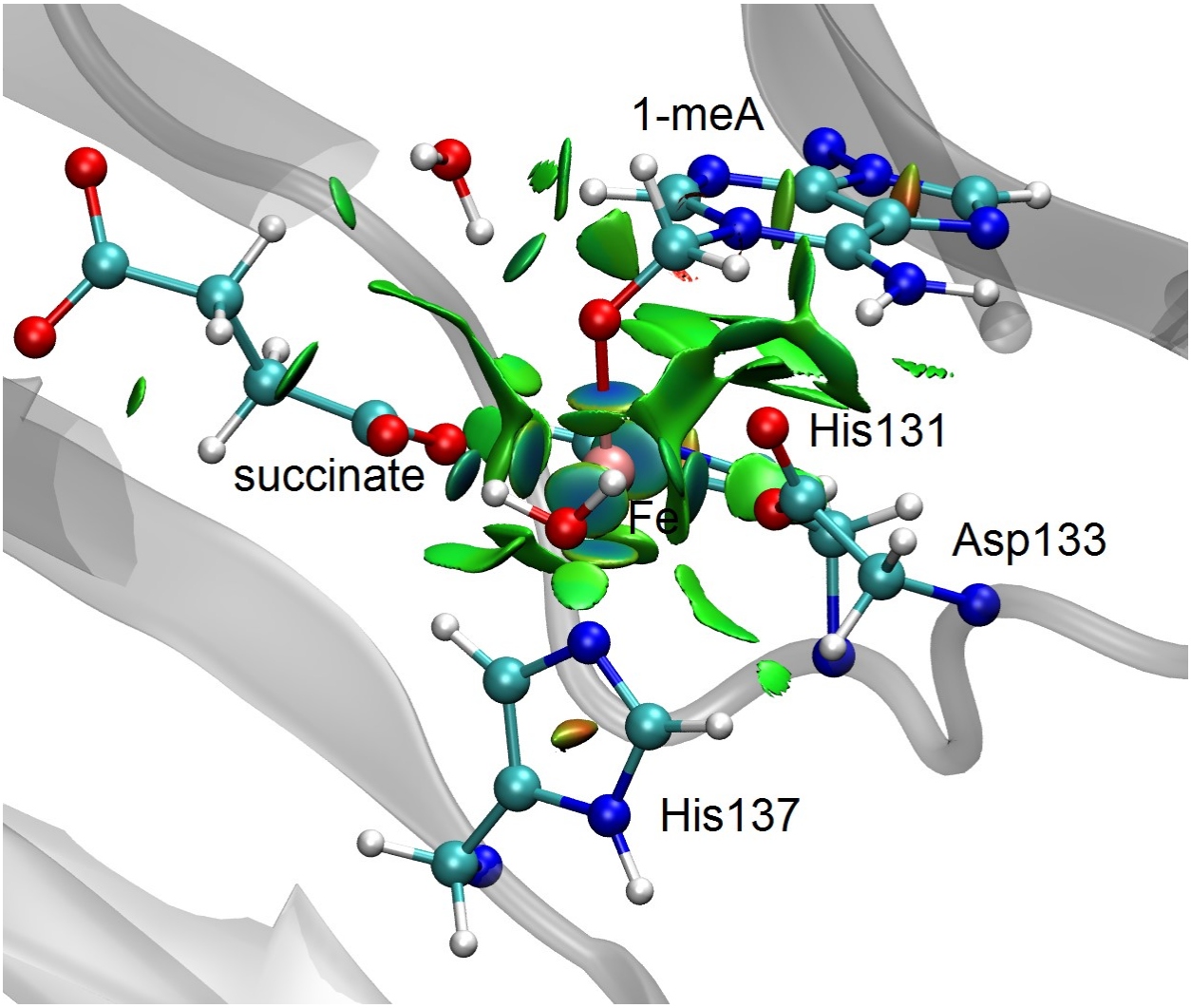


Figure 5.6 NCI surface (isovalue 0.5 au and a color scale -0.1<sign(λ_2)ρ<0.1 au) of the zwitterion structure for the OH^- pathway.

The lowest barrier for the OH^- pathway (TS_{I1I2}^{OH}, 12.0 kcal/mol) is close to the one for the H_2O pathway (TS_{I1I2}, 12.2 kcal/mol). The minor difference indicates these pathways are also equally favored for this step. For both pathways, the rebound step leads to an intermediate with much lower energy. To determine whether a zwitterion structure denoted as I5 also exists for the H_2O pathway, we carried out the optimization starting from a structure in which the proton is transferred from $–CH_3OH$ to Asp_{133}. During the optimization, the proton is spontaneously transferred back to

$–CH_3O^-$, which suggests that the pKa of the iron-bound $Asp_{133}H$ is larger than the iron-bound $–CH_3OH$ when a hydroxyl is coordinated to the Fe atom.

In order to investigate the effect from the positive charge on the nitrogen (N1) of 1-meA, we changed 1-meA to 1-deazameA by replacing N1 with C (see Figure 5.7a) in the optimized zwitterion $I2^{OH}$ (bound to the iron) and $I5^{OH}$ (unbound from the iron) structures, and carried out geometry optimizations. A stable structure for the $I2^{OH}$ analog (Figure 5.7b) for 1-deazameA was obtained on the PES. Conversely, in the case of the $I5^{OH}$ analog for 1-deazameA, the $–CH_2O^-$ group abstracts a proton spontaneously from the neighboring water that is coordinated to the iron (Figure 5.7c). This finding is consistent with previous results for 3-deazameC.[3] Additionally, different from 1-meA, we are able to obtain the $I3^{OH}$ analog (Figure 5.7d) for 1-deazmaeA, although its energy is 8.0 kcal/mol higher than for the $I2^{OH}$ analog. These optimized structures indicate that the proton accepting ability of iron-bound $–CH_2O^-$ group in 1-deazameA is weaker than when this group is not bound to the iron because of the stabilization effects from the metal. The pKa of iron-unbound $–CH_2OH$ in 1-deazameA is larger than the value of H_2O coordinated to the iron, and hence larger than 1meA. Since the proton transfer is a necessary step for the repair process and might become the rate-limiting step as discussed in the next section, the pKa difference of $–CH_2OH$ may partly explain why 1-meA and 3-meC, which bear a positive charge, are preferred by AlkB over 3-meT and 1-meG, which are charge neutral alkylated DNA bases.[3,8]

5.3.3 Unbinding from the iron center, proton transfer and the formation of formaldehyde

To study whether the unbinding from the metal (Figure 5.8; from I2 to I3 for the H_2O pathway; from $I2^{OH}$ to $I5^{OH}$ for the OH^- pathway) is a necessary step, we investigated the formation of the formaldehyde with $–CH_3OH$ (for the H_2O pathway) or $–CH_3O^-$ (for the OH^- pathway) being coordinated or unbound to the iron. In the case of the H_2O pathway with iron-coordinated $–CH_3OH$, the reaction happens in a concerted manner, where the proton transfer to Asp133 and the bond breaking between the C and N1 of 1-meA (C-N bond breaking, I2-P pathway in Scheme 2) leads

to the product with the formaldehyde bound to the iron (P). The calculated barrier for this step is 25.9 kcal/mol, which is higher than the hydrogen abstraction step. This result suggests that this step may become the rate-determining step under certain circumstances, such as if no better proton acceptors are available.

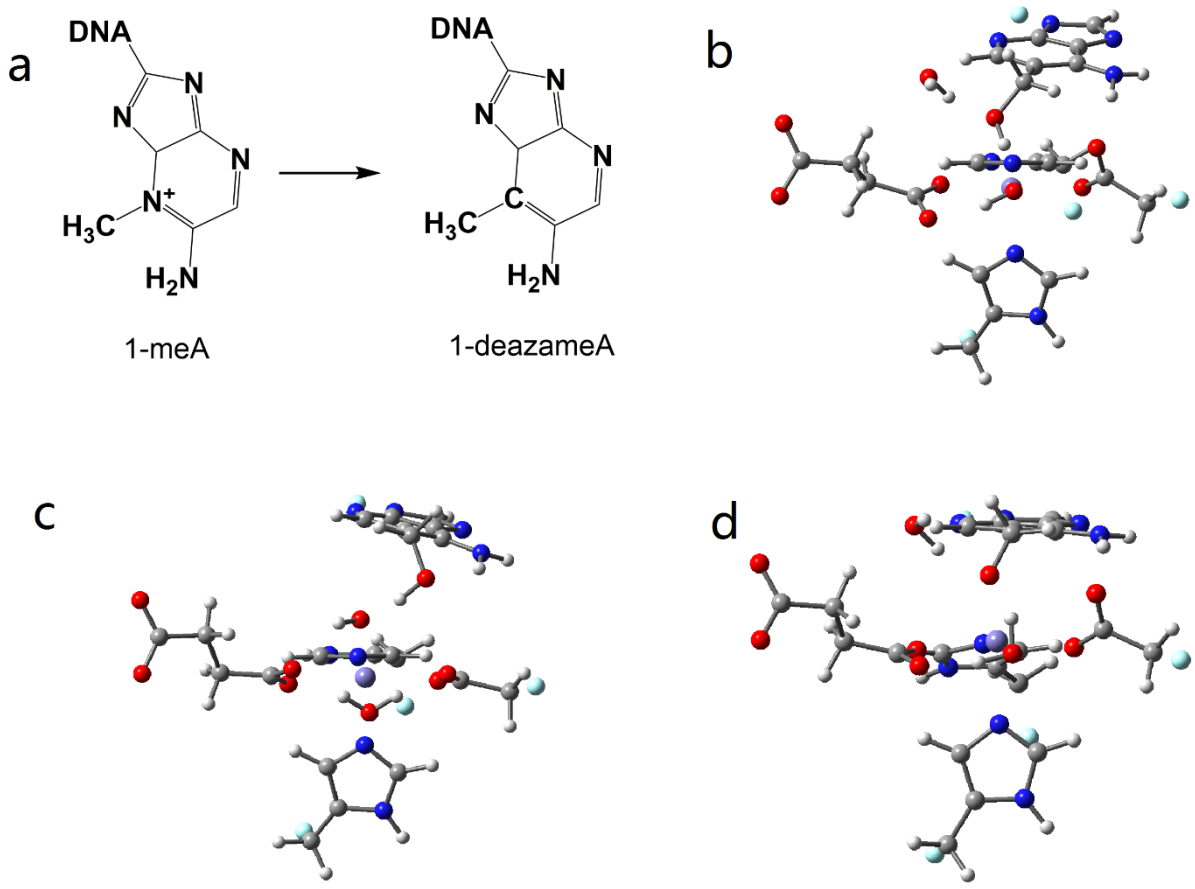


Figure 5.7 (a) the chemical structures of 1-meA and 1-deazameA. (b),(c),(d) are the 1-deazameA-related intermediates.

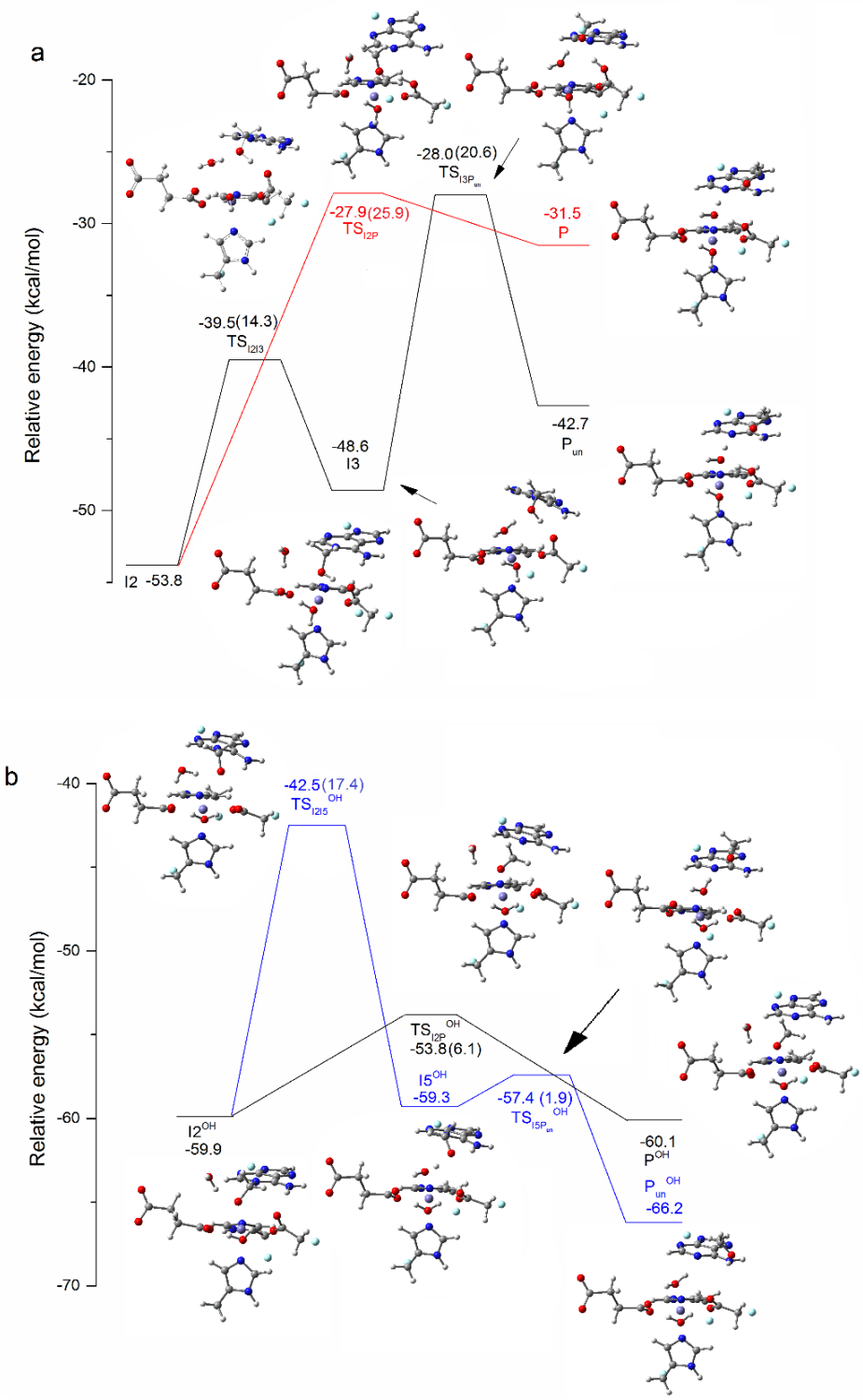


Figure 5.8 Relative energies (in kcal/mol) for the structures along the minimum energy path (MEP) for the detachment of the DNA base from Fe and the formation of formaldehyde in the quintet state for the H_2O pathway. The numbers in the parentheses are reaction barriers, which are the energy differences between intermediates and their corresponding TSs. (a) and OH^- pathway (b). The $^{IS}Fe^{III}$-O_F reactant (Figure 5.2) is taken as the reference for each pathway.

When the $–CH_3OH$ moiety is leaving the iron, I3 forms, and its energy is 5.2 kcal/mol higher than the iron-coordinated intermediate I2. The barrier for the unbinding process is 14.3 kcal/mol. The next step is the proton transfer and the C-N bond breaking. If these two processes are stepwise, a zwitterion intermediate (I4) that is unbound to the iron will form. The proton can then be transferred to Asp133 or some other neighboring residue, such as Glu136, in a direct-transfer pattern or via a water bridge. However, in the case of Asp133, the proposed zwitterion structure cannot be obtained, which is partly due to its weaker ability to accept a proton than for iron-unbound Asp133. In other words, the proton transfer to Asp133 has to be coupled with the C-N bonding and form the final product with the formaldehyde unbound to the iron (P_{un}). The barrier (TS_{I3P}) for this coupled process is 20.6 kcal/mol, and its relative energy is close to the unbound TS_{I2P}.

It is also possible that the proton is transferred to a neighboring residue instead of being transferred to Asp133. As shown in Figure 5.1, the nearest residue to the active site is Glu136, but its distance from the 1-meA suggests the proton transfer would likely have to occur via a water bridge. Before the proton transfer, the structure rearranges from $I3_{Glu136}$ to $I3'_{Glu136}$. (Figure 5.9). In $I3_{Glu136}$, the $-CH_3OH$ moiety forms a hydrogen bond with Asp133 while $-CH_3OH$ forms a hydrogen bond with the bridging water in $I3'_{Glu136}$. To check the existence of a zwitterion structure, we carried out the optimization of $I4_{Glu136}$ assuming the proton transfer from $–CH_3OH$ to Glu136 with a water molecule as the bridge. However, during the optimization, the structure changes back to $I3'_{Glu136}$ with the proton being spontaneously transferred back. This indicates that the pKa of $–CH_3OH$ is also larger than Glu136H. In other words, if the proton is transferred to Glu136, it has to be coupled with the C-N bond breaking. The calculated barrier for $TS_{I3'Pun-Glu136}$ for the concerted pathway from $I3'_{Glu136}$ to $P_{un-Glu136}$ is 7.4 kcal/mol, which is much lower than the barrier for the proton being transferred to Asp133. This suggests that Glu136 may be the final proton acceptor

when no better acceptor available. The proton acceptor role of Glu136 may partly account for the decreased activity of AlkB in repairing 1-meA when Glu136 is mutated to a leucine.[37]

The structure of the $TS_{I3'Pun\text{-}Glu136}$ shows a proton is first transferred from H_2O to Glu136, followed by the resulting OH^- accepting the proton from $–CH_3OH$. It is worth noting that Asp135 could be another possible proton acceptor as well. However, as its relative position to the DNA base in the crystal structure (Figure 5.1) is not conducive for the proton transfer, the proton transfer process may happen when the DNA base is leaving the active site. If a zwitterion structure (I4) indeed forms, the proton from $–CH_3OH$ could be transferred to a hydroxyl molecule in the solvent. In summary, for the H_2O pathway, after the hydrogen abstraction and the OH rebound step, the hydroxyl product first unbinds from the iron and loses a proton to Glu136 or Asp135 or solvent with concerted the C-N bond breaking.

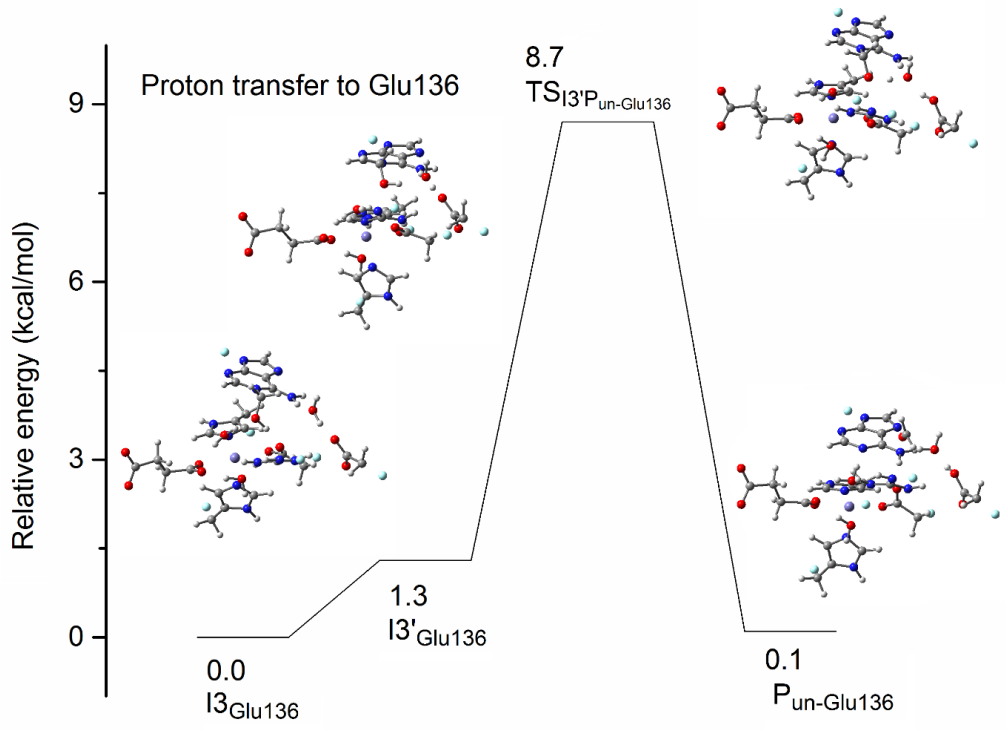


Figure 5.9 Relative energies (kcal/mol) for the structures along the minimum energy path (MEP) for the proton transferred to Glu136 in the quintet state for H_2O pathway. (Arg210, Glu136 and a bridging water were added to the QM subsystem for these structures)

For the OH^- pathway, as the proton has already been transferred to the iron-bound OH^- in the previous rebound step, the final step is only the C-N bond breaking with $–CH_3O^-$ bound or unbound to the iron. The calculated barrier when the methoxide is bound to the iron (TS_{I2P}^{OH}) is 5.9 kcal/mol, compared to the unbound structure (TS_{I2I5}^{OH}) that results in a barrier of 17.3 kcal/mol. This last barrier is 3.0 kcal/mol higher than the barrier for the unbinding process for $–CH_3OH$, which may be due to a stronger attraction between the iron and O of $–CH_3O^-$ than $–CH_3OH$ and the repulsion between the negatively charged O of $–CH_3O^-$ and one O of Asp133 during the unbinding process. The iron-unbound zwitterion intermediate $I5^{OH}$ is only 0.6 kcal/mol higher than the iron-bound one ($I2^{OH}$). The barrier for $I5^{OH}$ being finally dissociated into P_{un}^{OH} (TS_{I5P}^{OH}) is 1.9 kcal/mol. The imaginary frequency vibrational mode corresponds to the rotation of the formed formaldehyde, which suggests the energy for C-N bond breaking should be lower than 1.9 kcal/mol. Therefore, for the OH^- pathway, after the hydrogen abstraction and the OH rebound coupled with a proton transfer to the OH^-, the formed zwitterion structure prefers the C-N bond breaking directly over unbinding from the metal center first.

5.3.4 Comparison between the H2O and OH- pathways

For both the H_2O and OH^- MEPs (see Figure S4 and Figure S5 for the complete energy profile for these two pathways respectively), the rate-limiting step is the hydrogen abstraction step, and their barriers for this step are close to each other. Once the rate limiting step has been achieved, the mechanisms differ significantly. For the H_2O pathway, the last step with the lowest barrier is a proton transfer to Glu136 via a water bridge, and the hydroxyl product has to unbind from the iron center first. Since the water molecule that acts the bridge and Glu136 can move freely, they are not always in a perfect arrangement for the proton transfer. In that case, Asp133 may become the best choice. As the barriers for a proton being transferred to Asp133 coupled with the C-N bond breaking are higher than the hydrogen abstraction step, this step may become the rate-limiting step for the H_2O pathway. In contrast, the barrier for the last step for the OH^- pathway, the

C-N bond breaking leading to the formation of formaldehyde is much lower. It means the OH^- pathway should be preferred over the H_2O pathway, which may partly account for the basic optimal pH for the 1-meA repair catalyzed by AlkB.

The crystal structure for 3-meC (3OIS, Figure S1) is proposed to be a zwitterion structure similar to I4 based on QM/MM calculations.[3] However, those QM/MM calculations cannot rule out the possibility of an alcohol structure. According to the discussion above, if it is indeed a zwitterion in the crystal structure, the proton from the alcohol structure could be transferred to the solvent instead of to Asp133 or Glu136. The hydrogen abstraction, OH rebound and unbinding from the iron follow the H_2O pathway, and then the proton is transferred to OH^- to form a zwitterion structure. Finally, the zwitterion dissociates into the repaired DNA base and formaldehyde.

5.4. Conclusions

In this work, new pathways for the second part of the reaction mechanism, starting from the rate-limiting H atom abstraction, for the dealkylation of 1-meA catalyzed by AlkB have been proposed and investigated by QM/MM simulations based on the new experimental findings. For the hydrogen abstraction and the OH rebound step, the H_2O and OH^- pathways have close barriers and therefore are equally preferred. For the hydrogen abstraction step, different from most of the previous studies on model systems, the $^{HS}Fe^{III}$-O_{AF} state, where the iron (S=5/2) is antiferromagetically coupled with a oxyl (s=-1/2) becomes the ground state when the Fe-oxo distance, d(Fe-O_{oxo}), is long and similar to the $^{HS}Fe^{III}$-O_{AF} TS. This finding highlights the electronic structure change of the Fe^{IV}-oxo moiety under the enzymatic environment and with the binding of the equatorial OH^-. Regarding the OH rebound step, a hydroxyl structure forms for the H_2O pathway. In contrast, for the OH^- pathway, this step is coupled with the proton transfer from $–CH_3OH$ to the OH^- bound to the iron and forms a zwitterion structure bound to the iron, which can be also characterized as an Fe-O-C complex. Following the OH rebound step, the C-N bond between $–CH_3O^-$ and the DNA base can easily break to form the final product while the DNA base

is bound to Fe. In contrast, for the H_2O pathway, the hydroxyl complex needs to unbind from the iron center first and then transfer a proton to the neighboring residue Glu136 via a water bridge or Asp135 or lose it to the solvent. The proton transfer to Glu136 is coupled to the C-N bond breaking. The larger pKa value of $–CH_3OH$ in the hydroxyl intermediate for neutral DNA bases may account for its lower repair efficiency compared to positively charged DNA bases. The lower energy barrier for the last step in the OH^- pathway than the H_2O pathway when the proton has to be transferred to Asp133 in H_2O pathway may partly explain the basic optimal pH for the repair of 1-meA by AlkB.

5.5 References

(1) Zheng, G.; Fu, Y.; He, C. *Chem. Rev.* **2014**, *114*, 4602.

(2) Hausinger, R. P. *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39*, 21.

(3) Yi, C.; Jia, G.; Hou, G.; Dai, Q.; Zhang, W.; Zheng, G.; Jian, X.; Yang, C.-G.; Cui, Q.; He, C. *Nature* **2010**, *468*, 330.

(4) Grzyska, P. K.; Appelman, E. H.; Hausinger, R. P.; Proshlyakov, D. A. *Proc. Natl. Acad. Sci.* **2010**, *107*, 3982.

(5) Yu, B.; Edstrom, W. C.; Benach, J.; Hamuro, Y.; Weber, P. C.; Gibney, B. R.; Hunt, J. F. *Nature* **2006**, *439*, 879.

(6) Quesne, M. G.; Latifi, R.; Gonzalez-Ovalle, L. E.; Kumar, D.; de Visser, S. P. *Chem. Eur. J.* **2014**, *20*, 435.

(7) Liu, H.; Llano, J.; Gauld, J. W. *J. Phys. Chem. B* **2009**, *113*, 4887.

(8) Mishina, Y.; He, C. *J. Inorg. Biochem.* **2006**, *100*, 670.

(9) Fang, D.; Lord, R. L.; Cisneros, G. A. *J. Phys. Chem. B* **2013**, *117*, 6410.

(10) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.;

Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.; Gaussian, Inc.: Wallingford CT, 2009.

(11) Ponder, J. W.; Washington University, Saint Louis: MO, 2008.

(12) Zhang, Y.; Liu, H.; Yang, W. *J. Chem. Phys.* **2000**, *112*, 3483.

(13) Parks, J. M.; Hu, H.; Cohen, A. J.; Yang, W. *J. Chem. Phys.* **2008**, *129*.

(14) Burger, S. K.; Yang, W. *J. Chem. Phys.* **2006**, *124*.

(15) Chai, J.-D.; Head-Gordon, M. *J. Chem. Phys.* **2008**, *128*.

(16) Johnson, E. R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W. *J. Am. Chem. Soc.* **2010**, *132*, 6498.

(17) Fang, D.; Piquemal, J.-P.; Liu, S.; Cisneros, G. A. *Theo. Chem. Acc.* **2014**, *133*, 1.

(18) Gillet, N.; Chaudret, R.; Contreras-García, J.; Yang, W.; Silvi, B.; Piquemal, J.-P. *J. Chem. Theory Comput.* **2012**, *8*, 3993.

(19) Bellow, J. A.; Fang, D.; Kovacevic, N.; Martin, P. D.; Shearer, J.; Cisneros, G. A.; Groysman, S. *Chem. Eur. J.* **2013**, *19*, 12225.

(20) Fang, D.; Chaudret, R.; Piquemal, J.-P.; Cisneros, G. A. *J. Chem. Theory Comput.* **2013**, *9*, 2156.

(21) Contreras-García, J.; Johnson, E. R.; Keinan, S.; Chaudret, R.; Piquemal, J.-P.; Beratan, D. N.; Yang, W. *J. Chem. Theory Comput.* **2011**, *7*, 625.

(22) Shaik, S.; Chen, H.; Janardanan, D. *Nat Chem* **2011**, *3*, 19.

(23) Neidig, M. L.; Decker, A.; Choroba, O. W.; Huang, F.; Kavana, M.; Moran, G. R.; Spencer, J. B.; Solomon, E. I. *Proc. Natl. Acad. Sci.* **2006**, *103*, 12966.

(24) Geng, C.; Ye, S.; Neese, F. *Angew. Chem.* **2010**, *122*, 5853.

(25) Ye, S.; Neese, F. *Proc. Natl. Acad. Sci.* **2011**.

(26) Hirao, H.; Kumar, D.; Thiel, W.; Shaik, S. *J. Am. Chem. Soc.* **2005**, *127*, 13007.

(27) Kumar, D.; Hirao, H.; Que, L.; Shaik, S. *J. Am. Chem. Soc.* **2005**, *127*, 8026.

(28) Hirao, H.; Kumar, D.; Que, L.; Shaik, S. *J. Am. Chem. Soc.* **2006**, *128*, 8590.

(29) Shaik, S.; Hirao, H.; Kumar, D. *Acc. Chem. Res.* **2007**, *40*, 532.

(30) Janardanan, D.; Wang, Y.; Schyman, P.; Que, L.; Shaik, S. *Angew. Chem. Int. Ed.* **2010**, *49*, 3342.

(31) de Visser, S. P. *J. Am. Chem. Soc.* **2006**, *128*, 9813.

(32) Johansson, A. J.; Blomberg, M. R. A.; Siegbahn, P. E. M. *J. Phys. Chem. C* **2007**, *111*, 12397.

(33) Bernasconi, L.; Louwerse, M. J.; Baerends, E. J. *Eur. J. Inorg. Chem.* **2007**, *2007*, 3023.

(34) Wang, Y.; Han, K. *J Biol Inorg Chem* **2010**, *15*, 351.

(35) Decker, A.; Rohde, J.-U.; Klinker, E. J.; Wong, S. D.; Que, L.; Solomon, E. I. *J. Am. Chem. Soc.* **2007**, *129*, 15983.

(36) Sun, X.; Geng, C.; Huo, R.; Ryde, U.; Bu, Y.; Li, J. *J. Phys. Chem. B* **2014**, *118*, 1493.

(37) Zhu, C.; Yi, C. *Angew. Chem. Int. Ed.* **2014**, *53*, 3659.

CHAPTER 6

DFT STERIC BASED ENERGY DECOMPOSITION ANALYSIS OF INTERMOLECULAR INTERACTIONS

Reproduced with permission from *Theo. Chem. Acc.,* **2014,** 133,148

6.1. Introduction

The accurate determination of intermolecular interactions, such as hydrogen bonding, π-π stacking, etc. has attracted a lot of attention due to its importance in determining the physical and chemical properties of molecular systems, especially biomolecules. Various energy decomposition methods have been proposed for the purpose of understanding the physical meanings of these interactions more deeply, which is especially helpful in the development of force fields.[1-20] The energy decomposition methods can be roughly categorized into perturbative and variational ones. Among the perturbative methods, one of the widely used methods is the symmetry-adapted perturbation theory (SAPT)[8], in which the Hamiltonian is divided into Fock operators, the fluctuation potential for each monomer and the interaction potential. Physically meaningful components such as electrostatic, exchange-repulsion, induction, and dispersion can be obtained. Recently, Density Functional Theory (DFT) based SAPT methods have also been developed[12]. In the case of variational methods, several approaches have been proposed. The first such method was the Kitaura-Morokuma (KM) energy decomposition scheme[3]. Subsequently, a variety of new approaches that improve on the KM method have been proposed including the natural energy decomposition (NEDA)[7], the reduced variational space (RVS) analysis[6], the localized molecular orbital energy decomposition analysis (LMO-EDA)[17], the constrained space orbital variation (CSOV)[5,21], the absolutely localized molecular orbitals energy decomposition analysis (ALMO-EDA)[13,20,22], the block-localized wavefunction energy decomposition[10,19] to name a few. Some of these methods can only be applied to Hartree-Fock (HF) wave functions, while

others can be employed with post-HF methods. Some of these approaches have been extended to DFT by using Kohn-Sham orbitals. In addition, DFT based EDA methods have been developed as well, for example, fragment-localized Kohn-Sham orbitals via a single CI procedure[15] and the density-based energy decomposition with variationally determined intermediate state energies[18].

Recently, one of us has (SL) proposed the concept of DFT steric energy [14] in order to quantify the concept of steric effect, which is widely used in chemistry. It has been applied to a variety of systems to study conformational changes[23], the anomeric effect[24] and chemical reactions[25]. In this contribution we present an extension of DFT sterics to develop a novel EDA method, DFTs-EDA. This new DFTs-EDA method is applied to several molecular dimers and the results are compared to results obtained from CSOV decomposition analysis and QTAIM topological analysis[26] to gain a better understanding of the physical meanings of its components. The organization of the paper is as follows: in the following section, we introduce the theory for DFT steric and CSOV decomposition analyses. Subsequently, results for several molecular systems comprising 24 hydrogen-bonded water-containing dimers, DNA base pairs and two types of benzene dimers are presented and discussed, followed by the conclusion section.

6.2. Theory and Methodology

6.2.1 DFT Steric analysis

In DFT, the total energy is expressed as

$$E[\rho(\mathbf{r})]=T_s[\rho(\mathbf{r})]+E_e[\rho(\mathbf{r})]+E_{XC}[\rho(\mathbf{r})] =T_s[\rho(\mathbf{r})]+E_e[\rho(\mathbf{r})]+E_X[\rho(\mathbf{r})]+E_C[\rho(\mathbf{r})], \quad (1)$$

where T_s, E_e and E_{XC} are the noninteracting kinetic, electrostatic, and exchange-correlation energies respectively. The final term can be further separated into exchange (E_X) and correlation (E_C) components.

Recently, Liu has proposed a different partition method of the total energy[14] :

$$E[\rho(\mathbf{r})]=E_s[\rho(\mathbf{r})]+E_e[\rho(\mathbf{r})]+E_q[\rho(\mathbf{r})], \quad (2)$$

where $E_s[\rho(\mathbf{r})]$ is defined as the steric energy, $E_e[\rho(\mathbf{r})]$ is the same Coulomb energy term in equation 1, and $E_q[\rho(\mathbf{r})]$ represents the energy contributions from quantum effects due to the exchange-correlation effects. The first term is equal to the Weizsäcker kinetic energy[27]:

$$E_s[\rho(r)] = T_w[\rho(r)] = \frac{1}{8}\int \frac{|\nabla\rho(r)|^2}{\rho(r)} dr \quad (3)$$

The last term in equation 2 is defined as

$E_q[\rho(\mathbf{r})] = E_{XC}[\rho(\mathbf{r})] + E_{Pauli}[\rho(\mathbf{r})])$ (4),

where $E_{Pauli}[\rho(\mathbf{r})]$ represents the Pauli repulsion:

$E_{Pauli}[\rho(\mathbf{r})] = T_s[\rho(\mathbf{r})] - T_w[\rho(\mathbf{r})]$ (5).

For a bosonic system, the kinetic energy will become $T_w[\rho(\mathbf{r})]$. For fermions such as electrons, after excluding $E_{Pauli}[\rho(\mathbf{r})]$, $E_e[\rho(\mathbf{r})]$ and $E_{XC}[\rho(\mathbf{r})]$, the remaining term, $T_w[\rho(\mathbf{r})]$, should be a measure of steric effects, which is related to the spatial distribution of the electron density.

Applying this scheme to investigate interaction energies leads to following formulas:

$$\Delta E[\rho(\mathbf{r})] = \Delta T_s[\rho(\mathbf{r})] + \Delta E_e[\rho(\mathbf{r})] + \Delta E_{XC}[\rho(\mathbf{r})]$$

$$= \Delta T_s[\rho(\mathbf{r})] + \Delta E_e[\rho(\mathbf{r})] + \Delta E_X[\rho(\mathbf{r})] + \Delta E_C[\rho(\mathbf{r})] \quad (6)$$

and

$$E_{tot}[\rho(\mathbf{r})] = \Delta E[\rho(\mathbf{r})] = \Delta E_s[\rho(\mathbf{r})] + \Delta E_e[\rho(\mathbf{r})] + \Delta E_q[\rho(\mathbf{r})] \quad (7)$$

where for the terms with Δ, their values are obtained by subtracting the sum of the corresponding term for monomers from the one for the molecular complex.

It should be noted that the DFT steric effect is an indication of the size of the system, which is verified quantitatively by the linear relationship between the Fisher information. The Fisher

information is 1/8 of DFT steric effect, and the molecular and atomic volumes.[28] Therefore, a large negative value of the DFT steric effect means the dimer occupies less space than the sum of the monomers. Since no physical observable variable is directly linked to the steric effect, the definition of steric effect is not unique and may be different in some other approaches.[29-31] Thus, the DFT steric effect, which represents spatial effects, is slightly different from the common concept of the steric effect, which also includes some effects related to the Pauli exclusion.

6.2.2 Constrained Space Orbital Variation

The constrained space-orbital variation (CSOV) method [5,32-35] decomposes the intermolecular interaction energy into four components: Coulomb, exchange-repulsion, polarization and charge transfer. The first two, namely Coulomb and exchange-repulsion, are termed the frozen-core contribution since they involve unperturbed orbitals for the interacting monomers. The last two terms correspond to the induction effects allowing selected relaxation of the variational spaces.

$$E_{tot}=E_{FC}+E_{Pol}+E_{CT}=E_{Coul}+E_{Exch\text{-}Rep}+E_{Pol(A)}+E_{Pol(B)}+E_{CT(A\rightarrow B)}+E_{CT(B\rightarrow A)} \quad (8)$$

Following the Kitaura-Morokuma scheme, the frozen-core energy (E_{FC}) is calculated as the antisymmetrized Hartree product of the isolated monomer wave functions. When executing the SCF procedure, the occupied orbitals of monomers A and B remain frozen, which means there is no relaxation into the virtual orbitals. E_{Coul} corresponds to the classic Coulomb interaction, which includes the electron-electron repulsion, the electron-nucleus attraction and the nucleus-nucleus repulsion. For the interactions involving electrons, the unperturbed molecular orbitals of monomers are used. The difference between E_{FC} and E_{Coul} is defined as the exchange–repulsion term, which arises from the Pauli exclusion principle. It is worth noting that the CSOV values for E_{Coul} and $E_{Exch\text{-}Rep}$ are identical to the Coulomb and exchange terms obtained from the Kitaura-Morokuma decomposition if computed at the same level of theory.

The remaining terms, polarization (E_{Pol}) and charge transfer (E_{CT}), depend on the variation of the

molecular orbitals and their eigenvalues in different variational space. For example, $E_{Pol(A)}$ is the energy due to the relaxation of the occupied orbitals of molecule A in the electric field generated by molecule B. When performing the SCF procedure on A, Molecule B's occupied orbitals are kept frozen and its virtual orbitals are not included in the variational calculation. During the procedure, molecule A's orbitals are orthonormal to the frozen MOs of B. A similar procedure is carried out in order to calculate $E_{CT(A\rightarrow B)}$. However, this time the virtual orbitals of B are allowed to relax so that the electrons can be transferred from A to B. Antisymmetrized wave functions are used in the calculations on both polarization and charge-transfer components to eliminate short-range problems in dimers concerning strong polarizing fields, which are known to produce artifacts in the original Kitaura-Morokuma decomposition. For DFT, Kohn-Sham orbitals are used in CSOV calculations by following the procedure above.

6.2.3 Non-Covalent Interaction (NCI) Analysis

NCI[36] has been developed to visualize the non-covalent interaction by plotting the electron density versus the reduced density gradient. The peaks at low electron density characterize the desired non-covalent interaction. By multiplying the density by the sign of the second density Hessian eigenvalue (λ2), different types of interactions (attraction and repulsion) can be distinguished. Hence, it is a powerful tool to study the hydrogen bonding, which is classified as non-covalent attractive interaction. For NCI surfaces, red color is a sign of strong repulsion; blue color indicates strong attraction such as hydrogen bonding; green color means weak interaction.

6.2.4 Computational details

The geometries of the dimers presented in this work were optimized with B3LYP[37,38]/6-31G(d) level using the Gaussian 09 software suite[39]. The monomers were kept rigid in their respective monomer-optimized geometries during the geometry optimization of the dimers. Based on these optimized geometries, Kohn-Sham orbitals were obtained at the B3LYP level with two different basis sets, 6-31G(d) and 6-311++G(2d,2p), for subsequent EDA analysis for hydrogen-bonded

water-containing dimers. All reported corrections to supermolecular energies for basis set superposition error (BSSE) use the counterpoise method[40]. CSOV energy decomposition was performed at the same level of theory (B3LYP/6-31G(d) and 6-311++G(2d,2p)) using a modified version of the HONDO program[41]. Density Functional steric energy decomposition was done by a modified version of NWChem[42]. For the benzene dimers and DNA base pair only the 6-31G(d) basis was considered. The AIMPAC program was used to calculate the critical points for AIM analysis. NCI surfaces were generated by the NCIPLOT program[43] with its default parameters and visualized using VMD[44].

6.3. Results and Discussion

6.3.1. Hydrogen bonding dimers: Water-containing dimers and DNA base pair

Figure 6.1 shows all the 24 hydrogen-bonded water-containing dimers that have been studied in this paper and their NCI surfaces. The dimers include: formamide-water, N-methylformamide-water, methanol-water, fluomethane-water, dichloromethane-water, difluoromethane-water, methane-water, ammonia-water, methylamine-water and water-water. The results for the interactions obtained with the 6-311++G(2d,2p) basis sets are shown in Tables 6.1 and 6.2. As shown in Table S3, it is well known[45] that the BSSE energies play an important role in the intermolecular interaction and this effect is reduced with a larger basis set. After taking the BSSE corrections into account, the difference between the total interaction energies given by the two basis sets is reduced to an average of 1.0 kcal/mol. Comparison of the corresponding components calculated using these two different basis sets (Tables 6.1, 6.2, S1 and S2) shows that all the components are affected by the use of the larger basis set, however, the BSSE energy is largely distributed into the charge transfer term, which agrees with the previous studies.[15,35] Moreover, for different

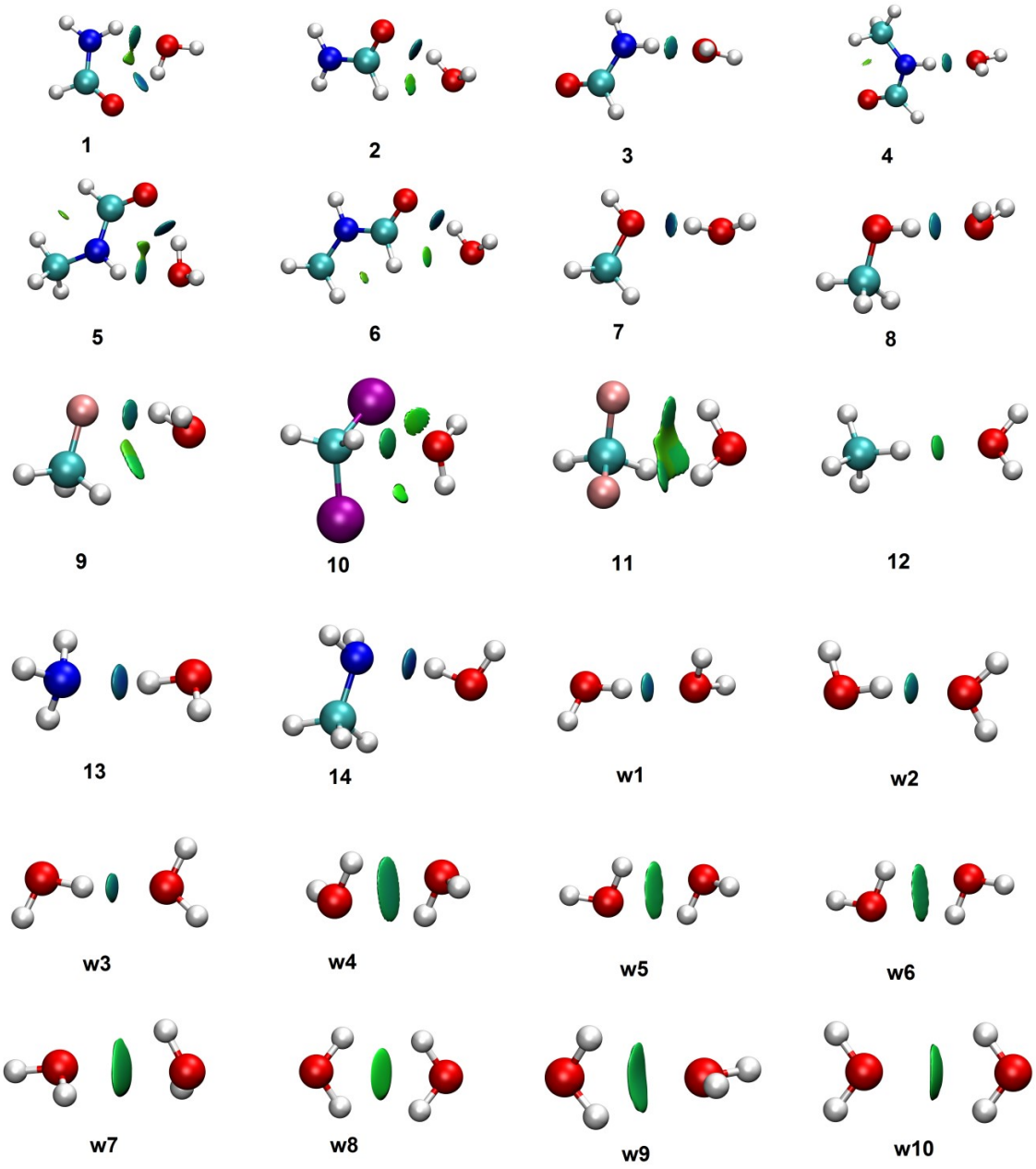


Figure 6.1 NCI plots and structures of 24 dimers (The molecule on the left is A and the right one is B). The isovalue for NCI surfaces is 0.5 au. The surfaces are colored on a blue-green-red scale with -0.04 au <sing(λ2)ρ< 0.04 au. (dimer 1,2,3: formamide-water; dimer 4,5,6: N-methylformamide-water; dimer 7,8: methanol-water; dimer 9: fluomethane-water; dimer 10: dichloromethane-water; dimer 11: difluoromethane-water; dimer 12: methane-water; dimer 13: ammonia-water; dimer 14: methylamine-water; dimer w1-w10: water-water.)

software, different SCF convergence criteria and integration grid points are used, which leads to small deviations between the total energies given by the three programs. The subsequent

discussion will be based on the results obtained with the 6-311++G(2d,2p) basis set for these water-containing dimers.

Table 6.1 shows the results from the CSOV decomposition on the 24 hydrogen-bonded dimers. It can be seen that all the components except the exchange-repulsion term stabilize the dimers. The absolute value of the Coulomb term is the largest among all the negative components, which means it contributes the largest stabilization to the dimers. However, its stabilization effect is largely cancelled out by the exchange-repulsion term, which originates from the Pauli repulsion. As shown in Table 6.1, the polarization energies for monomer A and B are close to each other. In contrast, the charge transfer energy from the hydrogen bond acceptor to the donor is much larger since the lone electron pairs of the acceptor are transferred to the donor. Using dimer 3 as an example, the charge transfer energy from formamide to water is much smaller than the one from water to formamide. This is because the electron of the oxygen atom of water acts as the hydrogen bond donor to formamide. This indicates that the charge transfer helps the formation of hydrogen bonds and stabilizes the dimers. This is consistent with the role of the charge transfer from the lone pair of the proton acceptor to the empty antibonding orbital of the proton donor in lowering the barrier for proton transfer.[46,47] The importance of the charge transfer contribution in intermolecular interactions has also been pointed out in the framework of the CSOV analysis asits weight increases when including dynamic corelation throught DFT.[35]

When scrutinizing the components from CSOV, they are highly related to the chemical properties of the dimers. Generally speaking, the dimers that exhibit more negative total interactions for the dimers show NCI surfaces that are deeper blue (see Figure 6.1). The deeper blue surfaces indicate stronger attractive interactions. For example, dimer 1 (formamide-water) and 5 (N-methylformamide-water), which have two dark blue surfaces are most strongly bound. In contrast, dimer 12 (methane-water), which only possess a small green surface indicating weak interaction, is least bound. All CSOV components for this dimer are small, especially the Coulomb term. This

Table 6.1 CSOV decomposition on water-containing dimers (unit: kcal/mol; B3LYP/6-311++G(2d,2p))

Dimers	E_{FC}	E_{Coul}	$E_{Exch-Rep}$	E_{PolA}	$E_{CT(A\rightarrow B)}$	E_{PolB}	$E_{CT(B\rightarrow A)}$	E_{tot}
1	-2.2	-16.4	14.3	-1.4	-2.3	-1.0	-1.5	-8.3
2	-2.1	-10.1	8.0	-0.9	-2.2	-0.5	-0.3	-5.9
3	-2.2	-8.1	5.8	-0.5	-0.2	-0.5	-1.4	-4.7
4	-2.0	-7.9	5.9	-0.5	-0.2	-0.5	-1.4	-4.5
5	-1.8	-17.4	15.6	-1.5	-2.6	-1.1	-1.5	-8.6
6	-2.1	-10.5	8.4	-0.9	-2.3	-0.5	-0.3	-6.2
7	-1.4	-9.8	8.5	-0.8	-2.2	-0.4	-0.3	-5.0
8	-1.7	-8.9	7.2	-0.5	-0.2	-0.5	-1.8	-4.7
9	-0.3	-6.4	6.2	-0.5	-1.2	-0.2	-0.4	-2.7
10	-0.5	-5.9	5.4	-0.4	-0.2	-0.2	-1.1	-2.5
11	0.3	-6.4	6.6	-0.3	-0.8	-0.2	-0.7	-1.7
12	0.3	-1.1	1.4	-0.2	0.0	0.0	-0.3	-0.2
13	-1.4	-12.4	10.9	-1.0	-2.9	-0.6	-0.2	-6.2
14	-0.7	-13.4	12.7	-1.3	-3.4	-0.7	-0.3	-6.4
w1	-1.9	-8.9	7.0	-0.4	-0.2	-0.5	-1.8	-4.9
w2	-1.7	-7.5	5.7	-0.4	-0.2	-0.5	-1.4	-4.2
w3	-1.8	-7.2	5.4	-0.4	-0.2	-0.5	-1.3	-4.1
w4	-2.1	-7.1	5.0	-0.3	-0.6	-0.3	-0.6	-3.8
w5	-2.0	-6.0	4.0	-0.3	-0.5	-0.2	-0.5	-3.4
w6	-2.0	-5.6	3.6	-0.3	-0.4	-0.2	-0.4	-3.3
w7	-1.8	-5.0	3.2	-0.1	-0.3	-0.2	-0.5	-2.8
w8	-0.5	-1.5	1.0	0.0	-0.2	0.0	-0.2	-1.1
w9	-1.7	-5.3	3.5	-0.2	-0.2	-0.1	-0.4	-2.7
w10	-1.2	-3.0	1.8	-0.1	-0.2	-0.1	-0.1	-1.8

is likely because methane is a nonpolar molecule. Comparing dimer 13 (ammonia-water) and 14 (methylamine-water), the addition of the methyl group, an electron donating group, increases the charge transfer from the nitrogen to the hydrogen of the water, and therefore makes the dimer more strongly bound. This type of stabilization effects from the methyl group can also be found by comparing dimer 7 (methanol-water) and the water dimers. However, if an electron donating group is added to the hydrogen bond acceptor, it will decrease the charge transfer ability, which can be verified by the charge transfer energy from methanol to water for dimer 7 is more negative

than the one from water to methanol for dimer 8(methanol-water). In addition to the relatively low attractions between the monomers, the repulsions between the oxygen of the water molecule and the carbon atom of the other monomer also lead to relatively high $E_{tot.}$ dimer 9 (fluomethane-water),10 (dichloromethane-water) and 11(difluoromethane-water) are the examples.

The DFTs-EDA results in Table 6.2 show that the electrostatic (E_e) and steric (E_s) interactions for the dimers are lower than the sum of these components from its two composite monomers. The stabilization effect is dominated by the change of E_s. In contrast, E_q, which includes E_{xc} and E_{Pauli}, increases when forming the dimers. As the value of ΔE_{xc} is negative, ΔE_{Pauli} contributes mostly to the destabilization effect. Regarding the traditional partitions, ΔT_s, which also contains ΔE_{Pauli}, is the only component destabilizing the dimers. The formation of dimers is favored by the steric, electrostatic, as well as correlation interactions while it is not favored by the quantum effects, which mainly come from the Pauli Exclusion Principle. This finding agrees well with the CSOV analysis with regarding to the destabilization effects arising from the Pauli repulsion. For the traditional DFT partition, ΔE_e has the largest stabilization effect. For the DFT steric analysis, the stabilization effect is dominated by ΔE_s. The plots of the total interaction energy versus E_{Coul}, ΔE_e and ΔE_s respectively are shown in Figure 6.2. Linear relationships can be found between these dominant components for each partition method with total interaction energy, andthe linear relationships between the DFTs-EDA components and the toal energy are conserved for water clusters.[48] Out of these three, E_{Coul} from CSOV and ΔE_s from DFT steric analysis have better linear relationships (R^2>0.93) with $E_{tot.}$ The former suggests the importance of the stabilization effects from the electrostatic interaction in forming hydrogen bonds, and the latter indicates the critical role of the spatial rearrangement of the electron in the formation of these hydrogen-bonded dimers, which is represented by ΔE_s.

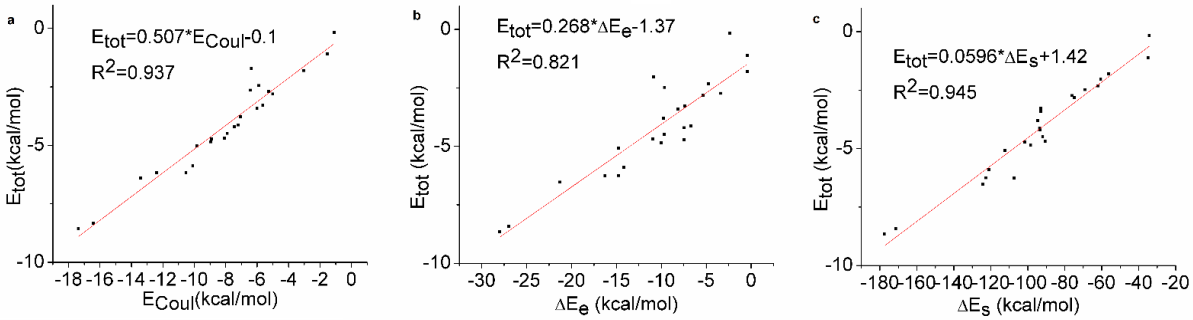


Figure 6.2 Linear fitting of E_e versus E_{Coul} (a), ΔE_e (b) and ΔE_s(c).

Table 6.2 DFT Steric Analysis on water-containing dimers (unit: kcal/mol; B3LYP/6-311++G(2d,2p))

Dimers	ΔTs	ΔEe	ΔExc	ΔEx	ΔEc	ΔEs	ΔE_{Pauli}	ΔEq	Etot
1	36.1	-27.0	-17.5	-12.9	-4.6	-171.4	207.5	190.0	-8.4
2	19.6	-14.1	-11.3	-8.1	-3.2	-120.9	140.5	129.2	-5.9
3	14.6	-11.0	-8.3	-6.0	-2.3	-90.6	105.2	96.9	-4.7
4	13.0	-9.7	-7.9	-5.6	-2.3	-91.9	104.9	97.1	-4.5
5	37.2	-28.0	-17.9	-13.2	-4.7	-177.6	214.8	196.9	-8.7
6	19.9	-14.8	-11.4	-8.2	-3.2	-122.5	142.4	131.0	-6.3
7	19.8	-14.7	-10.1	-7.3	-2.8	-112.2	132.0	121.9	-5.1
8	10.8	-7.4	-8.1	-5.6	-2.5	-101.5	112.3	104.2	-4.7
9	7.2	-4.8	-4.7	-3.3	-1.5	-62.2	69.4	64.6	-2.3
10	12.6	-9.6	-5.4	-3.6	-1.8	-69.0	81.6	76.1	-2.5
11	15.3	-10.8	-6.4	-4.8	-1.6	-60.6	75.9	69.4	-2.0
12	4.1	-2.4	-1.9	-0.9	-1.0	-34.3	38.4	36.4	-0.2
13	21.2	-16.3	-11.2	-8.3	-2.9	-107.2	128.4	117.2	-6.3
14	27.4	-21.3	-12.7	-9.3	-3.4	-124.3	151.7	139.1	-6.5
w1	13.6	-10.0	-8.5	-6.0	-2.4	-98.5	112.1	103.6	-4.9
w2	10.9	-7.4	-7.6	-5.4	-2.3	-93.4	104.3	96.6	-4.2
w3	10.2	-6.7	-7.6	-5.3	-2.2	-93.5	103.7	96.1	-4.1
w4	12.7	-9.8	-6.8	-4.3	-2.5	-94.6	107.3	100.6	-3.8
w5	11.1	-8.1	-6.4	-4.0	-2.4	-93.1	104.2	97.8	-3.4
w6	10.4	-7.4	-6.3	-4.0	-2.4	-93.0	103.4	97.1	-3.3
w7	6.9	-5.3	-4.4	-2.5	-2.0	-74.7	81.6	77.2	-2.8
w8	0.5	-0.4	-1.2	-0.3	-0.9	-35.0	35.5	34.3	-1.1
w9	5.3	-3.4	-4.7	-2.8	-1.9	-76.0	81.3	76.7	-2.7
w10	1.9	-0.4	-3.2	-1.9	-1.3	-56.3	58.2	54.9	-1.8

In addition, previous studies have shown that there is a direct relationship between the electron density at the intermolecular bond critical points and the total intermolecular interaction [49]. We performed Bader's QTAIM analysis on the ten water dimers to determine a possible relationship between the electron density at the intermolecular bond critical point (BCPs) and ΔE_s. We have considered only the water dimers in this case since they only have one inter-molecular BCP. To confirm our results, a good linear relationship between the intermolecular BCP's electron density (ρ_{BCP}) and the total interaction energy can be seen in Figure 6.3a. The relationship can be explained by that stronger hydrogen bonding dimers have more charge-transfer or (and) polarization and hence more electron mixing. As a result, the electron density at the critical point is larger. In addition, the definition of the DFT steric energy (equation 3) tells us its value is determined only by the electron density and its gradient. Is there a relationship between ΔE_s and the electron density at intermolecular BCPs where the electron density gradient is zero? The plot in Figure 6.3b indicates there is indeed a good linear relationship (R^2=0.904) between ΔE_s and 1/ ρ_{BCP}. The slope is positive, which means the dimer with a large ρ_{BCP} has smaller stabilization from ΔE_s. As monomers approach each other to form a dimer, the electron density and its gradient in space, especially around the area where the hydrogen bond forms, will be rearranged accordingly. As an example, Figure 6.4 compares the plots of the electron gradient versus the electron density for the water monomer (Figure 6.4a) and the canonical water dimer (w1, Figure 6.4b). It can be qualitatively seen that the electron density change mostly comes from the peak that belongs to the hydrogen bonding. This plot is similar to the plot between the reduced density gradient and the product of the sign of λ2 and the electron density obtained by NCI [36,43]. As mentioned before, stronger hydrogen bonding dimers have more mixing. The more relaxed the electron distribution is, the gradient becomes smaller and the electron density around the critical points is larger. According to equation 3, both make the steric energy smaller. Therefore, stronger hydrogen bonding dimers have lower steric energy. Similarly, for the water dimers, the change of DFT steric energy for forming dimers may be also mainly due to the interactions arising from the hydrogen

bonding region, which helps explain why ΔE_s correlates with the electron density at the intermolecular BCPs.

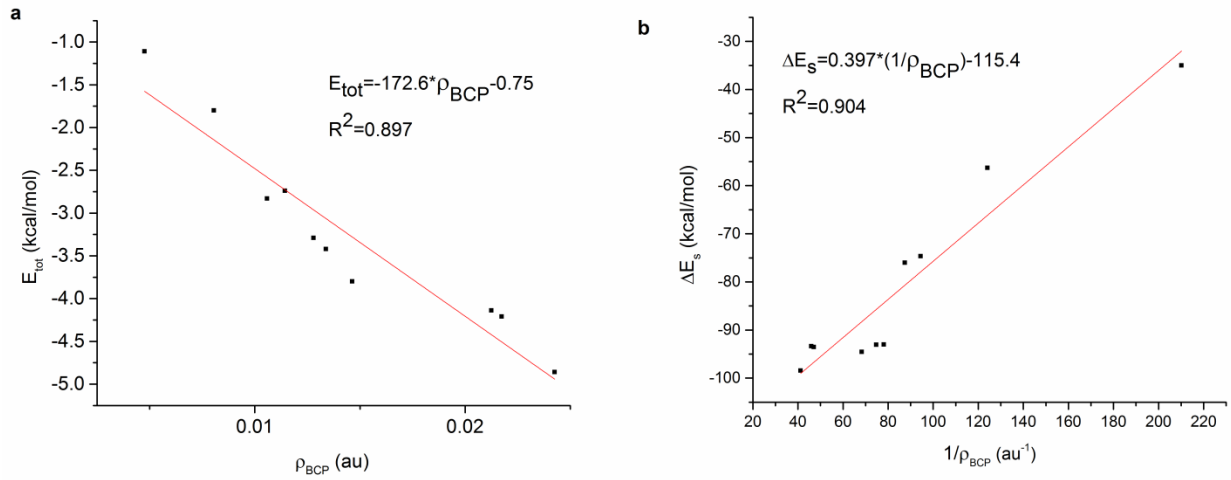


Figure 6.3 Linear fitting of E_{tot} versus ρ_{BCP}(a), and ΔE_s versus 1/ ρ_{BCP} (b) for water dimers.

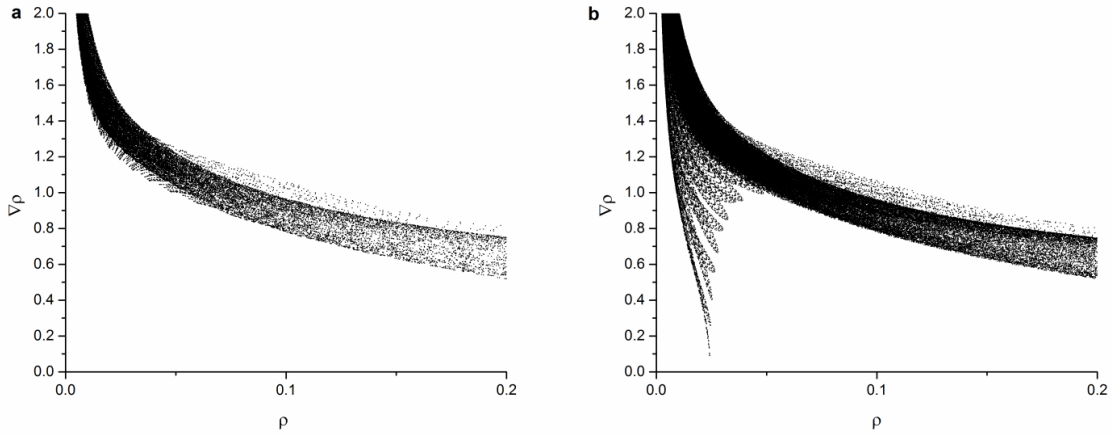


Figure 6.4 Plots of ∇ρ versus ρ for dimer w1: (a) monomer; (b) dimer.(Cufoff for ∇ρ and ρ are 2 and 0.2 respectively to show the intermolecular points)

To gain a further understanding of the components of the DFT steric analysis, we have compared the individual components from this decomposition to individual terms obtained with CSOV. For the electrostatic interactions a good linear relationship is observed (see Figure 6.5a) even though the magnitudes of the values differ between these two methods. The Coulomb interaction from CSOV is less negative than that from steric analysis because CSOV utilizes unperturbed orbitals of monomers while steric analysis uses completely relaxed orbitals of dimers. Since charge

transfer and polarization are due to the relaxation of the orbitals, the relationship between $E_{Coul}+E_{CT}+E_{Pol}$ and ΔE_e is plotted in Figure 6.5b. It is shown that their linear relationship is slightly better, which suggests charge transfer and polarization may include some electrostatic effects from the relaxation of the orbitals. The $E_{Exch\text{-}Rep}$ from CSOV, which is related to the Pauli Exclusion Principle, is comparable to ΔE_x and ΔE_{Pauli}. It can be seen that the $E_{Exch\text{-}Rep}$ interaction from CSOV is positive compared to ΔE_x from DFTs-EDA, which is negative (see Table 6.2). This suggests that the Pauli repulsion from CSOV is so positive that it cancels the stabilization effect from the exchange compared to the stabilization effect from the correlation energy. It should be mentioned here that the exchange from DFTs-EDA, which reflects the energy change when switching the spatial positions of the two electrons interchanged, does not include repulsion effects. The linear fitting of $E_{Exch\text{-}Rep}$ from CSOV versus E_x and E_{Pauli} is shown in Figures 6.6a and 6.6b. Figure 6.6c shows the linear relationship between $E_{Pol}+E_{CT}$ and the steric energy. For the uniform electron gas, because the gradient of the electron density is zero, the steric energy is zero according to equation 3. In this regard, the steric energy reflects the change of the electron density going from the ideal non-interacting electron gas to the molecular environment. For CSOV, as described in the previous section, the polarization and the charge transfer components describe the change in the orbitals (occupied and virtual) of one monomer due to presence of the perturbing field produced by the other monomer. The effects of these two energy components are also reflected on the change of the electron density and its gradient in the dimers. From the perspective of chemical meaning, the steric energy qualitatively represents the steric effect, which represents the effects from the spatial redistribution of the electron density when atoms or molecules are brought together [14,50]. Thus, the polarization and charge transfer from CSOV may capture the steric effect to some extent as well. As ΔE_{Pauli} represents the effects that Pauli exclusion principle take on the kinetic energy, and Pauli exclusion also takes part in the relaxation of the orbitals for calculation, the linear relationship between $E_{Pol}+E_{CT}$ and ΔE_{Pauli} has been fitted in Figure 6.6d.

$E_{Pol}+E_{CT}$ and the steric energy are negative while ΔE_{Pauli} is positive, which suggests that the steric energy contributes more to the $E_{Pol}+E_{CT}$ than ΔE_{Pauli}.

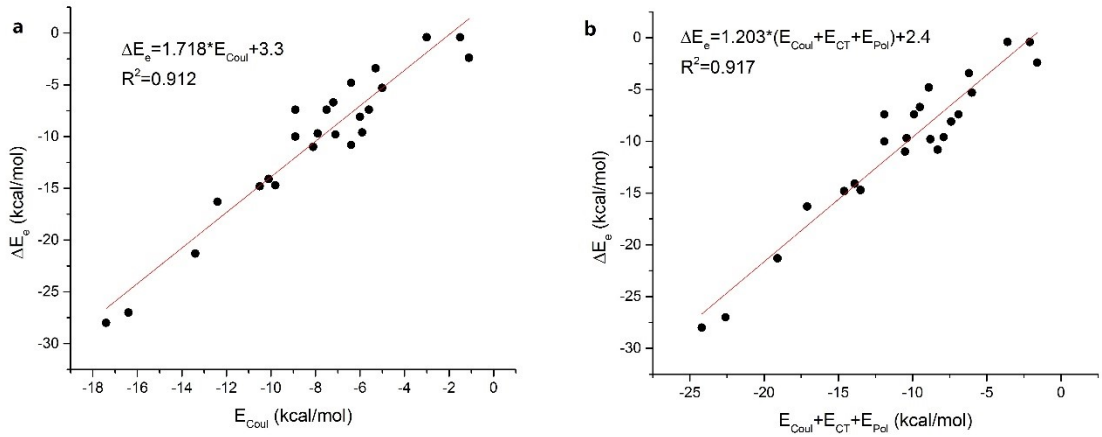


Figure 6.5 Linear fitting of ΔE_e versus E_{Coul} (a) and $E_{Coul}+E_{CT}+E_{Pol}$ (b) for water dimers.

Two DNA base pairs (Figure S1), guanine-cytosine and adenine-thymine paired in the canonical Watson-Crick conformation have also been studied. Because of the limitation on computation time needed by CSOV, the smaller basis set 6-31G(d) was adopted for these two dimers and for the benzene dimers discussed in the following sub-section. The results for CSOV and DFT steric analysis are summarized in Tables 3 and 4, respectively. As shown in Table 6.3, for multiple hydrogen-bonding dimers, the sum of the polarization and charge transfer energy increases and becomes comparable to the Coulomb interaction. Compared to the previous result [35] with a larger basis set, this increase of charge transfer is not only because of the increasing number of hydrogen bonds, but also due to the BSSE effect on the charge transfer component due to the small basis set employed. It can be seen from Table 6.4 that the pair with lower total interaction energy has a lower ΔE_s value, which is consistent with the hydrogen-bonded dimers above. In addition, after the mixing of the orbitals of the monomers, ΔE_e given by steric analysis for A-T is more negative than C-G while the trend is opposite with respect to $\Delta E_{Coulomb}$ from CSOV.

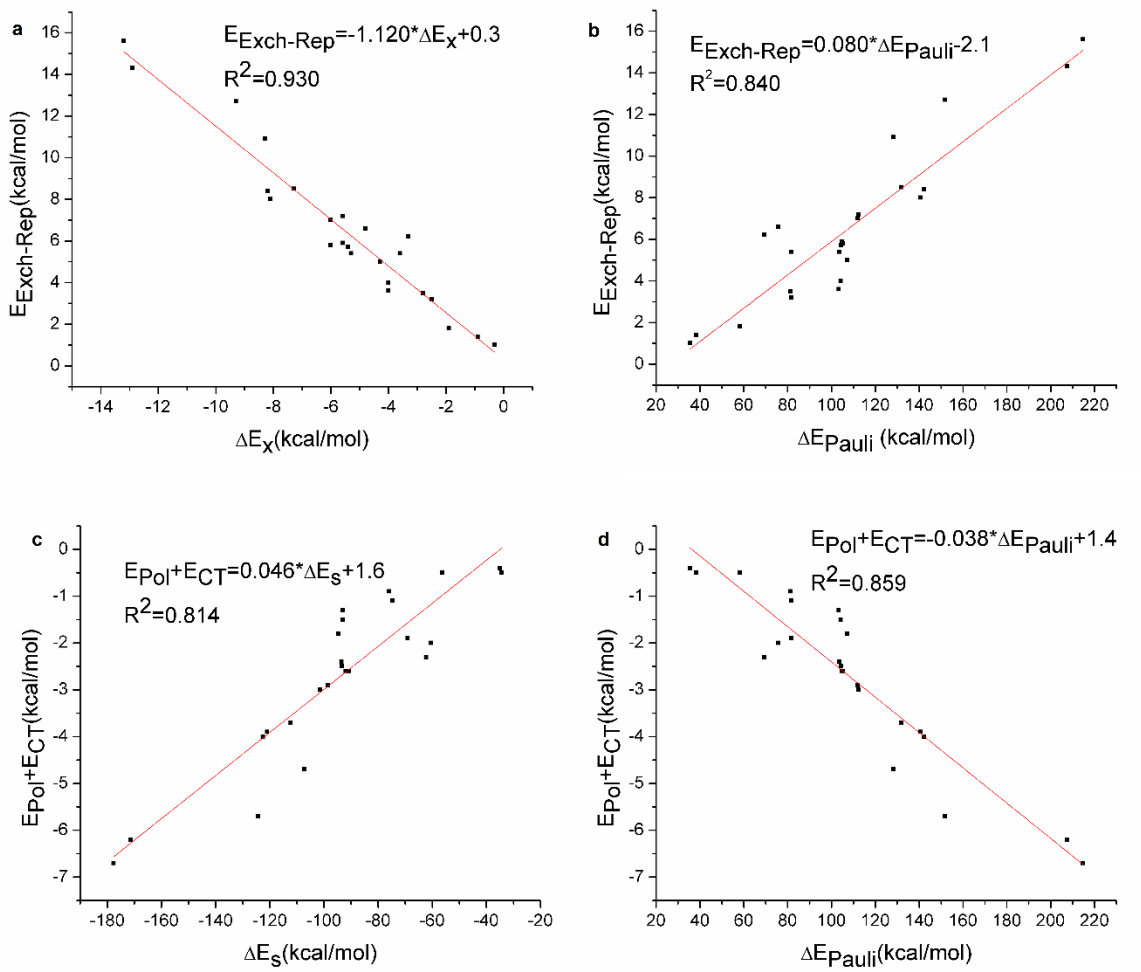


Figure 6.6 Linear fitting of $E_{exch-rep}$ versus ΔE_X (a), $E_{exch-rep}$ versus ΔE_S (b), $E_{Pol}+E_{CT}$ versus ΔE_S (c), and $E_{Pol}+E_{CT}$ versus ΔE_{Pauli} (d) for water dimers.

Table 6.3 CSOV decomposition on DNA base pair. (unit:kcal/mol; B3LYP/6-31G(d))

Dimers	E_{FC}	E_{Coul}	$E_{Exch-Rep}$	E_{PolA}	$E_{CT(A\rightarrow B)}$	E_{PolB}	$E_{CT(B\rightarrow A)}$	E_{tot}
A-T	-1.4	-23.9	22.6	-2.6	-6.7	-1.9	-5.1	-18.1
C-G	-7.1	-28.1	21.0	-3.8	-7.4	-2.6	-4.7	-25.5

Table 6.4 DFT Steric Analysis on DNA base pair. (unit:kcal/mol; B3LYP/6-31G(d))

Dimers	ΔTs	ΔEe	ΔExc	ΔEx	ΔEc	ΔEs	ΔEq	ΔEtot
A-T	50.5	-51.3	-17.5	-11.0	-6.4	-250.7	283.7	-18.3
C-G	41.6	-48.6	-19.1	-11.8	-7.3	-301.8	324.2	-26.2

6.3.2. Benzene dimers

Two types of benzene dimers (Figure 6.7), sandwich and T-shaped, calculated at varying intermolecular distances have been investigated. These two types of benzene dimers have been studied extensively by high-level methods such as MP2, CCSD(T), SAPT-DFT in order to obtain accurate intermolecular interaction energies.[51-60] It is known that traditional DFT methods like B3LYP fail to predict accurate interaction energy for these benzene dimers because of the lack of dispersion term. However, in most cases for DFT, the dispersion is treated by empirical corrections and it can be separated when performing EDA[18], which provides us the chance to do EDA analysis on the dimers using B3LYP without dispersion term.

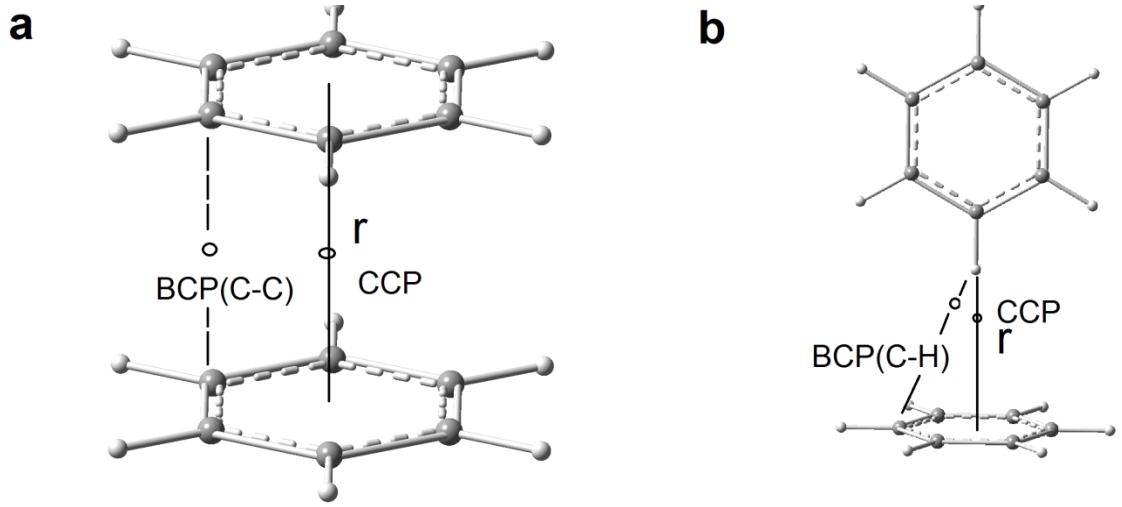


Figure 6.7 Structures of benzene dimers (r represents the distance. For a, it is the distance between the centers of two monomers. For b, it is the distance between one hydrogen atom of the upper benzene and the center of the other monomer. Black circles are critical points for AIM. For BCPs, only one is shown for illustration. BCP stands for bond critical point, and CCP is cage critical point.)

Regarding the sandwich benzene dimer, because of the missing dispersion term, which is one of the main sources to introduce the stabilization effects, the total energies are positive at all distances. It has been proposed that besides the dispersion term, electrostatic interaction and steric repulsion govern the total interaction energy as well.[52] As shown by the CSOV results in Table 6.5, the destabilization effects are mainly contributed by the exchange-repulsion term. It

deceases dramatically with the increasing of the distance. The Coulomb term also destabilizes the dimers. It first increases and then decreases with the increasing of the distance. Compared to the charge transfer term, the polarization energy is smaller. Both of them stabilize the dimers. On the other hand, in the case of the DFT steric results shown in Table 6.6, the electrostatic term (ΔE_e) is much larger than ($E_{Coulomb}$) from the CSOV analysis at short distances and then decreases to a negative value at long distances. When it comes to ΔE_s, instead of positive values that are indicative of the destabilization effects, the values of ΔE_s are negative. As shown in the definition, the DFT steric term excludes the effects from the Pauli Exclusion Principle and should be indicative of pure spatial effects from the rearrangements of the electrons. The so-called steric repulsion for the sandwich benzene dimer should refer to the Pauli repulsion, which is included in ΔE_q and similar to the proposed concept called steric strain.[31]

Table 6.5 CSOV decomposition on sandwich benzene dimer. (unit:kcal/mol; B3LYP/6-31G(d)

r(Å)	E_{FC}	E_{Coul}	$E_{Exch-Rep}$	E_{PolA}	$E_{CT(A\rightarrow B)}$	E_{PolB}	$E_{CT(B\rightarrow A)}$	E_{tot}
3.4	7.7	0.2	7.4	-0.1	-0.6	-0.1	-0.6	6.2
3.5	6.1	0.6	5.4	-0.1	-0.6	-0.1	-0.6	4.6
3.6	4.8	0.8	4.0	-0.1	-0.6	-0.1	-0.6	3.5
3.7	3.9	0.9	3.0	-0.1	-0.6	-0.1	-0.6	2.6
3.8	3.2	1.0	2.2	-0.1	-0.6	-0.1	-0.6	2.0
3.9	2.7	1.0	1.7	-0.1	-0.5	0.0	-0.5	1.6
4.0	2.3	0.9	1.4	0.0	-0.5	0.0	-0.5	1.3
4.1	2.0	0.9	1.1	0.0	-0.4	0.0	-0.4	1.1
4.2	1.7	0.9	0.9	0.0	-0.3	0.0	-0.3	1.0
4.3	1.4	0.8	0.6	0.0	-0.3	0.0	-0.3	0.8
4.4	1.2	0.8	0.5	0.0	-0.2	0.0	-0.2	0.8
5.0	0.7	0.5	0.2	0.0	0.0	0.0	0.0	0.6
5.5	0.4	0.4	0.0	0.0	0.0	0.0	0.0	0.4
6.0	0.3	0.3	0.0	0.0	0.0	0.0	0.0	0.3
8.0	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.1
9.0	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0
10.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 6.6 DFT Steric Analysis on sandwich benzene dimer. (unit:kcal/mol; B3LYP/6-31G(d))

r(Å)	ΔTs	ΔEe	ΔExc	ΔEx	ΔEc	ΔEs	ΔEq	Etot
3.4	-13.7	23.8	-4.0	-0.1	-4.0	-126.9	109.2	6.1
3.5	-7.3	15.0	-3.1	0.1	-3.2	-108.6	98.2	4.5
3.6	-2.7	8.5	-2.4	0.2	-2.6	-92.6	87.5	3.4
3.7	0.6	3.8	-1.8	0.3	-2.1	-78.6	77.4	2.6
3.8	2.8	0.6	-1.4	0.3	-1.7	-66.4	67.8	2.0
3.9	4.2	-1.6	-1.0	0.3	-1.3	-55.8	58.9	1.6
4.0	4.9	-2.9	-0.7	0.3	-1.0	-46.5	50.8	1.3
4.1	5.2	-3.7	-0.5	0.3	-0.8	-38.7	43.4	1.1
4.2	5.2	-4.0	-0.3	0.3	-0.6	-31.9	36.9	1.0
4.3	5.0	-4.0	-0.2	0.2	-0.4	-26.2	31.1	0.9
4.4	4.7	-3.8	-0.1	0.2	-0.3	-21.4	26.0	0.8
5.0	2.8	-2.3	0.1	0.1	0.0	-5.5	8.4	0.6
5.5	2.0	-1.7	0.1	0.1	0.1	-1.4	3.4	0.4
6.0	1.5	-1.3	0.1	0.0	0.0	-0.2	1.7	0.3
8.0	0.5	-0.4	0.0	0.0	0.0	0.1	0.4	0.1
9.0	0.3	-0.3	0.0	0.0	0.0	0.0	0.3	0.1
10.0	0.2	-0.2	0.0	0.0	0.0	0.0	0.2	0.0

We have also calculated intermolecular BCPs and their electron density for these dimers. Two types of intermolecular critical points (CP), BCPs and Cage Critical Points (CCP) are depicted in Figure 6.7. It can be seen that these two types of points have a good linear relationship with each other as shown in Figure 6.8. It can be seen that in contrast to the water dimers discussed above, $\rho_{BCP(C\text{-}C)}$ and E_{tot} do not show a linear relationship. However, $\rho_{BCP(C\text{-}C)}$ correlates linearly with ΔE_s with a negative slope. This is consistent with the decreasing trend in ΔE_s as the electron density increases at intermolecular BCPs, similar to the effect observed in the water dimers above.

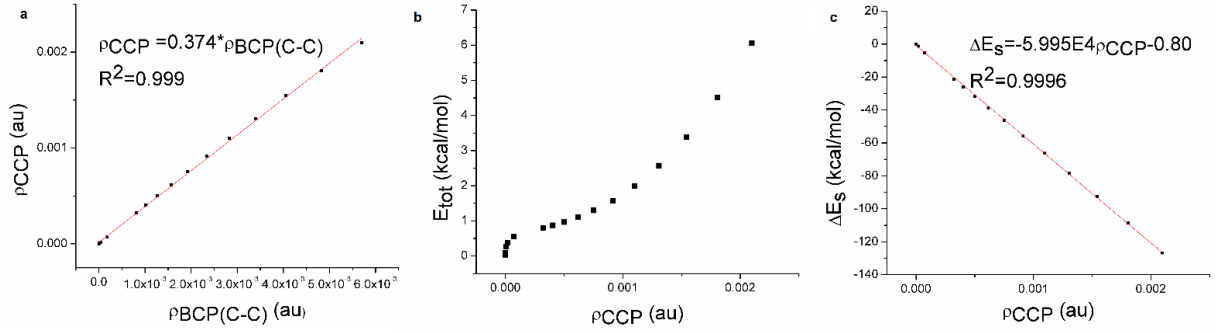


Figure 6.8 (a) Linear fitting of ρ_{CCP} versus $\rho_{BCP(C\text{-}C)}$. (b) Plot of E_{tot} versus ρ_{CCP}. (c) Linear fitting of ΔE_s versus ρ_{CCP} for the benzene dimer in the sandwich conformation.

Since dispersion is not the main contribution to the total interaction energy for the T-shaped benzene dimer, the values of the total interaction energy obtained by B3LYP (see Table 7 and Table 8) are negative and closer to the results from the high-level methods. [51] Within the separation distances we studied, the results from CSOV show that Coulomb interaction stabilizes the dimer, which is consistent with the previous SAPT-DFT result[55]. However, after the mixing of the orbitals of the dimers, the energy becomes positive and destabilizes the dimer, as indicated by the results from DFT steric analysis. Both methods indicate the destabilization effects arise from the exchange term. As shown by the results from CSOV, the charge transfer from the bottom benzene to the upper benzene, mainly π electron transferred to the empty orbital of the hydrogen atom, is comparable to the Coulomb term. The plots of E_{tot} and ΔE_s versus ρ_{CCP} are shown in Figure 6.9. It is observed that ΔE decreases with the increasing of ρ_{CCP}, which correlates linearly with ΔE_s.

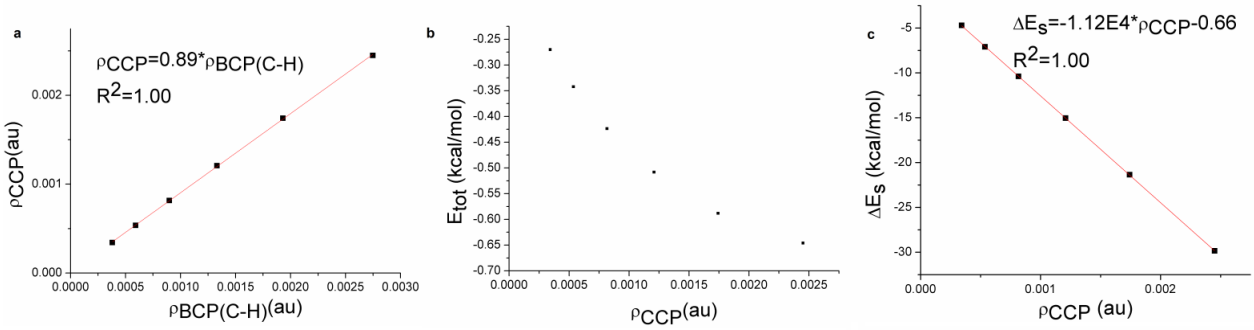


Figure 6.9 (a) Linear fitting of ρ_{CCP} versus $\rho_{BCP(C-H)}$. (b) Plot of E_{tot} versus ρ_{CCP}. (c) Linear fitting of ΔE_s versus ρ_{CCP} for the T-shaped benzene dimer.

Table 6.7 CSOV decomposition on T-shaped benzene dimers with B3LYP/6-31G(d) (unit: kcal/mol; the upper benzene is denoted as A.)

r(Å)	E_{FC}	E_{Coul}	$E_{Exch-Rep}$	E_{PolA}	$E_{CT(A\rightarrow B)}$	E_{PolB}	$E_{CT(B\rightarrow A)}$	Etot
3.0	0.2	-0.5	0.7	-0.1	-0.1	0.0	-0.7	-0.6
3.2	0.2	-0.4	0.5	-0.1	-0.1	0.0	-0.5	-0.5
3.4	-0.1	-0.3	0.2	-0.1	0.0	0.0	-0.4	-0.5
3.6	-0.1	-0.2	0.2	0.0	0.0	0.0	-0.3	-0.5
3.8	0.0	-0.2	0.2	0.0	0.0	0.0	-0.2	-0.3
4.0	0.0	-0.2	0.2	0.0	0.0	0.0	-0.2	-0.2

Table 6.8 DFT Steric Analysis on T-shaped benzene dimers with B3LYP/6-31G(d) (kcal/mol)

r(Å)	ΔTs	ΔEe	ΔExc	ΔEx	ΔEc	ΔEs	ΔEq	Etot
3.0	-2.2	1.5	0.0	0.8	-0.8	-29.9	27.7	-0.7
3.2	-1.4	0.6	0.3	0.8	-0.5	-21.4	20.2	-0.6
3.4	-1.0	0.2	0.3	0.6	-0.3	-15.0	14.4	-0.5
3.6	-0.7	0.1	0.3	0.5	-0.2	-10.4	9.9	-0.4
3.8	-0.6	0.1	0.2	0.3	-0.1	-7.1	6.6	-0.3
4.0	-0.5	0.2	0.0	0.1	-0.1	-4.7	4.2	-0.3

6.4. Conclusions

Herein we have applied Density Functional steric analysis, a new energy decomposition method, combined with the CSOV and QTAIM to investigate the components that contribute to the total interaction for dimers from a different perspective. One can gain more understandings of the origins of the formation of dimers by combining these two analyses. It has been shown that there is a good linear relationship between the change of DFT steric energy and the total interaction energy for hydrogen-bonding dimers. In addition, comparisons between components of CSOV and DFT steric energy decomposition show the effects of the relaxation of the orbitals of monomers and the linear relationship between the polarization and charge transfer from CSOV and the steric and Pauli energy from DFT steric analysis. In combination with QTAIM analysis, it has been found that the change of DFT steric energy directly correlates with the electron density in the critical points. This suggests the steric energy is indicative of the energy change arising from the spatial distribution of the electron density and also helps explain the connection between

the change of DFT steric energy and the total interaction energy. This new decomposition method provides a way to quantify the spatial effects on the electron density change during the formation of dimers.

6.5 References

(1) Hirschfelder, J. O. *Chem. Phys. Lett.* **1967**, *1*, 363.

(2) Murrell, J. N.; Shaw, G. *J. Chem. Phys.* **1967**, *46*, 1768.

(3) Kitaura, K.; Morokuma, K. *Int. J. Quantum Chem.* **1976**, *10*, 325.

(4) Ziegler, T.; Rauk, A. *Inorg. Chem.* **1979**, *18*, 1558.

(5) Bagus, P. S.; Hermann, K.; Bauschlicher, J. C. W. *J. Chem. Phys.* **1984**, *80*, 4378.

(6) Stevens, W. J.; Fink, W. H. *Chem. Phys. Lett.* **1987**, *139*, 15.

(7) Glendening, E. D.; Streitwieser, A. *J. Chem. Phys.* **1994**, *100*, 2900.

(8) Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994**, *94*, 1887.

(9) Chen, W.; Gordon, M. S. *J. Phys. Chem.* **1996**, *100*, 14316.

(10) Mo, Y.; Gao, J.; Peyerimhoff, S. D. *J. Chem. Phys.* **2000**, *112*, 5530.

(11) Mayer, I. *Chem. Phys. Lett.* **2003**, *382*, 265.

(12) Hesselmann, A.; Jansen, G.; Schutz, M. *J. Chem. Phys.* **2005**, *122*, 014103.

(13) Khaliullin, R. Z.; Cobar, E. A.; Lochan, R. C.; Bell, A. T.; Head-Gordon, M. *J. Phys. Chem. A* **2007**, *111*, 8753.

(14) Liu, S. *J. Chem. Phys.* **2007**, *126*, 244103.

(15) Reinhardt, P.; Piquemal, J.-P.; Savin, A. *J. Chem. Theory Comput.* **2008**, *4*, 2020.

(16) G. A. Cisneros, T. A. D., N. Gresh, P. Reinhardt, O. Parisel, J. Pilmé and J-P Piquemal *Design of next generation polarizable force fields from ab initio computations: beyond point charges*; Springer-Verlag: London, 2009.

(17) Su, P.; Li, H. *J. Chem. Phys.* **2009**, *131*, 014102.

(18) Wu, Q.; Ayers, P. W.; Zhang, Y. *J. Chem. Phys.* **2009**, *131*, 164112.

(19) Steinmann, S. N.; Corminboeuf, C.; Wu, W.; Mo, Y. *J. Phys. Chem. A* **2011**, *115*, 5467.

(20) Azar, R. J.; Head-Gordon, M. *J. Chem. Phys.* **2012**, *136*, 024103.

(21) Szczesniak, M. M.; Chalasinski, G.; Cybulski, S. M. *J. Chem. Phys.* **1992**, *96*, 463.

(22) Khaliullin, R. Z.; Bell, A. T.; Head-Gordon, M. *Chem. Eur. J.* **2009**, *15*, 851.

(23) Liu, S.; Govind, N. *J. Phys. Chem. A* **2008**, *112*, 6690.

(24) Huang, Y.; Zhong, A.-G.; Yang, Q.; Liu, S. *J. Chem. Phys.* **2011**, *134*, 084103.

(25) Liu, S.; Hu, H.; Pedersen, L. G. *J. Phys. Chem. A* **2010**, *114*, 5913.

(26) Bader, R. F. W. *Acc. Chem. Res.* **1985**, *18*, 9.

(27) Weizsacker, C. F. v. *Z. Phys. A Hadrons Nucl.* **1935**, *96*, 431.

(28) Rong, C.; Lu, T.; Liu, S. *J. Chem. Phys.* **2014**, *140*, 024109.

(29) Badenhoop, J. K.; Weinhold, F. *J. Chem. Phys.* **1997**, *107*, 5406.

(30) Nagy, Á. *Chem. Phys. Lett.* **2007**, *449*, 212.

(31) Pinter, B.; Fievez, T.; Bickelhaupt, F. M.; Geerlings, P.; De Proft, F. *Phys. Chem. Chem. Phys.* **2012**, *14*, 9846.

(32) Bagus, P. S.; Illas, F. *J. Chem. Phys.* **1992**, *96*, 8962.

(33) Neyman, K. M.; Ruzankin, S. P.; Rösch, N. *Chem. Phys. Lett.* **1995**, *246*, 546.

(34) Márquez, A. M.; Lopez, N.; García-Hernández, M.; Illas, F. *Surf. Sci.* **1999**, *442*, 463.

(35) Piquemal, J.-P.; Marquez, A.; Parisel, O.; Giessner–Prettre, C. *J. Comput. Chem.* **2005**, *26*, 1052.

(36) Johnson, E. R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W. *J. Am. Chem. Soc.* **2010**, *132*, 6498.

(37) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(38) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(39) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.;

Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.; Gaussian, Inc.: Wallingford CT, 2009.

(40) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.

(41) Dupuis, M.; Marquez, A.; Davidson, E. R.; Quantum Chemistry Program Exchange (QCPE): Bloomington IN, 1999.

(42) Valiev, M.; Bylaska, E. J.; Govind, N.; Kowalski, K.; Straatsma, T. P.; Van Dam, H. J. J.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T. L.; de Jong, W. A. *Comput. Phys. Commun.* **2010**, *181*, 1477.

(43) Contreras-García, J.; Johnson, E. R.; Keinan, S.; Chaudret, R.; Piquemal, J.-P.; Beratan, D. N.; Yang, W. *J. Chem. Theory Comput.* **2011**, *7*, 625.

(44) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.

(45) Balabin, R. M. *J. Chem. Phys.* **2008**, *129*, 164101.

(46) Saritha, B.; Durga Prasad, M. *J Chem Sci* **2012**, *124*, 209.

(47) Saritha, B.; Durga Prasad, M. *J. Phys. Chem. A* **2011**, *115*, 2802.

(48) Wang, Y.; Zhao, D.; Rong, C.; Liu, S. *Acta Phys.-Chim. Sin.* **2013**, *29*, 2173.

(49) Parthasarathi, R.; Subramanian, V.; Sathyamurthy, N. *J. Phys. Chem. A* **2006**, *110*, 3349.

(50) Ess, D. H.; Liu, S.; De Proft, F. *J. Phys. Chem. A* **2010**, *114*, 12952.

(51) Jaffe, R. L.; Smith, G. D. *J. Chem. Phys.* **1996**, *105*, 2780.

(52) Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* **2006**, *110*, 10656.

(53) Zhikol, O. A.; Shishkin, O. V. *Int. J. Quantum Chem.* **2012**, *112*, 3008.

(54) Zhikol, O. A.; Shishkin, O. V.; Lyssenko, K. A.; Leszczynski, J. *J. Chem. Phys.* **2005**, *122*, 144104.

(55) Podeszwa, R.; Bukowski, R.; Szalewicz, K. *J. Phys. Chem. A* **2006**, *110*, 10345.

(56) van der Avoird, A.; Podeszwa, R.; Szalewicz, K.; Leforestier, C.; van Harrevelt, R.; Bunker, P. R.; Schnell, M.; von Helden, G.; Meijer, G. *Phys. Chem. Chem. Phys.* **2010**, *12*, 8219.

(57) Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. *J. Am. Chem. Soc.* **2002**, *124*, 10887.

(58) Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* **2004**, *108*, 10200.

(59) Tauer, T. P.; Sherrill, C. D. *J. Phys. Chem. A* **2005**, *109*, 10475.

(60) Sherrill, C. D.; Takatani, T.; Hohenstein, E. G. *J. Phys. Chem. A* **2009**, *113*, 10146.

ABSTRACT

COMPUTATIONAL STUDIES OF ENZYMATIC REACTIONS AND INORGANIC COMPOUNDS

by

DONG FANG

August 2014

Advisor: Professor G. Andrés Cisneros

Major: Chemistry (Physical)

Degree: Doctor of Philosophy

Quantum mechanics (QM) calculations have been demonstrated to give very accurate results in studying a variety of chemical reactions. However, the current computing resources limits its use in large-size systems such as enzymes. Compared to QM, Molecular Mechanics (MM) have an advantage in computing time but disadvantage in accuracy, and MM cannot be used in reactions involving bonding forming and breaking. To combine the advantages of these two types of methods, QM/MM has been introduced to study reactions and processes in large-size systems, and it has been proven to give accurate results with moderate computational expense. When studying enzymatic reactions, a comprehensive understanding of the interactions in an intuitive way is highly instructive. The combined ELF/NCI analysis meets the demand. In this dissertation, we have discussed the applications and development of the combined ELF/NCI analysis and QM/MM method. The combined ELF/NCI analysis has been extended to study two enzymatic reactions in Chapter 2, the synthesis of DNA by Human DNA polymerase λ and the ε subunit of DNA Polymerase III and help reveal the factors that account for the higher reaction barrier for Mg^{2+} than Mn^{2+} and different coordination numbers for two metal centers with respect to the interactions. It also has been used to investigate the interactions that result in the unusual

seesaw shape of several metal clusters in Chapter 3. QM/MM simulations has been carried out to study the steps after the formation of Fe^{IV}-oxo for the repair of 1-methyladinine (1-meA) catalyzed by AlkB. First, the rate-limiting step, the hydrogen abstraction step for the H_2O pathway has been fully studied in Chapter 4. The results shed lights on the electronic structure of Fe^{IV}-oxo. Residues that may affect the reaction barrier has been obtained by EDA. The combined ELF/NCI analysis also reveals important interactions and their changes along the reaction. We have also proposed and studied a new OH^- pathway in Chapter 5. Comparison with the traditional H_2O pathway provides more details on the mechanism. The consistency between our theoretical results and experimental findings have been found. In Chapter 6, a newly proposed EDA method, DFT-steric analysis is performed on several types of dimers. Comparison between it and another EDA method CSOV and QTAIM provides a deep understanding of the components of the interactions between monomers from different perspectives.

AUTOBIOGRAPHICAL STATEMENT

Dong Fang

Education

2009-2014 Ph.D., Physical Chemistry

Department of Chemistry, Wayne State University

2005-2009 B.S., Chemistry

Department of Chemistry, University of Science and Technology of China

Honours and awards

Member of Phi Lambda Upsilon (PLU)

Thomas C. Rumble University Graduate Fellowship (2013-2014)

Dan Trivich Memorial Award, 2014 (WSU)

Publications

1. "Toward a Deeper Understanding of Enzyme Reactions Using the Coupled ELF/NCI Analysis: Application to DNA Repair Enzymes", D. Fang, R. Chaudret, J.-P. Piquemal, and G. A. Cisneros, J. Chem. Theory Comput. 2013 9 (5), 2156-2160
2. "Ab Initio QM/MM Calculations Show an Intersystem Crossing in the Hydrogen Abstraction Step in Dealkylation Catalyzed by AlkB", D Fang, R. L. Lord, and G. A. Cisneros, J. Phys. Chem. B 2013 117 (21), 6410-6420
3. "Novel Alkoxide Cluster Topologies Featuring Rare Seesaw Geometry at Transition Metal Centers", J. A. Bellow, D. Fang, N. Kovacevic, P. D. Martin, J. Shearer, G. A. Cisneros, S. Groysman, Chem. Eur. J. 2013, 19 (37), 12225-12228
4. "DFT-steric-based energy decomposition analysis for intermolecular interactions", D. Fang, J.-P. Piquemal, S. Liu, G. A. Cisneros, Theo. Chem. Acc. 2014, 133 (5), 1484
5. "Ab initio QM/MM investigation of the catalytic steps following hydrogen abstraction in the reaction catalyzed by DNA dealkylase AlkB", D. Fang, G. A. Cisneros, to be submitted.
6. "A new smoothing function to account for long-range electrostatics in QM/MM", D. Fang, R. E. Duke, G. A. Cisneros, in preparation
7. "Understanding selectivity of hard and soft metal cations within biological systems using the subvalence concept II. Factors Controlling the Subvalence effects and compared visions of ELF and EPLF", R. Chaudret, D. Fang, R. L. Lord, G. A. Cisneros, B. Courcy, J.-P. Piquemal, in preparation.