Computational modeling of photolysis and catalysis reactions in vitamin b12-dependent enzymes.

Abdullah Al Mamun

University of Louisville

Follow this and additional works at: https://ir.library.louisville.edu/etd

Part of the Physical Chemistry Commons

Recommended Citation
https://doi.org/10.18297/etd/3406

This Doctoral Dissertation is brought to you for free and open access by ThinkIR: The University of Louisville's Institutional Repository. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of ThinkIR: The University of Louisville's Institutional Repository. This title appears here courtesy of the author, who has retained all other copyrights. For more information, please contact thinkir@louisville.edu.
COMPUTATIONAL MODELLING OF PHOTOLYSIS AND CATALYSIS REACTIONS IN VITAMIN B$_{12}$-DEPENDENT ENZYMES

By
Abdullah Al Mamun
B.Sc., University of Dhaka, 2013
M.Sc., University of Dhaka 2015
M.S., University of Louisville 2018

A Dissertation
Submitted to the Faculty of the
College of Arts and Sciences of the University of Louisville
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy in Chemistry

Department of Chemistry
University of Louisville
Louisville, Kentucky

May 2020
COMPUTATIONAL MODELLING OF PHOTOLYSIS AND CATALYSIS REACTIONS IN VITAMIN B\textsubscript{12}-DEPENDENT ENZYMES

By

Abdullah Al Mamun
B.Sc., University of Dhaka, 2013
M.Sc., University of Dhaka 2015
M.S., University of Louisville 2018

A Dissertation Approved on

February 11, 2020

By the following Dissertation Committee

______________________________
Dr. Pawel M. Kozlowski (Dissertation Director)

______________________________
Dr. Craig A. Grapperhaus

______________________________
Dr. Lee M. Thompson

______________________________
Dr. Noppadon Sathithsuksanoh

______________________________
Dr. Justyna Jaroszyńska-Wolińska
DEDICATION

This dissertation is dedicated to the people who
sacrificed their lives during the liberation war of Bangladesh in 1971
ACKNOWLEDGEMENTS

I would like to thank so many people for their incredible support, help, and encouragement. First, I am very grateful to my thesis supervisor Dr. Pawel M. Kozlowski. It was a wonderful experience working with him. I have learned so many lessons from his guidance which I will cherish in my whole life. His guidance, encouragement, dedication and perseverance were the best thing for a graduate student like me. Hopefully, I will also get his insights and suggestions for my future journey.

I want to acknowledge my committee members Dr. Craig Grapperhaus, Dr. Lee Thompson, Dr. Noppadon Sathitsuksanoh for their continued assistance, and advice during my Ph.D. I am grateful to them for their support. I would like to thank my ad-hoc committee member Dr. Justyna Jaroszyńska-Wolińska, for her willingness to be a committee member in my dissertation. I am extremely grateful to the Department of Chemistry for offering me the teaching assistantship position to obtain my doctorate. It would not be possible without the generous support from the department. In addition to these, I am grateful to a number of people in the department of chemistry, Sherry Nally, Dr. Frank Zamborini, Dr. Jinjun Liu, Dr. Baldwin, Dr. Powe for their help during my stay at the Department of Chemistry. Harrison Simral from Cardinal Research Cluster (CRC) was an invaluable support for me. I took his help so many times to address the technical problems while using CRC.

Collaboration is an essential thing for doing research. I am thankful for having the chance to collaborate with so many tremendous people. Dr. Piotr Lodowski is a wonderful collaborator who is always ready to provide valuable suggestions and insights for many
complex problems. I am grateful to him for his incredible support and knowledge that was shared to enhance my research. I would also like to thank Dr. Maria Jaworska, Dr. Craig Grapperhaus, Dr. Robert Buchanan, Dr. Takasi Ogura, Dr. Teizo Kitagawa, Dr. Rahul Jain Steven Cronin for being a wonderful collaborator.

My stay at Louisville was pleasant because of so many friends I have in my school. I would like to acknowledge my friend Dr. Anam Chandra Paul for his support and friendship, Dr. Ramya Billur, a terrific friend who supported and encouraged me in so many cases, Jahangir Alom, Anwar Hossain, Ram Hona, Rocky, Likhon, Sukanto, and so many others. My fellow lab-mate and friend Dr. Brady Garabato helped me a lot in adapting to my initial research at the University of Louisville. At last but not least my two important lab-mate, Megan J. Toda and Arghy Pratim Ghosh who are excellent collaborators, friends and a wonderful colleague. I had the most pleasant experience working with them.

Finally, my sincere thanks and gratitude is to my family members who are constantly supporting me since my childhood. I cannot express my gratitude in words to my elder brother who is always helping me and supporting me, my wonderful sister and nephews, my father, and my mother. They are the inspiration for any of my success and deserves my greatest admiration.
ABSTRACT

COMPUTATIONAL MODELLING OF PHOTOLYSIS AND CATALYSIS REACTIONS IN VITAMIN B_{12}-DEPENDENT ENZYMES

Abdullah Al Mamun

February 11, 2020

Vitamin B_{12} is a complex organometallic molecule, the derivatives of which such as adenosylcobalamin (AdoCbl) and methylcobalamin (CH_{3}Cbl) act as a cofactor in numerous enzymatic reactions. These two biologically active cofactors contain a unique organometallic Co-C σ bond. Important feature of this Co-C bond is that it can be activated by both thermally and photolytically inside the enzymatic environment as well as in the solution. In the case of enzymatic reactions where AdoCbl molecule act as a cofactor, the cleavage of the Co-C bond constitutes the key catalytic step. The most intriguing features of this cleavage is that upon binding with a substrate the cleavage of the Co-C bond is 10^{12}-fold rate accelerated inside enzyme compared to the isolated cofactor. There are number of factors responsible behind this trillion-fold rate acceleration which are still under investigation. Alternatively, the Co-C bond in AdoCbl can also be cleaved by light to generate the same radical pair (RP) as in the case of enzymatic catalysis. The light sensitivity of these cobalamins was known for more than five decades only until recently it has been associated with controlled reactivity, namely optogenetic regulation and light-
activated drug delivery. Moreover, the ability to probe photolytic cleavage of the Co–C bond for enzyme bound AdoCbl is of particular relevance in enzymatic catalysis. Herein this study the photolysis and native catalysis mechanism for vitamin B_{12}-dependent enzymes will be investigated using a combined quantum mechanics/molecular mechanics (QM/MM) approach. According to our studies, it appears that the enzymatic environment controls the cleavage of Co–C bond by exerting an electrostatic interaction. In addition, the enzymatic environment controls the generated Ado radical path by providing a cage. This ultimately effects the formation of the product in native catalysis as well as in photolysis. Finally, I will be discussing a plausible connection between photolysis and native catalysis mechanism in vitamin B_{12}-dependent enzymes based on the idea that the protein environment reduces the cofactor. The reaction mechanism based on the reduced cofactor in native catalysis was found to be similar with the photo-dissociation mechanism.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................................................................................ iv

ABSTRACT .......................................................................................................................... vi

LIST OF TABLES .................................................................................................................. xii

LIST OF FIGURES ............................................................................................................... xiii

LIST OF SCHEMES .............................................................................................................. xx

## CHAPTER PAGE

1 INTRODUCTION 1

1.1 Overview of the dissertation ................................................................. 1

1.2 Vitamin B$_{12}$ and its derivatives ............................................................. 3

1.3 Structural model of vitamin B$_{12}$ derivatives ............................................ 5

1.4 Vitamin B$_{12}$-dependent enzymes .............................................................. 9

1.5 Native catalysis mechanism in AdoCbl-dependent enzymes ...................... 13

1.6 Photolysis of Co-C bond in AdoCbl-dependent enzymes ......................... 18

1.7 Motivation towards understanding catalysis and photolysis ...................... 25

2 COMPUTATIONAL METHODOLOGY 29

2.1 Background ................................................................................................. 29

2.2 QM/MM theory ............................................................................................ 30

2.3 QM/MM model preparation ....................................................................... 35

2.4 DFT/MM calculations .................................................................................. 36

2.5 TD-DFT/MM calculations ........................................................................... 41
2.6 Long-range charge transfer (LR-CT) excitations in TD-DFT

3 MECHANISM OF LIGHT INDUCED RADICAL PAIR FORMATION IN COENZYME B$_{12}$-DEPENDENT ETHANOLAMINE AMMONIA LYASE

3.1 Background

3.2 Computational details

3.2.1. QM/MM set up

3.2.2. DFT/MM calculations

3.2.3. TD-DFT/MM calculations

3.3 Results and discussions

3.3.1 Ground state geometry

3.3.2 Electronically excited states along the Co-C bond distance

3.3.3 Potential energy surfaces as a function of axial bond lengths

3.3.4 Analysis of long-range charge transfer excitations

3.3.5 Pathways of photodissociation

3.3.6 Implications of AdoCbl photochemistry inside EAL

3.3.7 Comparison with experiments

3.3.8 New perspective for photochemistry of AdoCbl inside EAL

3.4 Summary and conclusions

4 PHOTOLYTIC CLEAVAGE OF THE CO-C BOND IN COENZYME B$_{12}$-DEPENDENT GLUTAMATE MUTASE

4.1 Introduction

4.2 Computational details

4.3 Results and discussions
4.3.1 Ground state geometry.................................................................85
4.3.2 PECs as a function of Co-C bond length..................................89
4.3.3 S\textsubscript{1} PES as a function of axial bond lengths............91
4.3.4 Photodissociation mechanism based of S\textsubscript{1} PES........94
4.3.5 Comparison of photolytic mechanism in GLM with EAL..........96
4.3.6 Comparison with experiments..................................................99
   4.3.6.1 Transient absorption spectroscopy and EPR studies..........99
   4.3.6.2 Resonance Raman spectroscopy........................................102
   4.3.6.3 Oxidation states of Co ion in MLCT state.......................103
   4.3.6.4 Quantum yield and relaxation rates................................106
4.3.7 Implications of GLM photochemistry......................................108
4.4 Conclusions..................................................................................108

5 A COMPUTATIONAL INVESTIGATION OF NATIVE CATALYSIS
MECHANISM IN ADOCBL-DEPENDENT ETHANOLAMINE AMMONIA
LYASE .........................................................................................110

5.1 Background....................................................................................110
5.2 Computational details.................................................................114
   5.2.1. System preparation...............................................................114
   5.2.2. Molecular dynamics simulation..........................................115
   5.2.3. QM/MM calculations............................................................116
5.3 Results and discussions..............................................................117
   5.3.1 Ground state geometry........................................................117
   5.3.2 Cleavage of the Co-C bond....................................................120
6 CAN PHOTOLYSIS OF THE CO-C BOND IN COENZYME B\textsubscript{12}-DEPENDENT ENZYMES BE USED TO MIMIC THE NATIVE REACTION?

135

6.1 Introduction.................................................................135
6.2 Computational details..................................................138
6.3 Results and discussions................................................140
   6.3.1 Potential energy curves as a function of Co-C bond lengths........140
   6.3.2 Low-lying excited state PES as a function of axial ligands.........142
   6.3.3 Ground state PES as a function of two axial coordinates..........144
   6.3.4 Assessment of our hypothesis with literature.......................146
   6.3.5 The connection between photolysis and catalysis....................149
   6.3.6 Potential role of the reduced cofactor in native catalysis.........152
6.4 Concluding remarks....................................................153

7 CONCLUSIONS AND FUTURE DIRECTIONS

REFERENCES

APPENDIX

LIST OF ABBREVIATIONS

CURRICULUM VITAE
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>5</td>
</tr>
<tr>
<td>3.2</td>
<td>9</td>
</tr>
<tr>
<td>3.3</td>
<td>63</td>
</tr>
<tr>
<td>4.1</td>
<td>8</td>
</tr>
<tr>
<td>4.2</td>
<td>8</td>
</tr>
<tr>
<td>4.3</td>
<td>90</td>
</tr>
<tr>
<td>4.4</td>
<td>9</td>
</tr>
<tr>
<td>5.1</td>
<td>118</td>
</tr>
</tbody>
</table>

3.1 Selected geometrical parameters for optimized and crystal structures of AdoCbl cofactor in EAL and in solution.

3.2 Selected structural data of AdoCbl in different environment and the conformations of the ribose ring in Ado ligand.

3.3 Eight lowest vertical singlet electronic transitions and orbital characterizations based on the single point TD-DFT/MM calculations in EAL.

4.1 Selected geometrical parameters like bond length, bond angles, torsional angles for the optimizes structure of AdoCbl in GLM without the substrate, optimized structure of GLM with the substrate, crystal structure of GLM and the optimized structure of EAL without the substrate.

4.2 Selected geometric parameters like exocyclic angles and endocyclic angle of the Ado ligand to determine the ribose orientation using pseudorotation phase (P) and pseudorotation angle (Ωm).

4.3 Eight lowest vertical singlet excited states characterized by the detail orbital analysis.

4.4 Orbital characterization and DCT diagnostics for several points from the MLCT and LF regions of the S1 PES.

5.1 Structural parameters of AdoCbl cofactor in different environment.
**LIST OF FIGURES**

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>The molecular structure of Cbl and the upper axial ligands</td>
</tr>
<tr>
<td>1.2</td>
<td>Base-on and base-off structural analog of Cbl</td>
</tr>
<tr>
<td>1.3</td>
<td>Structural model of vitamin B$_{12}$ and its analog</td>
</tr>
<tr>
<td>1.4</td>
<td>Reactions catalyzed by Cbl-dependent enzymes</td>
</tr>
<tr>
<td>1.5</td>
<td>Native catalytic cycle of AdoCbl-dependent enzymes</td>
</tr>
<tr>
<td>1.6</td>
<td>Proposed mechanism for the Co-C bond cleavage in AdoCbl enzymes</td>
</tr>
<tr>
<td>1.7</td>
<td>Different orientations of the ribose ring in AdoCbl cofactor found in AdoCbl-dependent enzymatic catalysis</td>
</tr>
<tr>
<td>1.8</td>
<td>Schemes of photolysis for AdoCbl in solution and in its dependent enzymes</td>
</tr>
<tr>
<td>1.9</td>
<td>Photochemical mechanism of Co-C bond cleavage and the conformational and structural changes in AdoCbl based photoreceptor CarH</td>
</tr>
<tr>
<td>2.1</td>
<td>Schematic diagram of protein with different partition</td>
</tr>
<tr>
<td>2.2</td>
<td>A comparison of bond dissociation energy (BDE) and the bond lengths using different DFT functional</td>
</tr>
<tr>
<td>3.1</td>
<td>Native catalysis cycle for AdoCbl-dependent EAL</td>
</tr>
<tr>
<td>3.2</td>
<td>Photolysis of Co-C bond in AdoCbl-dependent EAL</td>
</tr>
</tbody>
</table>
3.3 The structures of EAL. a) crystal structure of EAL obtained from the protein data
bank containing two asymmetric units (PDB ID: 3abs). b) one asymmetric unit of
the crystal structure used for the calculations. c) Structure of EAL optimized with
QM/MM level of theory. d) the active site of AdoCbl used for the calculations in
the high layer model.................................................................51

3.4 Truncated structure of AdoCbl cofactor, containing atom numbers, which was used
in the QM region (high layer). (b) Model structure of RibCbl..............................52

3.5 (a) Crystal structure of EAL (PDB id: 3ABS) with AdePeCbl cofactor where
adenine-9-yl-pentyl (AdePe) is the upper axial ligand. (b) Optimized structure of
EAL with reconstructed AdoCbl cofactor where AdePe is replaced with the Ado
ligand.......................................................................................................57

3.6 Structure of Ado ligand with glycosyl rotation angle $\chi_{CN}=O1'-C1'-N9'-C8'$
exocyclic ribose angle $\Phi_2=C1'-C2'-O2'-O2'H$, and exocyclic ribose angle $\Phi_3=C2'$-
$C3'-O3'-O3'H$..........................................................................................58

3.7 Potential energy curves for the ground (black) and lowest, vertical singlet (red) and
triplet (blue) excited states as functions of Co-C bond length for (a) AdoCbl inside
EAL and (b) Im$-[CoIII(corrin)]^+Ado^+$ base-on model complex in water..........61

3.8 Potential energy curves for the ground and lowest, vertical singlet and triplet
excited states as functions of Co-C bond length for (a) AdoCbl inside EAL and (b)
H$_2$O-[CoIII(corrin)]$^-Ado^+$ base-off model complex.......................................62

3.9 HOMO and LUMO molecular orbitals for the lowest vertical singlet electronic
transitions..............................................................................................64
Potential energy surface of $S_0$ and $S_1$ with vertical projections plotted as a function of axial bond lengths. (a) AdoCbl inside EAL (b) $\text{Im}^-[\text{CoIII(corrin)}]$−Ado$^+$ base-on model complex in water and (c) $\text{H}_2\text{O}^-[\text{CoIII(corrin)}]$−Ado$^+$ base-off model complex in water

(a) Scheme of photoreaction of AdoCbl inside EAL (b) minimum energy paths (Path A and B) between MLCT and LF state plotted as a function of energy versus Q value

HOMO and LUMO molecular orbitals involved in electronic excitations corresponding to selected points on the $S_1$ PES

Potential energy surfaces for the singlet ground $S_0$ and $S_1$ states for RibCbl inside EAL

Potential energy surfaces of $S_1$ state as function of axial bond lengths with photoreaction pathways depicted by red arrows for (a) AdoCbl inside EAL (b) $\text{Im}^-[\text{CoIII(corrin)}]$−Ado base-on model complex in water and (c) $\text{H}_2\text{O}^-[\text{CoIII(corrin)}]$−Ado base-off model complex in water

Mechanism of AdoCbl photolysis in EAL and post-homolysis photophysical events (cage escape, internal conversion), as described in Section 3.3.7

a) Mechanism of Co-C bond cleavage in AdoCbl-dependent GLM. Substrate is $S$-glutamate. The reaction catalyzed by GLM is the conversion of $S$-glutamate to $(2S,3S)$-3-methylaspartate b) Photolytic cleavage of Co-C bond in AdoCbl-dependent GLM without the binding of a substrate

a) Crystal structure of AdoCbl-dependent GLM (PDB ID:1I9C) obtained from the protein data bank. b) Structure of one asymmetric unit which is used for the
calculations. c) Optimized structure of GLM obtained from the QM/MM calculations………………………………………………………………………………………83

4.3 GLM optimized and crystal structure were shown to depict the role of environment upon binding with the enzyme………………………………………………………………………………………………………86

4.4 PEC of ground state $S_0$ as well as low-lying excited states as a function of Co-C bond length for AdoCbl-dependent GLM…………………………………………………………….89

4.5 PES of $S_0$ and $S_1$ plotted as a function of axial ligand bond distances for a) AdoCbl-dependent GLM and b) AdoCbl-dependent EAL………………………………………………….92

4.6 Molecular orbitals of AdoCbl-dependent GLM involved in electronic excitations from $S_0 \rightarrow S_1$ for the selected points on the $S_1$ PES……………………………………………………93

4.7 a) Photodissociation scheme for AdoCbl-dependent GLM based on the $S_1$ PES. minima b) The energetics associated with Path A and Path B were plotted as a function of Q vs energy……………………………………………………………………………….95

4.8 $S_1$ PES as a function of axial bond lengths Co-C and Co-N$_{Im}$ with the photodissociation pathways depicted by red arrows for a) AdoCbl-dependent GLM and b) AdoCbl-dependent EAL………………………………………………………97

4.9 Relative energetics of the MLCT and LF state in AdoCbl cofactor for a) AdoCbl-dependent GLM and b) AdoCbl-dependent EAL along the coordinate Q………98

4.10 The photodissociation mechanism in AdoCbl-dependent GLM………………100

4.11 PES of $S_0$ and $S_1$ electronic states plotted as a function of axial bond lengths along with the contour representation of $S_0$ and $S_1$ states as a function of axial bond lengths to make the comparison of $S_1$ PES and the photodissociation mechanism in different environment………………………………………………………….101
4.12 Energy diagram of frontier KSOs for Im-[CoIV(corrin)]-Ado$^{2+}$ (labeled 1e-Ox) model complex at BP86/TZVPP level of theory………………………………………104

4.13 Isosurface of spin density obtained from BP86/TZVPP calculations for the Im-[CoIV(corrin)]-Ado$^{2+}$ model complex (labeled 1e-Ox). (b) Spin density distribution on fragments of complex based on Mülliken population analysis…105

4.14 Electron density differences between $S_0$ and $S_1$ states for the AdoCbl-dependent GLM at $S_{1(min)}$ MLCT point…………………………………………………………..106

4.15 Mechanism of photodissociation for AdoCbl-dependent GLM compared with TAS experiment……………………………………………………………………..107

5.1 The crystal structure and the optimized structure of EAL depicted in ribbon and ball and stick model……………………………………………………………..120

5.2 a) Bond dissociation energy curve of the Co-C obtained using different level of theory and compared with. b) Spin density distribution of the Co-C bond throughout the cleavage. c) MM energy contributions throughout the cleavage of the Co-C bond……………………………………………………………………..121

5.3 Potential energy curve of EAL along the glycosyl angle ($\chi$) to depict the hydrogen abstraction by Ado radical from the substrate………………………………………..124

5.4 Optimized structures of intermediates in the reaction pathway. a) Optimized structure of diradical state as described [$\text{Ado}^{•} + \text{Co(II)}$] the relative energetics of this intermediates depicted in figure 5.7. b) Optimized structure of intermediate 1d in figure 5.7. This intermediate is the first detectable species by spectroscopy in the EAL native catalytic cycle…………………………………………………………..125
5.5 a) The optimized structure of intermediate as shown in figure 5.7. The notation of this intermediate is [Ade + Prod\textsuperscript{*}]. b) The energy diagram of the hydrogen re-abstraction as a function $\chi$ angle with energy…………………………………126

5.6 The optimized structure of intermediates species in EAL catalytic cycle. a) Optimized structure of diradical after the hydrogen re-abstraction. b) Optimized structure of the final product formed in the EAL catalytic cycle…………………127

5.7 The overall energetics of EAL catalytic cycle. The individual intermediates are depicted by 1a, 1b, etc. The blue color indicates the ONIOM energy whereas the red color is for the QM energy. QM energy is the energy of only the model system which includes the active site of the cofactor………………………………………128

6.1 Potential energy curves (PECs) of lowest singlet state along with the low-lying singlet excited state as a function of Co-C bond length for AdoCbl-dependent EAL without the substrate…………………………………………………………...141

6.2 Potential energy surface along with the vertical projections of $S_0$ and $S_1$ surface for the substrate free model of AdoCbl-dependent EAL…………………………143

6.3 The 2D contour representation of the first singlet excited state ($S_1$). The PES was represented as a function of two axial bond lengths…………………………144

6.4 Active site figures of AdoCbl-dependent EAL. AdoCbl cofactor is depicted as ball and stick model and the protein residue using ribbon. a) Crystal structure of AdoCbl-dependent EAL. b) Optimized structure of AdoCbl-dependent EAL with the substrate molecule…………………………………………………………145

6.5 The 2D contour representation of the PESs. The relative energies were shown above all the figures along with the color code. The relative energy unit is in eV a)
The S₀ PES of the model system without the substrate molecule. b) The S₀ PES of the substrate bound model system.

6.6 The PESs of S₁ and D₁ state along with the contour representation. a) The excited state (S₁) PES of the model system without the substrate molecule.

6.7 The optimized structures of the two minima. a) Optimized structure of AdoCbl-dependent EAL model system with the substrate at π* minimum (I(D₁min), Figure 6.6b) in D₁ PES. b) Optimized structure of the model system in σ* minimum (III(D₁min), Figure 6.6b).

A.1 Structure of Ado ligand. The exocyclic and endocyclic angles were shown in detail. Glycosidic χ angle is defined as O1'-C1'-N9'-C8'.

A.2 Crystal structure of AdoCbl with corrin ring, the side chains were removed, and the axial base is simplified. Crystal structure obtained from Randaccio et al.

A.3 HOMO and LUMO molecular orbitals for selected points in the S₁ PES (Figure 3.11) in addition to the molecular orbitals displayed in figure 3.12.

A.4 The ground state optimized geometries for the selected points in S₁ PES shown in figure 3.11.

A.5 (a) Optimized structure of EAL without the substrate. (b) Optimized structure of GLM with the substrate. The structural analysis is shown in Table 4.1.

A.6 Detail molecular orbital pictures to characterize the several points in the S₁ PES as shown in figure 4.7 in addition to the orbital analysis in figure 4.6.

A.7 Energy diagram of frontier KSO for Im-[CoIV(corrin)]-Ado²⁺ model complex (labeled 1e-OX) at B3LYP/TZVPP level of theory.
A.8  (a) Isosurface of spin density obtained from B3LYP/TZVPP calculations for the Im-[CoIV(corrin)]-Ado²⁺ (labeled 1e-Ox) model complex. (b) Spin density distribution on fragments of complex based on Mülliken population analysis….187

A.9  The active site figure of EAL. Ado ligand is shown in the figure along with the H-bonding network with the surrounding protein residue…………………………187
# LIST OF SCHEMES

<table>
<thead>
<tr>
<th>SCHEME</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>136</td>
</tr>
<tr>
<td>a) The homolytic cleavage of the Co-C bond in presence of a substrate for AdoCbl-dependent enzymes. b) The homolytic cleavage of the Co-C bond to generate Co(II) and Ado radical in presence of light without the binding of a substrate molecule.</td>
<td></td>
</tr>
</tbody>
</table>

| 6.2    | 147  |
| The ground state PES depicted as a function of the Co-C bond and the protein configuration, adopted from Warncke et al. |
CHAPTER I
INTRODUCTION

1.1. Overview of the dissertation

This dissertation is based on the photolysis and the native catalysis study of vitamin B\textsubscript{12} dependent enzymes. Two different B\textsubscript{12} dependent enzymes namely ethanolamine ammonia lyase (EAL) and glutamate mutase (GLM) were studied in this dissertation. Specifically, the mechanism of photolysis and the catalysis were investigated using the combined quantum mechanics/molecular mechanics (QM/MM) method. The plausible connection between photolysis and native catalysis were also studied in this dissertation. Overall this dissertation has seven chapters. Chapter one covers the introduction and the brief history of vitamin B\textsubscript{12} chemistry and enzymology. Since the discovery of vitamin B\textsubscript{12} in 1926, it is an active field of research in chemistry. This was discussed in more detail in chapter one. Chapter two introduces the computational methods that were applied for this study. Specifically, the QM/MM methodology, a brief description of the density functional theory (DFT) method, and time-dependent DFT (TD-DFT) were discussed in this chapter. The light induced reactions of EAL is described in chapter three. GLM, another B\textsubscript{12}-dependent enzyme also undergoes the photolysis which is discussed in chapter four. Overall in these two chapters, it was shown that the mechanism of photolysis in two different classes of AdoCbl-dependent enzymes and their influences
in the mechanism. The mechanism of photodissociation for AdoCbl cofactor in solution
were also discussed and compared with enzymatic environment. Based on these studies, it
appears that the mechanism of photodissociation differs based on the environment of the
cofactor and binding of the axial ligands.

Chapter five includes the native catalysis mechanism of EAL using QM/MM
method. The native catalysis of EAL involves different intermediates in the reaction. Using
QM/MM method different intermediates were optimized and studied. The detail energetics
and structure of the intermediates were analyzed. The factors responsible for the rate
enhancement were analyzed in chapter five. Chapter six includes the connection between
photolysis and native catalysis mechanism based on the idea that in both ways the
generated radical pairs are similar. However, photolysis and native catalysis involves
different electronic states and timescales. The photolysis of the Co-C bond is mediated by
the low-lying electronic excited states which is the basis for the construction of excited
states potential energy surface (PES) to describe the photodissociation mechanism.
However, this cannot be directly used to compare with the native catalysis as the electronic
states involves in both mechanisms are different. In chapter six it was shown that addition
of an electron to the cofactor is required to construct a reduced state PES for the ground
state. The topology of the reduced ground state PES and the excited state PES are similar.
Thus, photolysis and native catalysis can be compared in a way that involves a reduced
state. The possibility of proton-coupled electron transfer (PCET) is described in chapter
six for the reduction of the cofactor. Chapter seven includes the overall summary of the
study and the future directions. The important outcome of this study is that it provides a
basis for the understanding of photochemistry inside enzymatic environment. The native
catalysis mechanism of EAL was investigated and several plausible factors responsible for the unusual reactivity were discussed. Based on the study of catalysis and photolysis, a probable connection between these two mechanisms were established. Overall these studies are the fundamental basis for more detail understanding of B_{12} chemistry and enzymology. These insights can be used to utilize B_{12} chemistry for many different purposes such as drug delivery, biomimetic catalyst design and for many other optogenetic tools.

1.2. Vitamin B_{12} and its derivatives

The derivatives of vitamin B_{12}, otherwise known as cobalamin (Cbl), are nature’s most beautiful cofactor, which participates in numerous enzymatic reactions in both bacteria and mammals.\textsuperscript{1-5} Two important physiological reactions catalyzed by Cbl derivatives in mammals are methyl (CH\textsubscript{3}) group migration through methylcobalamin (CH\textsubscript{3}Cbl) in methionine synthase (MetH) and the H/X exchange in -CH-CX linkage of organic compounds through adenosylcobalamin (AdoCbl) in methyl-malonyl CoA mutase (MCM).\textsuperscript{5-9} Although it appears that the derivatives of vitamin B_{12} play an important physiological function for both prokaryotes and eukaryotes, only prokaryotes can synthesize it directly in their metabolism.\textsuperscript{10,11} The total synthesis of Cbl occurs in a complex pathway that requires approximately thirty enzymatic steps for prokaryotes.\textsuperscript{11} A higher organism that requires Cbls uptake it from food source through a unique evolutionary mechanism or from a natural symbiotic relationship. Mammalians have a complicated Cbl uptake route in their metabolism.\textsuperscript{10,12} Upon intake, Cbls have to be resorbed from the source and transformed physiologically into biologically active vitamin B_{12} derivatives and then actively transported to the cells.\textsuperscript{3,10,13} Three different proteins play a role in this total
transportation such as haptocorrin (HC), transcobalamin (TC), and intrinsic factor (IC). Crystal structures of these transport proteins have been investigated in detail and discussed somewhere else. In short, B\textsubscript{12} derivatives from the food source form a complex with HC in saliva under acidic conditions. This HC-Cbl complex then migrates to the duodenum, where the pancreatic enzymes release the Cbls. Later on, Cbls binds with the IF to facilitate transport to TC to form a TC-Cbl complex. The TC-Cbl complex carries Cbls into the cell where lysosomes degrade the complex and releases Cbl to be converted into biologically active vitamin B\textsubscript{12} derivatives such as AdoCbl and CH\textsubscript{3}Cbl. The deficiency of B\textsubscript{12} derivatives in these metabolic processes may result from an unbalanced diet or from genetic disorders, which leads to severe neuropsychiatric and hematological problems such as pernicious anemia.

The discovery of vitamin B\textsubscript{12} occurred in a fortuitous event by three notable scientists, namely Whipple, Minot, and Murphy in 1926, while treating pernicious anemia patients with raw liver extracts. This significant discovery brought enormous success in treating the pernicious anemia and heralded the Nobel prize in physiology in 1934. These discoveries were the beginning era of B\textsubscript{12} chemistry since then it has been flourished with many notable contributions. Two decades later, of this pioneering work, the isolation and the crystallization of vitamin B\textsubscript{12} cofactor was achieved by Folkers, Sharp, Dohme, and Smith in 1948. It was shown that the isolated compound contained Co atoms and has remarkably increased the response towards pernicious anemia with limited dosages such as 4 \( \mu \)g of vitamin B\textsubscript{12}. This molecule later named as cyanocobalamin, but the detailed structural formula of this molecule was still unknown. In 1954, a famous English scientist named Dorothy Hodgkin resolved the complex multiring structure of vitamin B\textsubscript{12}
which is still considered as a breakthrough in $\text{B}_{12}$ chemistry.\textsuperscript{31} She was awarded Nobel prize in chemistry in 1964 due to this pioneering work on the $\text{B}_{12}$ crystal structure.\textsuperscript{32}

### 1.3. Structural model of vitamin $\text{B}_{12}$ derivatives

The structure of Cbls consists of a negatively charged, nineteen carbon corrin ring, which is equatorially coordinated with octahedral cobalt (Co) ion in the center of the corrin ring.\textsuperscript{13,31,33-36} The structure of corrin resembles porphyrin, as both belong to the family of tetrapyrrolic macrocycles, with two noticeable exceptions; in former one-carbon (C$_{\text{meso}}$) is missing and corrin ring is more saturated. Consequently, corrin is “less aromatic” and slightly non-planar in comparison to porphyrin macrocycle and contains four partially saturated pyrroles.\textsuperscript{37-39} The central Co ion in Cbl is always coordinated equatorially by these four pyrrolic nitrogen atoms regardless of its oxidation state.\textsuperscript{36} In Co(III) oxidation state, the Co ion is hexacoordinated with four innermost nitrogen’s along with two axial ligands. The upper axial position is commonly known as the $\beta$ axial site, whereas the lower axial position generally referred to as $\alpha$ axial site (Figure 1.1). Upon the conversion of the Co(III) to Co(II), the $\beta$ axial ligand is lost, and the Co ion adopts $d^7$ configuration with an unpaired electron in the $d_{z^2}$ orbital.\textsuperscript{5,13,36,40} Another important structural feature of Cbls includes several amide groups project upward and downward from the corrin ring, which is known as side chains a-f, shown in figure 1.1 where one sidechain known as f is connected to a nucleotide through an amide bond.
Figure 1.1. a) The molecular structure of Cbl and atomic numbering of corrin ring. b) The molecular formula of the upper axial ligand, c) The insertion denotes the LA and SA polarization axes of ππ* transitions of corrin ring and the fat line indicates the conjugated -C=N=C- bonds.

The dimethyl-benzimidazole (DBI) base from the nucleotide loop acts as lower axial ligand and coordinates with the central Co ion from the α axial position of the corrin macrocycle (Figure 1.1). A structural analog of this kind is known as base-on Cbl. In strongly acidic medium, the DBI base detached from the lower axial position and form base-off Cbl (Figure 1.2b).
Figure 1.2. Structural analog of Cbls. a) Base-on form of Cbl where the lower axial ligand is a dimethyl-benzimidazole. b) Base-off Cbl where the lower axial ligand is Im moiety. This structural analog represents the base-off/his-on form found in some AdoCbl and CH$_3$Cbl dependent enzymes.

The upper axial ligand varies in different Cbls. Some of the commonly known upper axial ligand in Cbls is cyanide (CN$^-$), methyl (CH$_3$), 5$'$-deoxy 5$'$-adenosyl (Ado), water (H$_2$O), hydroxo (OH$^-$), sulphito (SO$_2^-$), etc.$^{1,2,34-36,40,44-46}$ The nature and the naming of the Cbl is mostly dependent on the upper axial ligand.$^{1,10,34-36,47,48}$ For instance, if the upper axial ligand is CN$^-$ then it is called as cyanocobalamin (CNCbl), which is also known as vitamin B$_{12}$. An interesting fact about the CNCbl is that it is biologically inactive, but it has two biologically active derivatives.$^{26}$ In enzymatic reactions, the upper axial ligand is replaced by an Ado group, yielding an AdoCbl, or by a CH$_3$ group, yielding a CH$_3$Cbl, while DBI retains its coordination or replaced by another ligand (Figure 1.3).$^{49-51}$
These two are the biologically active B$_{12}$ derivatives that act as a cofactor in numerous enzymatic reactions.$^{3-7,10,49,52-61}$

![Figure 1.3. Structural model of vitamin B$_{12}$ analog.](image)

The crystal structure of these two essential metabolites was first resolved in 1961 for AdoCbl and 1985 for CH$_3$Cbl.$^{62,63}$ The overall structure of these two cofactors is quite similar, except for a few subtle differences in the corrin ring puckering and the axial ligands. In the case of CH$_3$Cbl, the fold angle of corrin is 14.8°, which is 10° in AdoCbl.$^{62,63}$ In the case of lower axial position, both cofactors constitute DBI with slightly different pKa values for protonation. In CH$_3$Cbl, the pKa value for the protonation of the DBI base is 2.9, which is 3.7 for AdoCbl.$^{1,2}$ The most striking similarity between these two biological cofactors is an organometallic Co-C σ bond, which is the first example of the metal-carbon bond that plays a vital role in enzymatic reactions.$^{1,64-66}$ The cleavage of this bond constitutes the critical step in B$_{12}$ enzymatic catalysis.
1.4. Vitamin B12-dependent enzymes

Since the discovery of vitamin B12, it has been a subject of extensive research and development in chemistry and biochemistry. However, the exact biochemical role of the cofactor was not clear for a long time. In 1958 renowned scientist H.A. Barker first discovered AdoCbl-dependent enzyme glutamate mutase (GLM).\textsuperscript{67,68} It was shown that GLM catalyzes the conversion of L-glutamate to L-threo-3-methyl aspartate during the fermentation of glutamate by \textit{clostridium teteanomorphum}.\textsuperscript{67} The essential catalytic step in this reaction is the homolytic cleavage of the Co-C σ bond to generate the Ado radical. The Ado radical formed by the homolytic cleavage is very reactive and then undergo subsequent steps. Because of this substantial reactivity of this radical, AdoCbl-dependent enzymes exert strict control on the radical intermediate to prevent unwanted side reactions.\textsuperscript{7,9,49,54,61,69} This kind of enzymatic reaction utilizing radical chemistry is relatively rare, and all AdoCbl-dependent enzymes known to date regarded as a primary example of utilizing this radical chemistry.\textsuperscript{3,4,6,60,61} On the other hand, another crucial vitamin B12 derivative CH3Cbl, undergoes heterolytic cleavage of the Co – C bond to catalyze the transfer of a CH3 group from a CH3 donor to a CH3 acceptor.\textsuperscript{57,58}

Depending on the type of reactions they catalyze, all B12-dependent enzymes can be classified into three main classes.\textsuperscript{3,7,70} First, CH3Cbl-dependent methyltransferases (MeT) that catalyze CH3 transfer reactions,\textsuperscript{57-59} second AdoCbl-dependent isomerases,\textsuperscript{5,53,71,72} which catalyze 1,2 rearrangement reactions and reductive dehalogenases (RDase)\textsuperscript{59,73,74} which utilizes Cbl cofactors to catalyze the dehalogenation reactions. In the case of MeT, cofactor CH3Cbl bind in base-off mode with the enzyme where the lower axial ligand DBI is replaced with the Histidine residue commonly known
as base-off/his-on mode and catalyze many critical biochemical reactions (Figure 1.2b). Some of the reactions catalyzed by MeT utilizing CH₃Cbl are methionine biosynthesis in both mammals and bacteria, carbon dioxide fixing pathway in anaerobic acetogenic bacteria, and methanogenic archaea. Among different CH₃Cbl-dependent MeT, methionine synthase (MetH) is the most well-studied enzyme that is found in mammals and bacteria. MetH catalyzes the transfer of a CH₃ group from methyltetrahydrofolate (CH₃-H₄folate), to cob(I)alamin, then to the thiolate of homocysteine (Hcy) to generate H₄folate and methionine, an essential amino acid. The initial step of this reaction is the heterolytic cleavage of the organometallic Co-C bond, which is enhanced 10⁵-fold by MetH. Several studies have been conducted to understand the detailed mechanism of this enzymatic catalysis. However, there is no unanimous consensus about the mechanism.

AdoCbl-dependent enzymes catalyze the 1,2-rearrangement or elimination reactions between a hydrogen (H) atom and an electronegative group from an adjacent carbon atom. To date, twelve AdoCbl-dependent enzymes have been identified, and the crystal structure was resolved. These enzymes employ the AdoCbl cofactor as a source of organic radicals to catalyze the reactions. Based on the three-dimensional structure of these enzymes, it appears that the binding of the protein with the AdoCbl cofactor occurs in two ways, base-on mode and the base-off mode (Figure 1.2). In the case of base-on mode the lower axial DBI ligand retains its position, whereas in base-off mode the DBI ligand is replaced by a histidine residue known as base-off/his-on binding mode (Figure 1.2b).
Based on the binding of the cofactor with the proteins and the type of reactions they
catalyze, the AdoCbl-dependent enzymes can be classified into two main groups, mutases,
and eliminases.\textsuperscript{5,70,72} The eliminases group participates in 1,2- rearrangement reactions of
1,2-diols or 1,2-amino alcohols to 1,1-diols or 1,1-amino alcohols, which undergoes further
elimination of water to generate aldehyde and ammonia as the final products (Figure 1.4).\textsuperscript{72}
This type of reaction is significant for many industrial purposes. The cofactor binds with
the enzyme in base-on/his-off mode, and the distance between the substrate 1,2-diols or
1,2-amino alcohols and the cofactor is longer compare to other mutase class of AdoCbl-
dependent enzymes.\textsuperscript{53,71,80-82} Members of this group include ethanolamine ammonia-lyase
(EAL), diol dehydratase (DDH), glycerol dehydratase (GDH), which are bacterial
enzymes.\textsuperscript{53,71,83,84} In mutases, the cofactor binds in base-off/his-on mode and can be further
divided into two groups, carbon skeletal mutase, and aminomutases.\textsuperscript{3,6,70} Carbon skeletal
mutase participates in the reactions which involve the interchange of an H-atom from one
carbon atom to another carbon atom. The members of this subgroup are glutamate mutase
(GLM), methyl-malonyl-CoA mutase (MCM), methyleneglutarate mutase (MGM),
isobutyrl-CoA mutase (ICM) etc.\textsuperscript{70,72,80,81} Among these enzymes, the most well studied is
MCM, which is found in both mammals and bacteria and also plays an essential biophysical
role in catalyzing the methylmalonyl-CoA to succinyl-CoA (Figure 1.4).\textsuperscript{80} On the other
hand, the aminomutase enzyme catalyzes 1,2 migration of amino groups to the adjacent
atoms. In those cases, there is an additional pyridoxal phosphate molecule acts as a cofactor
along with AdoCbl. The two-known enzyme of this group is ornithine aminomutase
(OAM) and d-lysine aminomutase (5,6-LAM).\textsuperscript{82,85,86}
Figure 1.4. Reactions catalyzed by Cbl-dependent enzymes. Different color schemes indicate different enzymatic reactions. Among these enzymes, EAL, GLM and MCM are AdoCbl-dependent enzymes and MetH is CH$_3$Cbl-dependent enzyme. In addition, EAL and GLM are bacterial enzymes and MCM and MetH found in both bacteria and mammals.
1.5. Native catalysis mechanism in AdoCbl-dependent enzymes

The most striking similarities in all AdoCbl-dependent enzymes are the initial homolytic cleavage of the Co-C bond to generate the Co(II) and Ado radical pair.\(^4,10,87\) The generated Ado radical is a very reactive species and instantly abstracts an H-atom from the substrate to generate a substrate derived radical and 5′-deoxyadenosine. The substrate radical then undergoes rearrangement to generate the rearranged substrate radical, otherwise known as product radical. The product radical re-abstract the H-atom from 5′-deoxyadenosine, which leads back to the formation of Ado radical and the final product. The Ado radical will then recombine back with Co(II) and complete the catalytic cycle.\(^3,6,7,10,14,55,88,89\) The detail catalytic mechanism is shown in figure 1.5. This catalytic mechanism has been known to be followed in all AdoCbl-dependent enzymes and seems beautiful and straightforward at initial observation. An interesting fact about this seemingly simple mechanism is that the essential details of the total mechanism are mostly unknown. For instance, it has been known for a long time that the isolated AdoCbl cofactor undergoes thermolysis to homolytically cleave the Co-C bond to generate Co(II) and Ado radical.\(^1,65,66,90,91\) The bond dissociation energy (BDE) of the Co-C bond at room temperature is 32 kcal/mol, and the rate constant for this uncatalyzed thermally induced reaction is \(3.8 \times 10^{-9}\) S\(^{-1}\).\(^65,66,87,91\) The half-life of this homolysis is estimated to be six months, which appears that at the uncatalyzed condition, the reaction is significantly prolonged.\(^66,87\)
Figure 1.5. Native catalytic cycle of AdoCbl-dependent enzymes. Intermediate IV is the product that will undergo further elimination or migration process to yield the final product.

In the case of an enzyme, the cofactor AdoCbl and the substrate binds with the protein, and the interaction between the protein and the cofactor leads to the dissociation of Co-C bond to generate Co(II) and Ado radical.\textsuperscript{1,10,92} In the absence of a substrate, this reaction is slower, and only a small fraction of the cofactor undergoes Co-C bond dissociation. However, upon binding with the substrate, the cleavage of the cofactor’s Co-C bond occurs in a millisecond time scale, and the internal equilibrium constant for homolysis is nearly one.\textsuperscript{87,91} The rate constant for the cleavage of the Co-C bond is estimated to be (10-100) S\textsuperscript{-1}. Thus, it appears that in the enzymatic environment, the cleavage of the Co-C bond is accelerated by \(10^{12}\)-fold compared to the isolated cofactor.\textsuperscript{65,87,91,93} This mechanism is followed in all the AdoCbl-dependent enzymes such as MCM, EAL, GLM, etc and suggests that the enzymatic environment plays a significant
role in this unusual activation. Several hypotheses have been proposed in order to explain this unusual reactivity. Such as the cleavage of the Co-C bond is not the rate-determining step, thus to achieve such reactivity, the BDE should be reduced by 50% inside the enzymatic environment. Another critical hypothesis considered was large-scale conformational change induced by substrate binding, which ultimately weakens the Co-C bond. However, the attempt to verify this hypothesis was regarded as barely successful and also not widely accepted for many AdoCbl-dependent enzymes. Based on the resolved crystal structure of AdoCbl enzymes, it appears that there is no significant structural distortion of the protein upon substrate binding. Similarly, using many other experimental techniques such as resonance Raman, UV-visible spectroscopy, and magnetic circular dichroism it was shown that the AdoCbl structure is not perturbed upon the binding of an enzyme. This issue remains poorly understood and needs further investigation.

Another critical aspect of the catalytic cycle is the H-atom abstraction steps by the Ado radical (Figure 1.5). As stated earlier, Ado radical is a very reactive species and has not been observed before, so a detailed analysis is also not properly understood. After the generation of Ado radical, it instantly abstracts an H-atom from the substrate and forms substrate derived radical and 5'-deoxyadenosine(Figure 1.5). Based on electron paramagnetic resonance (EPR) it was shown that substrate radical and Co(II) is the first detectable species in the enzymatic catalysis. Besides, it was also shown that Co-C bond cleavage and the H-atom abstraction steps were coupled based on kinetic isotopic effect (KIE) observation through pre-steady state stopped-flow experiments. KIE is a powerful tool to provide a detail importation about complex chemical reactions by
identifying the rate-limiting steps through the comparison of reaction rate between unlabeled and isotopically labeled molecules. In the case of GLM and MCM, the substrate was deuterated, and then the Co-C bond cleavage was observed. Based on the UV-visible spectral changes throughout the reaction, it was shown that the rate is significantly lower in the deuterated sample, which indicates the H-atom abstraction steps as the rate-limiting step in the reaction that confirms the kinetic coupling between two steps.7,9,88,106

A concerted or stepwise mechanism for the Co-C bond cleavage and the H-atom abstraction is also another debated issue in B\textsubscript{12} chemistry (Figure 1.6). Based on the computational study conducted for the GLM model system, it was shown that the concerted mechanism is energetically favorable over the stepwise mechanism by 4.0 kcal/mol.107 Similarly, another computational study based on empirical valence bond (EVB) theory also supports the concerted mechanism in the MCM enzyme.108,109 In contrast to these, there are also several other studies conducted on MCM, OAM, which supported the stepwise mechanism.61,85,110,111 A critical point to consider in certain AdoCbl enzymes such as eliminases the distance between the substrate and the cofactor is longer.53,84,97 This indicates that the generated Ado radical has to migrate a long distance to abstract the H-atom from the substrate. Thus, it indicates that in the case of a concerted mechanism, a significant reorganization of protein is required to facilitate the migration of Ado, which is unlikely based on the available crystal structure data.53,95
Figure 1.6. Proposed mechanism for the Co-C bond cleavage in AdoCbl-dependent enzymes.

Nevertheless, other theories were suggested based on experimental and computational studies. For example, upon radical generation through homolytic cleavage, Ado undergoes specific conformational changes, such as pseudo-rotation of the glycosyl moiety, to migrate a long distance. In the case of GLM, it was suggested that ribose moiety undergoes puckering, such as before hydrogen abstraction, the conformation of Ado was shown to be 2'-endo whereas after hydrogen abstraction, it changes to 3'-endo (Figure 1.7).\textsuperscript{81,112} The information about 2'-endo and 3'-endo is shown in figure A1. In MCM, there is an adenine ribose rotation occurs to facilitate the hydrogen abstraction.\textsuperscript{113} Based on MCM crystal structure, it appears that when Ado bound to Cbl, it adopts syn conformation, and after the cleavage of the Co-C bond, it adopts anti conformation.\textsuperscript{113} Despite this many valuable hypotheses, there is still a lacking’s in formulating a clear picture to understand
the mechanism of AdoCbl-dependent enzymes and requires further experimental and computational investigations.

![Figure 1.7](image)

**Figure 1.7.** Different orientation of ribose ring in the AdoCbl cofactor found in AdoCbl-dependent enzymatic catalysis.

**1.6. Photolysis of Co-C bond in AdoCbl-dependent enzymes**

Although the known biological function of Cbl is to act as a cofactor in thermally driven enzymes, it also exhibits a rich and complex photochemistry.\(^{1,64,114-125}\) Like many other biomolecules such as porphyrins, chlorophyll, flavins, Cbls also respond to light and participate in diverse chemical reactions.\(^{11,50,126-133}\) The light sensitivity of Cbls was first
reported in the 1950s, but for many years it was considered as an unwanted side reaction.\textsuperscript{67,134} In a study conducted by Barker et al. in 1958 showed, that the pseudo vitamin B\textsubscript{12} undergoes decomposition through light irradiation.\textsuperscript{67} This was considered as irritation of coenzyme B\textsubscript{12} by light. Later on, in 1960, a more detail study was conducted by the same group of scientists to investigate the photochemical cleavage and the photolysis products.\textsuperscript{68,114,135} In another study conducted by Johnson, Hogenkamp, Schrauzer, and others in 1960 showed that the fate of the photolysis of coenzyme B\textsubscript{12} depends on the presence of O\textsubscript{2}.\textsuperscript{64,115,136-139} The photochemical study of Cbls has continued to gain interest among the scientific community since these critical contributions. Different research groups made significant contributions to understand various aspects of Cbls photochemistry. In the 1960s the photochemical reactions of alkylCbls and non-alkylCbls were studied by Dolphin, Pratt, and co-workers.\textsuperscript{90,140,141} It was shown that the non-alkylCbls were insensitive to light, whereas the alkylCbls are photoactive.

One of the first theoretical efforts to understand the optical and electronic properties of vitamin B\textsubscript{12} derivatives was first reported by Day in 1967 using Huckel and Pariser-Parr-Pople methods.\textsuperscript{142,143} The primary outcome of these studies was the quantitative explanation of the absorption bands such as α, β, γ for different types of Cbls. Schrauzer conducted another critical theoretical study of Cbls in the 1970s using a self-consistent HMO method.\textsuperscript{115} It was shown that the photolysis of the Co-C bond is induced by charge transfer (CT) from a bonding axial molecular orbital to the antibonding axial orbital. The extent of the photo resistivity seems to increase with the increase of the Co-C charge transfer extent. Lately, the study of light sensitivity increases exponentially with the advancement of many spectroscopic techniques. Steady-state absorption (Abs)
spectroscopy, electron paramagnetic resonance (EPR), resonance Raman (rR), circular dichroism (CD), magnetic circular dichroism (MCD), and most recently transient absorption spectroscopy (TAS) were quite frequently used to understand the nature of the electronically excited states of Cbls.

In the 1990s, Chen and Chance used modern spectroscopic techniques such as TAS to investigate the electronically excited states of Cbls in nanosecond time scale.\textsuperscript{116,117} It was shown that the photochemical and photophysical properties of Cbls are dependent on the binding of the axial ligands. The significant outcome of these studies was a proposed scheme to understand the photochemical and photophysical properties of Cbls clearly. The effect of the magnetic field in the photolytic properties of Cbls was also an essential factor. In the 1990’s Grissom and co-workers showed that the photolytic cleavage of the Co-C bond varies with the varying magnitude of the magnetic field.\textsuperscript{144-147} The recombination rate of the radical pairs increases by applying hundreds of milli tors of the magnetic field which decreases the quantum yield of photolysis. This is due to the Zeeman effect interacting on the geminate radical pairs. Similarly, another group conducted a more interesting study of the magnetic field effect in AdoCbl-dependent enzymes, EAL, which was the first known study of magnetic field effect in the biological system.\textsuperscript{148}

The photochemistry of Cbls has gained much more interest than ever before at the beginning of the twenty-first century. There are several reasons for this, such as the advancement of ultrafast TAS techniques, along with the improvement of computational power to investigate many complex reactions. To that end, many research groups such as Sension and co-workers, Scrutton and co-workers, Jones and co-workers, Warncke and co-workers, and our group and others actively involved in elucidating the details regarding the
photolysis of Cbls in solutions as well as inside enzymes.\textsuperscript{42,82,100,123-125,148-176} Based on these studies, it was shown that the photochemical and photophysical properties of Cbls depend on the nature of the axial ligands and the environment of the cofactor.\textsuperscript{42,50,124,176} For instance, the alkylCbls such as CH\textsubscript{3}Cbl and AdoCbl are photoactive whereas non-alkylCbls such as CNCbl are photostable.\textsuperscript{124,158} The photolysis of the Co-C bond was also shown to be wavelength dependent for CH\textsubscript{3}Cbl.\textsuperscript{120} Such as at 400 nm excitation, the quantum yield of the radical pair is higher, whereas, at 520 nm, it is lower. In the case of AdoCbl, the photolysis quantum yields are unaffected by the excitation wavelength.\textsuperscript{119} In both AdoCbl and CH\textsubscript{3}Cbl, the environment of the cofactor was also shown to play a critical role.\textsuperscript{118,152,155,177-180} These studies laid the firm foundation for a more detail investigation of the photolysis mechanism. Based on these studies, it appears that the low-lying electronic states are very crucial to understand the details of Co-C bond photolysis.

Upon excitation, with light, the biologically active forms of Cbls undergoes homolytic cleavage of their organometallic Co-C bond to generate singlet born alkyl/Co(II) radical pairs. In isolated AdoCbl cofactor, the light excitation in the 500-330 nm range cleaves the Co-C bond with near unit quantum yield within 100 ps.\textsuperscript{119,121,122} There is a rapid relaxation then from the excited state to a meta-state which resembles a species similar to the Co(II)alamin without the lower axial ligand binding. This is caged by the solvent molecules which prevents the free radical escape. The metastable Co(II) then reattach with lower axial DBI base and the Ado radical to return to the ground state known as geminate recombination or can recombine with DBI and escape the solvent cage to dissociate as free radicals Co(II) and Ado.\textsuperscript{119,153,156} The quantum yield of the photolysis is dictated by the
competition between geminate recombination and radical escape.\textsuperscript{153,156} The scheme of this photodissociation mechanism is shown in figure 1.8.

![Figure 1.8](image_url)

**Figure 1.8.** Scheme of photolysis for AdoCbl in solution and its dependent enzymes.

The kinetics of the geminate recombination is also very important, which is dependent upon several factors such as the geometry of the upper axial ligand, the presence of a lower axial base, the effect of the magnetic field and the cofactor environment. For example, in Ado radical, the spin density is mostly located on the 5'-carbon which points towards Co(II) and favors recombination. On the other, the in CH\(_3\) radical is planar which points the spin density from the opposite face of the Co(II). This ultimately decreases the geminate recombination.\textsuperscript{124,150,151,181} The photolysis of base-off AdoCbl is significantly lower compared to the base-on AdoCbl cofactor. The absence of a lower axial base opens up a channel for a fast nonradiative decay to the ground state. The rate of geminate recombination also impacted in base-off AdoCbl. In the base-off AdoCbl cofactor, the
quantum yield of the solvent separated radical pairs is in the range of 0.045, whereas in the case of base-on AdoCbl, it is in the range of 0.20-0.24.\textsuperscript{43,122,157,176}

As discussed before, the environment of the cofactor also plays a role in the rate of geminate recombination as well as the mechanism of the photodissociation. Inside the enzyme, the AdoCbl also undergoes photolysis to generate the same radical pairs [Co(II) and Ado] as in the case of isolated cofactor with lower photolysis yield.\textsuperscript{152,177} Of particular importance is that the mechanism of photodissociation in enzymatic environment follows similar trends as in the case of the solution environment. However, several things to be noted, such as in the case of AdoCbl dependent EAL and GLM, it was shown that the photolysis of the Co-C bond also generates similar radical pairs, without the subsequent coupling of the H-atom abstraction steps.\textsuperscript{166,177,178} As a matter of fact, AdoCbl photolysis has proven to be a useful tool to probe the influence of the cofactor environment in radical pair generation and their subsequent relaxation dynamics. More recently, it was shown that the photolysis of Co-C bond in AdoCbl based photoreceptor known as CarH is involved with large scale conformational changes.\textsuperscript{11,132,182,183} CarH is a bacterial enzyme that mediates light-dependent gene regulation in photoprotective cellular responses. At dark intact AdoCbl cofactor binds with the enzymes to form a tetramer, which can attach with DNA to block the transcription. Upon shinning the light, the Co-C bond of AdoCbl cleaves to generate the radical pairs, which then break the tetramer to monomer and detach from DNA. This mechanism is shown in figure 1.9. Thus, the photolysis of AdoCbl in biological medium plays a crucial role and can be harnessed for many different purposes.
Figure 1.9. Photochemical mechanism of Co-C bond cleavage in AdoCbl-dependent CarH. Different steps were shown in the above figure. In the case of A, the cofactor is not bound with the CarH. At dark, the cofactor and enzyme bound together to attach with DNA, which is shown in B. In that case, shining with light cleave the Co-C bond and generate the radical pairs. The detachment from DNA then follows this, this is shown in C. Different types of intermediated such as AdoCbl and the Co(II) are shown in D.

These underlying photochemical processes in AdoCbl biological systems are controlled by the relaxation dynamics that occur on picosecond (ps) timescales and mediates by the low-lying electronic states. Thus, the complexity of these systems is many folds, which makes them challenging to study. To this date, various experimental techniques have been used to study the photochemical and photophysical properties of Cbls. Ultrafast TAS techniques have evolved to be the best reliable method to study the Cbls photochemical properties. The TAS spectrum probes the electronic states of Cbls in
a conformation space, which appears to be the dark state and unobserved using the linear absorption spectrum. Thus, using TAS complements the study from many other experimental techniques such as resonance Raman, steady-state absorption spectroscopy and circular dichroism. The major outcome based on these studies is that modulation of the axial ligation, and the presence of different environments, controls the dynamics of low-lying first singlet excited state (S₁), which dictates the underlying photophysical properties. These issues need to be investigated in more detail using state of the art theoretical calculations. Theoretical calculations can provide insights about the energy cascades of different metastable states, the importance of the charge transfer states, and the effect of the environment in the photochemical mechanism.

1.7. Motivation towards understanding catalysis and photolysis

This study involves quantum mechanics (QM), molecular dynamics (MD), combined quantum mechanics/molecular mechanics (QM/MM) study in biologically active vitamin B₁₂ cofactors, AdoCbl, and its dependent enzymes GLM and EAL, to understand their catalytic and photolytic properties.¹⁷⁸,¹⁷⁹,¹⁸⁴ The use of computational chemistry methods, in that case, is very crucial. Using different computational methods can give detail mechanistic ideas as well as the energetics of multiple intermediates, unstable radicals, meta-stable states. These features are very significant to investigate for many different purposes and also very complicated to understand by only experimental investigation. For instance, AdoCbl-dependent enzymes employ radical chemistry to catalyze chemically essential reactions such as dehydration of glycerol, conversion of alcohol to aldehyde.⁴⁹,¹⁸⁵ Glycerol is a significant by-product of biofuel conversion. With increasing concern about climate change, the production of biofuels has increased which
may lead to an oversupply of glycerol. In that case, the conversion of glycerol to its many viable precursors is paramount. In recent years the conversion of glycerol to 1,3-propanediol has received significant attention.\textsuperscript{186} AdoCbl dependent enzymes employ radical based chemistry to catalyze this type of reaction.\textsuperscript{72} The most critical step in these catalytic reactions is to reduce the activation energy to a certain degree and to increase the rate of reactions enormously to facilitate the generation of radical pairs. Upon radical pair generation enzymatic environment, exert strict control over the radicals to catalyze these chemically important reactions. Owing to this, an enormous amount of enzyme can be used as a green catalyst in the near future. Thus, the artificial design of enzymes can provide a viable solution that requires detailed understanding of the existing enzymatic systems such as coenzyme B\textsubscript{12}-dependent enzymes.

The photochemistry of Cbls is also a vital topic to understand as it has the potential to utilize for many therapeutic applications. Many biological systems respond to light and perform sophisticated light-induced reactions.\textsuperscript{127,128} One of the challenges, in that case, is to design correct light-activated groups that fall within the range of tissue’s optical window, which lies in the region of 600-900 nm.\textsuperscript{15} Less than 600 nm wavelength light is absorbed by the hemoglobin and melanin of the body whereas more than 900 nm wavelength light is interfered by water.\textsuperscript{187} To date, commonly known prodrugs respond to light irradiation at a wavelength less than 450 nm which makes it challenging for practical applications in photodynamic therapy and light-activated drug delivery.\textsuperscript{188,189} In biologically active Cbls, the photo lability is dependent on the upper axial ligand. Different drug molecules can be attached with the upper axial ligand and can be activated by light to cleave the Co-C bond. The fundamental challenge in those cases is to find a viable candidate that can be activated
by light within the tissue's optical window. This issue can be addressed by attaching fluorophore molecules with Cbls. The fluorophores molecule can act as an antenna to capture the long-wavelength light such as 600-900 nm and transmit the energy to Cbls to activate the Co-C bond and hence liberate the drug.\textsuperscript{189} This can be a potential therapy for a variety of disorders such as cancer, autoimmune and vascular diseases and requires further investigation. A Molecular-level understanding of Cbls photochemistry is required to gain deep insight about these issues.

In recent years the photobiology of B\textsubscript{12} has also gained significant interest. Though the photolysis of Co-C in EAL and GLM was shown to play no significant physiological role, it can be used to probe the native catalysis.\textsuperscript{184} Several studies have been conducted to elucidate these details and require further investigation.\textsuperscript{168,184} On the other hand, understanding the photobiology of AdoCbl is also very crucial. AdoCbl acts as a cofactor in certain enzymes that uses the photolysis of the Co-C bond for gene regulation. These special enzymes, known as a photoreceptor, senses light to alters the cellular response. One of its kind is called CarH that binds with AdoCbl in dark state, attaches with DNA, and utilizes the photolysis of its organometallic Co-C bond. After photolysis of the Co-C bond, there is a large-scale conformational change occurs in the enzyme which is essential for DNA transcription.\textsuperscript{133} The details is shown in figure 1.9. Thus, AdoCbl based photoreceptors unwind the new field of vitamin B\textsubscript{12} photobiology, which utilizes a different photochemical mechanism of Cbls. As a result, it is crucial to understand the photochemistry of AdoCbl in different mediums at a molecular level. Here in this study, the photochemical mechanism of AdoCbl-dependent enzymes EAL and GLM, the native catalysis mechanism of EAL and the connection between catalysis and photolysis will be
discussed using methods of computational chemistry. Different levels of theories were applied including, QM, MD, and QM/MM to understand the properties of these two enzymes. These two enzymes EAL and GLM, were chosen for this study for several reasons. First, EAL and GLM represent two major classes of AdoCbl-dependent enzymes. Understanding the mechanism for these two enzymes will help to get a clear insight into many other AdoCbl-dependent enzymes. Second, crystal structures of EAL and GLM were available in the protein data bank (PDB) repository, which is the basis for our studies. Most importantly, the photochemistry of EAL and GLM were studied using TAS which can be compared with the computational data to get more detail insights.
CHAPTER II

COMPUTATIONAL METHODOLOGY

2.1. Background

The use of computational methods has become an invaluable tool in modern-day science. It has been widely used to validate the experimental phenomena such as reactivity, molecular geometries, and spectra with the help of arbitrary models to provide new insights in the field of physics, chemistry, and biology.\textsuperscript{190} There are different methods in computational chemistry based on a different level of theories. Such as MD simulation is widely used for a complex biological system, which is based on classical mechanical force fields to describe the molecular vibrations, torsions, van der Waals and electrostatic interactions. These force fields are unable to describe the changes that occur in the electronic structure of a system such as bond breaking and formation, electron transfer, or electronic excitations. On the other hand, QM methods based on the approximate solutions of the Schrodinger equation allows one to calculate the electronic properties of a system; however, it has some limitations to consider while applying for biological systems such as protein, enzymes, etc.\textsuperscript{190-194} Enzymes consist of thousands of atoms and different conformations to consider and thus challenging to model using the QM method alone. In 1976, Ariel Warshel and Levit accomplished this by combining the QM and MM method which is known as combined QM/MM methods that allows one to calculate the electronic properties in complex environments.\textsuperscript{193} Here in this section, a brief overview of the QM/MM methodology and the procedure that was taken to prepare the QM/MM
inputs for the ground and excited-state calculations of AdoCbl-dependent enzymes will be discussed.

2.2. QM/MM theory

In order to establish a basic understanding of the QM/MM method, it is necessary to look at the foundation of computational chemistry based on the underlying principle of quantum mechanics. The quantum nature of the particles can be described using the complex Schrödinger equation which is the basis of quantum mechanics and gives a complete description of the electronic structure of the molecule.\textsuperscript{190,195,196} The formula of the simplest non-relativistic time-independent Schrodinger equation is,

\[ \hat{H}\psi = E\psi \]  (2.1)

In the above equation, \( \hat{H} \), is the Hamiltonian operator which has kinetic and potential energy terms. \( \hat{H} \) can be written as \( \hat{H} = \hat{E} + \hat{V} \), where \( \hat{E} \) is the kinetic energy and \( \hat{V} \), is the potential energy term. For a many-electron system, the Hamiltonian \( \hat{H} \) can be written as

\[ \hat{H} = -\frac{\hbar^2}{2m}\sum_i \frac{1}{m_\alpha} \nabla^2_\alpha - \frac{\hbar^2}{2m_e} \sum_l \nabla^2_l + \sum_\alpha \sum_\beta \frac{Z_\alpha Z_\beta e^{r^2}}{r_{\alpha\beta}} - \sum_\alpha \sum_l \frac{Z_\alpha e^{r^2}}{r_{l\alpha}} + \sum_\alpha \sum_{j>i} \frac{e^{r^2}}{r_{ij}} \]  (2.2)

Here in this equation, \( \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \), and \( -\frac{\hbar^2}{2m}\sum_i \nabla^2_i \) is nuclear kinetic energy, \( -\frac{\hbar^2}{2m_e} \sum_l \nabla^2_l \) is the kinetic energy of electrons, \( \sum_\alpha \sum_{\alpha > \beta} \frac{Z_\alpha Z_\beta e^{r^2}}{r_{\alpha\beta}} \) is the nuclear-nuclear repulsion term, \( \sum_\alpha \sum_l \frac{Z_\alpha e^{r^2}}{r_{l\alpha}} \) is electron-nuclear attraction term, \( \sum_\alpha \sum_{j>i} \frac{e^{r^2}}{r_{ij}} \) is electron-electron repulsion term. Based on equations 2.1 and 2.2, it is clear that the exact solution of the Schrödinger equation is not possible because of the electron-electron interaction. In
order to solve the equation numerically for many-electron systems, there used to be many assumptions and approximations.\textsuperscript{190,195-197} One of the very first approximations for the simplification of the Schrödinger equation is known as Born-Oppenheimer Approximation. Based on that approximation, it is possible to separate the electronic and nuclear motion to a high degree of accuracy. This alone is not sufficient to solve the Schrödinger equation for a many-electron system, which further requires the complete expansion of the many-electron wavefunction in terms of Slater determinants. Based on these many approximations there were many important theories were established such as Hartree-Fock (HF), Coupled Cluster (CC), Density Functional Theory (DFT) that can quite accurately calculate the properties of chemical systems.\textsuperscript{190,195,196,198,199} This level of theory can be used for a system containing hundreds of atoms with varying levels of accuracy and computational cost. In the case of protein which contains thousands of atoms, this QM level of theory is still impractical. On the other hand, molecular dynamics simulation (MD) based on the force field is quite famous for the large molecular systems irrespective of some limitations.\textsuperscript{190}

The MD method is developed based on the empirical molecular mechanical (MM) force field, and the accuracy of the MD is sharply limited. An example of the typical MM force field that is implemented in Amber is given below,

\[
E_{total} = \sum_{bonds} K_r (r - r_{eq})^2 + \sum_{angles} K_\theta (\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{V_n}{2} [1 + \cos(n\Phi - \gamma)] + \sum_{i<j} S_{ij}^{vdw} \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) + S_{ij}^Q \frac{q_i q_j}{\varepsilon r_{ij}}
\]

(2.3)
In equation 2.3, the first terms describe the bonded interactions formed by all the chemical bonds in the system; the second and third terms are the angular and dihedral interactions in the system. The fourth term denoted in the equation represents all the nonbonded interactions such as Vander-Waals and the Coulomb interactions of the chemical system. In the case of enzymatic catalysis, bond breaking and making are involved which is quite complex to explain with this level of theory. These difficulties can be addressed using the combined QM/MM method where the chemically important part is treated with high-level QM theory and the protein part which plays the role of an environment can be treated with the low level molecular mechanical method (MM) (Figure 2.1).\textsuperscript{192-194} In our studies, we have applied this QM/MM method to understand the photophysical and photochemical properties of vitamin B\textsubscript{12}-dependent enzymes, which involve the investigation of the ground and excited electronic states of the chemical system. Here in this section, a brief overview of the QM/MM methodology will be discussed.

![Figure 2.1. Schematic diagram of a protein with different partitions.](image-url)
In the QM/MM method, the entire system is divided into certain layers through a boundary between the QM and MM subsystem. The energy of the entire system can be calculated in two ways, the subtractive and the additive method.\textsuperscript{192,194} In the additive method, an MM calculation is performed for the outer MM region with force field, a QM calculation for the inner QM region and the interaction between the QM and the MM is considered through an explicit treatment of the coupling terms between the regions.\textsuperscript{192,194,200} This is a robust method and does not require the MM parameters for the QM partitions. The total energy of the additive scheme is calculated through the following equation,

\begin{equation}
E_{\text{total}} = E_{\text{QM}} + E_{\text{MM}} + E_{\text{QM/MM}}
\end{equation}

In this equation $E_{\text{QM}}$ is the energy of the high layer portion calculated by the QM level of theory, $E_{\text{MM}}$ is the energy of the low layer part calculated through the MM level of theory, $E_{\text{QM/MM}}$ is the electrostatic coupling between the QM and MM region. This electrostatic coupling can be addressed in two ways, electronic and mechanical embedding. In mechanical embedding (ME), the interactions between the QM and MM partition is considered through classical point charge, whereas in electronic embedding (EE), MM point charges are added to the QM Hamiltonian as an additional part. EE is more accurate in the sense that the surrounding environment polarizes the electron density in the QM region due to the external potential added in the QM Hamiltonian. For the subtractive QM/MM method the total energy is calculated based on the following equation,

\begin{equation}
E_{\text{total}} = E_{\text{QM,model}} + E_{\text{MM,real}} - E_{\text{MM,model}}
\end{equation}
This subtractive method is also known as ONIOM (Our Own N-layered Integrated Molecular Orbital and Molecular Mechanics) implemented in the Gaussian software suite. In that case, the real part contains all the atoms in the system and calculated with MM level of theory; the model system contains the atom in the high layer, which constitutes the atoms in the cofactor (Figure 2.1). Both QM and MM level of theory is applied for the model system. Hydrogen link atom is used to treat the covalent bond between the model system and the real system. The total energy is then calculated based on the equation 2.5. One of the advantages of the ONIOM system is that the calculations can be carried out using multiple layers. In the case of a three-layer calculation, an additional middle layer is used for further simplification of the active site. The middle layer can be treated with MM or an inexpensive QM level of theory. For a three-layer ONIOM calculation, the energy of the total system can be calculated based on the following equation,

\[
E_{total} = E_{QM,model} + E_{QM,middle} + E_{MM,real} - E_{MM,middle} - E_{MM,model} \tag{2.6}
\]

Equation 2.6 represents three-layer ONIOM calculations where the middle layer is being treated with the less expensive QM level of theory. Both ME and EE embedding can be employed in the ONIOM method. The ME scheme is much simpler and computationally less expensive because QM/MM electrostatic is treated by using a fixed atomic charge in the core and environment. The polarization of the QM wavefunction is not considered in this method. An interesting point to consider is that the geometry of the system is less affected by the core charges of the system, whereas it affects the energetics. In that case, the computationally effective way is to optimize the geometries with the ME and then single-point energy calculations can be performed with the EE scheme. In the following sections the methodology applied to prepare the QM/MM input will be discussed in detail.
2.3. QM/MM model preparation.

In order to perform the QM/MM calculations, the initial step is to prepare a model system based on the available X-ray structure. We have studied two coenzyme B$_{12}$-dependent enzymes, namely GLM and EAL. The initial structure of those systems was obtained from the protein data bank.$^{53,81}$ This is then subject to further modifications. At first, the crystal structure is protonated in the PDB2PQR web server using the Propka suite of program and the visual inspection.$^{201,202}$ The protonated structure is then used for further calculations. At first the amber parameters for the model system, which is AdoCbl in our case is prepared with metal center parameter builder (MCPB) implemented in the Amber program.$^{203-205}$ For the protein residue in the system, Amber FF14SB force field was employed.$^{206}$ In order to solvate the system, an octahedral box of TIP3P water was used. The total charge of the system was neutralized by adding Na$^+$ ions. After this initial setup, the whole system was minimized using steepest descent and conjugate gradient method for 10000 steps. A series of Harmonic restraints were applied, starting from 15 to 0.05 kcal/(mol.$\cdot$Å$^2$) to constraint the enzyme while keeping the TIP3P waters free to move. Finally, the constraint was removed, and the minimized structure was first heated for 100 ps from 0k to 300k with a time step of 1fs at constant volume and temperature (NVT) using Langevin dynamics.$^{207}$ Periodic boundary conditions were applied while performing the MD simulation. After the initial heating, the structure was equilibrated for 100 ps at 300k. Finally, a 200 ns production MD was performed at the NVT ensemble. Shake algorithm was applied to constraint the bonds and angles involving H atoms.$^{203}$

In order to obtain a better initial structure for the QM/MM calculations, the obtained results from the MD simulation were analyzed using the visual program VMD.$^{208}$ Based
on the rmsd analysis, certain snapshots from the production MD simulation were chosen for further QM/MM calculations. MolUp was used for the QM/MM input preparation from the MD snapshot.\textsuperscript{209} It is a VMD utility tool that can be effectively used for the Gaussian input preparation. Certain atoms from the center of the active site based on the distance were allowed to move whereas the rest of the system was frozen. In addition, TAO, a toolkit to assist ONIOM calculation was also used for the further modification of the input files.\textsuperscript{210} Specifically, the preparation of new inputs from previous calculations, analyzing the QM/MM results were done using the TAO program. The details methodology and the level of theory applied for our QM/MM calculations will be discussed in the next few sections.

2.4. DFT/MM calculations.

In order to investigate the ground-state properties of the enzymes, DFT/MM level of theory was applied. For a three-layer ONIOM calculation, DFT/semi-empirical/MM and for a two-layer system, the DFT/MM level of theory was applied. DFT is also an alternative approach to derive the approximate solutions of the time-independent Schrodinger equation. There are certain advantages of using DFT for a sizeable molecular system. In many-body wavefunction, the explicit use of wave function increased the degrees of freedom with the increasing number of particles which turns out to be impractical to solve the Schrodinger equation.\textsuperscript{190,191,195,196} On the other hand, unlike many ab-initio methods such as CI, HF, MRCI, MP2, DFT uses the concept of electron density which simplifies the degrees of freedom. The electron density is a function of position only; no matter how large the system is, it is only dependent on (x, y and z) coordinates whereas the wavefunction of n electron molecule is a function of 3n variables. The electron density of a system, $\rho(r)$ can be written as,
\[ \rho(r) = \sum_{i=1}^{N} N_i |\Psi_i|^2 \]  

(2.7)

Hohenberg and Kohn set the fundamental basis of the modern DFT calculation in 1964.\(^{198}\) The theorem suggests that the ground-state electron density function determines all the 
ground state properties of the molecule.\(^{198}\) For a system with electron density \(\rho_0\), the ground 
state energy is related to \(\rho_0(x, y, z) \rightarrow E_0\). The energy \(E_0\) is a functional of the electron 
density, \(E_0 = F[\rho_0]\). Thus, a functional exists, but the exact formula of that is unknown.

The second theorem of Hohenberg and Kohn suggest that, any trial electron density 
function will give energy higher than the exact ground state energy if the true electron 
density is known, \(E_v[\rho_t] \geq E_0[\rho_0]\).\(^{199}\) Based on this equation, it appears that the true 
ground state energy can be obtained if the accurate molecular electron density function and 
the exact energy functional is known. However, the challenges are that both the accurate 
electron density \([\rho_0(r)]\) and the exact energy functional is unknown. One of the ways to 
address this issue is to consider a fictitious non-interacting reference system in which 
electrons do not interact and in which ground-state electron density \(\rho_t\) is exactly the same 
as the ground state electron density of the real system, \(\rho_t = \rho_0\). Based on this approach there 
are two systems to consider, real electrons and the non-interacting electrons. The ground 
state electronic energy of the real system can be calculated based on the following equation,

\[ E_0 = \langle T[\rho_0]\rangle_{real} + \langle V_{Ne}[\rho_0]\rangle_{real} + \langle V_{ee}[\rho_0]\rangle_{real} \]  

(2.8)

Here in equation 2.8, the first term \(\langle T[\rho_0]\rangle\) is sum of the electron kinetic energies, \(\langle V_{Ne}[\rho_0]\rangle\) 
is the nuclear-electron attraction energies, \(\langle V_{ee}[\rho_0]\rangle\) is the electron-electron repulsion 
energies. The electron-nuclear potential energy can be calculated based on the simple
The challenges are to find the $\langle T[\rho_0] \rangle_{\text{real}}$ and $\langle V_{ee}[\rho_0] \rangle_{\text{real}}$, since functional for the kinetic and the potential part is not known. This can be addressed using the idea of the fictitious system, which can be considered as a reference system and the deviation from the real system kinetic energy can be written as,

$$\Delta \langle T[\rho_0] \rangle = \langle T[\rho_0] \rangle_{\text{real}} - \langle T[\rho_0] \rangle_{\text{ref}}$$  \hspace{1cm} (2.10)

or this can be rearranged to,

$$\langle T[\rho_0] \rangle_{\text{real}} = \langle T[\rho_0] \rangle_{\text{ref}} + \Delta \langle T[\rho_0] \rangle$$  \hspace{1cm} (2.11)

The deviation in electron-electron repulsion energy can also be calculated using the similar formula above,

$$\Delta \langle V_{ee}[\rho_0] \rangle = \langle V_{ee}[\rho_0] \rangle_{\text{real}} - \langle V_{ee}[\rho_0] \rangle_{\text{ref}}$$  \hspace{1cm} (2.12)

Equation 2.12, can be rewritten as,

$$\langle V_{ee}[\rho_0] \rangle_{\text{real}} = \langle V_{ee}[\rho_0] \rangle_{\text{ref}} + \Delta \langle V_{ee}[\rho_0] \rangle$$  \hspace{1cm} (2.13)

Now eqn., 2.9, 2.11, 2.12 can be replaced into Eqn. 2.8 for the energy expression,

$$E_0 = \langle T[\rho_0] \rangle_{\text{ref}} + \Delta \langle T[\rho_0] \rangle + \int \rho_0(r)v(r)dr + \langle V_{ee}[\rho_0] \rangle_{\text{ref}} + \Delta \langle V_{ee}[\rho_0] \rangle$$  \hspace{1cm} (2.14)

It appears from this Eqn. that most of the terms are easy to calculate; however, the difficulty is to calculate the deviation of kinetic energy and the electron-electron repulsion energy
together which is known as exchange correlation energy, which is a functional of the electron density,

\[ E_{XC}[\rho_0] = \Delta\langle T[\rho_0] \rangle + \Delta\langle V_{ee}[\rho_0] \rangle \]  \hspace{1cm} (2.15)

The other essential terms in the Eqn. are \( \langle T[\rho_0] \rangle_{ref} = -\frac{1}{2} \sum_{i=1}^{2n} \langle \psi_{i1}^{KS}(1) | \psi_{i2}^{KS}(2) \rangle \).

The nuclear electron attraction energy, \( \int \rho_0(r) v(r) dr = -\sum_{nuclei} Z_A \int \frac{\rho_0(r_1)}{r_{1A}} dr_1 \),

The electron-electron repulsion energy for the reference system,

\[ \langle V_{ee}[\rho_0] \rangle_{ref} = \frac{1}{2} \iint \frac{\rho_0(r_1)\rho_0(r_2)}{r_{12}} dr_1 dr_2 \]

Using all these terms to Eqn. 2.14 turns to,

\[ E_0 = -\frac{1}{2} \sum_{i=1}^{2n} \langle \psi_{i1}^{KS}(1) | \psi_{i2}^{KS}(2) \rangle - \sum_{nuclei} Z_A \int \frac{\rho_0(r_1)}{r_{1A}} dr_1 + \frac{1}{2} \iint \frac{\rho_0(r_1)n_0(r_2)}{r_{12}} dr_1 dr_2 + E_{XC} \] \hspace{1cm} (2.16)

Here in that Eqn. the term most susceptible to error is \( E_{XC} \) and also it is unknown. There are certain ways to solve this equation. Different approximations have been used for this exchange-correlation such as local density approximation (LDA) where it uses locally uniform electron gas to approximate the exchange functional. In the case of generalized gradient approximation (GGA), non-locally uniform electron gas approximation is being used. Some of the popular GGA functionals are BP86, BLYP, PBE91. Another popular functional is known as hybrid functional, where it mixes GGA with the HF exchange. One of the most popular hybrid DFT functionals is B3LYP. The accuracy of the DFT calculations are mostly dependent on the choice of a proper functional (Figure 2.2). For example, in the case of a transition metal system, it has been shown that the use of hybrid
The problem in using the B3LYP functional is coming from the inclusion of HF exchange in the hybrid functional. This is also consistent for the Cbl cofactors such as CH$_3$Cbl and AdoCbl. On the other hand, using pure GGA functional such as BP86, PBE correctly reproduce the BDE of the M-C bond. It is also important to mention the reproducibility of the structural parameters using that functional. Based on our studies of Cbls it appears that the use of BP86 is also crucial to correctly reproduce the structural parameters such as bond lengths, bond angles etc. Several other studies including Jensen et al., Smith et al, also showed the similar trend of DFT functionals. Based on these benchmark studies, we applied BP86 functional to our QM/MM studies. In the case of the QM layer, the BP86 functional was used with TZVP/P basis sets which is shown to produce an excellent agreement with the experiment.

Figure 2.2. A comparison of BDE and the bond lengths using different DFT functionals. A) Comparison of Co-C BDE of CH$_3$Cbl using different DFT functionals with experiment and high-level ab-initio calculations. b) Structural parameters comparison for the CH$_3$Cbl with experiment and different computational approaches.
2.5. TD-DFT/MM calculations

DFT method has been successfully used to understand the ground state properties of a molecule. Whereas in the case of excited-state properties, a different approach is required. A time-dependent version of DFT namely TD-DFT method has been employed to understand the optical properties of the molecular system based on the Runge-Gross theorem.\textsuperscript{218} The main idea based on the Runge-Gross theorem states that there is a one to one mapping between the external potential and the electronic one body density from a fixed initial state. According to that theory, the density determines the external potential which can be used to obtain the time-dependent wavefunctions and the observables of the system. Similar to the DFT method, the time-dependent potential can also be calculated by using a non-interactive system where the density is perturbed in an effective time-dependent potential. Based on the idea of Runge-Gross theorem, the time-dependent Kohn Sham equation can be written as similar to the ground state Kohn-Sham equation, such as,

\begin{equation}
  i \frac{\partial}{\partial t} \phi_i(\vec{r}, t) = \left[ -\frac{\nabla^2}{2} + v_{KS}(\vec{r}, t) \right] \phi_i(\vec{r}, t) \tag{2.17}
\end{equation}

The time-dependent density of the interacting system can be obtained using the following formula,

\begin{equation}
  \rho(\vec{r}, t) = \sum_{t}^{occ} |\phi_i(\vec{r}, t)|^2 \tag{2.18}
\end{equation}

The Kohn-Sham potential $v_{KS}$ from equation 2.17 can be separated in the following term,

\begin{equation}
  v_{KS}(\vec{r}, t) = v_{ext}(\vec{r}, t) + v_{Hartree}(\vec{r}, t) + v_{XC}[\rho(\vec{r}, t)] \tag{2.19}
\end{equation}
Here in this equation, the first and second term have their usual meaning, whereas the third term is the exchange-correlation potential. Similar to DFT the exact form of this functional is unknown and hence it requires different approximations. Hence TD-DFT calculate the properties of the excited state considering the response of the system’s ground state to an external perturbation. One of the common ways to calculate the response properties such as excitation energies, oscillator strength, dynamic polarizability etc. is to apply linear response formalism in TD-DFT which is known as LR-TDDFT. Similar to DFT, the TD-DFT method is also widely used because of the computational cost and comparable accuracy. Since Cbls model system contains hundreds of atoms, we applied TD-DFT along with the MM level of theory to calculate the excited state properties.

The validity of using TD-DFT for Cbls has already been explored in many studies. It was shown that the accuracy of the ground and excited state properties based on DFT and TD-DFT calculations are dependent on the choice of a proper functional. Hence a proper benchmark study is crucial to identify the correct functionals for specific studies. For Cbls, TDDFT calculations were benchmarked against high level ab-initio calculations as well as Abs and TAS studies. CNCbl has been referred as a paradigm system for the evaluation of the excited states of Cbls, so it can be used as a basis for the understanding of other Cbls. The manifold of low-lying excited states was calculated using TDDFT for CNCbl and compared with high level ab-initio calculations. Three different functionals B3LYP, BP86, and LC-BLYP were used and compared. The energy gap between occupied and the virtual orbital using BP86 functional was more accurate compare to the hybrid functional B3LYP. However, the electronic transitions in the CNCbl were mostly predicted as $\pi \rightarrow \pi^*$ transitions by using B3LYP. This is
contradictory with CD and MCD studies. On the other hand, the BP86 results indicate that the α/β region of the CNCbl is dominated by the mixed contributions of \( \pi \rightarrow \pi^* \) transitions with the transitions associated from the d orbitals of the Co ion. This is known as Ligand-to-metal charge transfer (LMCT) consistent with the high level second order multi-configurational quasi-degenerate perturbation theory (MC-XQDPT2) and complete active space self-consistent field (CASSCF) level of theory.\textsuperscript{221,222}

TD-DFT method was also benchmarked for the CH\(_3\)Cbl cofactor to provide the nature of the low-lying excited state specifically S\(_1\) state and compared with TAS.\textsuperscript{173} Based on the TAS studies by Sension and co-workers, it was shown that the S\(_1\) state in CH\(_3\)Cbl is a metal-to-ligand charge transfer (MLCT) type transition.\textsuperscript{124,155} This is contradictory to the previously reported \( \pi \rightarrow \pi^* \) transitions localized on the corrin. Theoretical calculations using high-level CASSCF/MC-XQDPT2 theory provided a consistent description of the MLCT transitions, in agreement with the TAS experiments.\textsuperscript{222} On the other hand, using TDDFT provides variable conclusions depending on the use of the functional. Hybrid functional such as B3LYP, TPSSH, MPW1PW91 predicted the S\(_0\)→S\(_1\) as \( \pi \rightarrow \pi^* \) transitions in contrast to the TAS and high-level ab-initio calculations. On the other hand, GGA type functional otherwise known as pure functional such as BP86, BLYP characterize the transitions as a mix contribution of \( \pi \rightarrow \pi^* \) and \( \pi \rightarrow \text{metal d} \) orbitals consistent with the MLCT transitions. Based on these studies, it appears that the use of BP86 functional for the studies of Cbls using TDDFT is very crucial.\textsuperscript{124,155,173,220-223}

2.6. Long-range charge transfer (LR-CT) excitations in TD-DFT

It is clear that the TDDFT is a widely used method for the excited state calculations because of its extremely favorable accuracy to the cost ratio. However, this method has
some severe flaws, underestimation of which can cause a serious error in the calculations. One of the commonly known is the LR-CT issue, associated with the underlying exchange correlational functional used in the TD-DFT methodology. These issues originate from the incorrect estimation of the electronic excitations from the unphysical low-lying electronic states otherwise referred to as ghost states. In the case of Cbls, these errors are caused by long-range electronic excitations which have significantly poor overlap between occupied and the virtual orbitals.\textsuperscript{223} Different diagnostics and tests have been proposed in order to address this issue based on the occupied and virtual orbitals such as \( \Lambda \) index developed by Tozer and co-workers.\textsuperscript{224} This index is used to measure the degree of the orbital overlap between occupied and the virtual orbitals with a numerical value range from 0 to 1. A lower value of \( \Lambda \) indicates poor overlap which corresponds to the problematic LR-CT transition. For pure GGA functional, it was shown that the \( \Lambda < 0.4 \) and for hybrid functional \( \Lambda < 0.3 \) indicates significantly underestimated electronic transitions.\textsuperscript{224} The electronically excited states of isolated Cbls were evaluated using the \( \Lambda \) diagnostics.\textsuperscript{222,223}

There is another method to investigate the extent of LR-CT issues in TD-DFT using a density-based matrix otherwise known as the D\textsubscript{CT} index developed by Ciofini and co-workers.\textsuperscript{225,226} This density-based index is computed from the electron density associated with the ground state \( \rho_{GS}(r) \) and excited state \( \rho_{EX}(r) \). The change in density can be calculated by this equation, \( \Delta \rho(r) = \rho_{EX}(r) - \rho_{GS}(r) \). Two functions were defined \( \rho_+(r) \) and \( \rho_-(r) \) based on the density augmentation and the depletion zone associated with the electronic transition. Two barycenter’s of the spatial regions \( R_+ \) and \( R_- \) were defined based on the \( \rho_+(r) \) and \( \rho_-(r) \) based on the following equations,
\[ R_+ = \frac{\int r \rho_+(r) dr}{\int \rho_+(r) dr} = (x_+, y_+, z_+) \quad \text{and} \quad R_- = \frac{\int r \rho_-(r) dr}{\int \rho_-(r) dr} = (x_-, y_-, z_-) \]

The \( D_{CT} \), can be calculated by taking the difference between these two barycenters density distribution, \( D_{CT} = |R_+ - R_-| \). In order to calculate the transferred charges \( (q_{CT}), \rho_+(r) \) or \( \rho_- (r) \) can be integrated over-all space. Such as, \( q_{CT} = \int \rho_+(r) dr = \int \rho_-(r) dr \)

For one-electron excitation the value of \( q_{CT} \) between 0 and 1. The variation of the dipole moment \( (\mu_{CT}) \) between the ground and excited state can thus be defined \( \mu_{CT} = D_{CT} q_{CT} \).

The centroid of an axis associated with the positive and negative charges defined as \( C_{+/-} \) is an important parameter to understand the center of the charges, the formula of which is below,

\[
C_+(r) = A_+ e\left(- \frac{(x - x_+)^2}{2\sigma_{+x}^2} - \frac{(y - y_+)^2}{2\sigma_{+y}^2} - \frac{(z - z_+)^2}{2\sigma_{+z}^2}\right)
\]

\[
C_-(r) = A_- e\left(- \frac{(x - x_-)^2}{2\sigma_{-x}^2} - \frac{(y - y_-)^2}{2\sigma_{-y}^2} - \frac{(z - z_-)^2}{2\sigma_{-z}^2}\right)
\]

Here in the above equations, \( A_{+/-} \) is the normalization constant, \( \sigma_{+/-} \) is the root mean square deviations along the three axes. The half of the sum of the centroid axis along any specific directions such as x can be defined as, \( H = \frac{\sigma_{+x} + \sigma_{-x}}{2} \). These mathematical equations have some useful meaning and can be used to understand the charge transfer excitations. Such as, \( H \geq D_{CT} \) implies a significant overlap between the centroids along the chosen direction. So, the electronic transitions occur through the \( \pi \) space of the system. On the other hand, if \( H \leq D_{CT} \) indicates poor overlap between the density augmented and depleted zone which is the indication of charge transfer through space. This is usually an indication of the LR-CT states. The value of charge transfer distance is also another indication of the
LR-CT states. For a significantly higher value of $D_{CT}$ indicates poor overlap between the states. Thus $D_{CT}$ parameters can be a useful tool to measure the charge transfer extent in different systems. This is widely used in the case of a push-pull dye system, whereas a careful execution of this index can make it valuable tools for any other purpose. In the case of Cbls, the LR-CT issue mainly arise due to the electronic transitions between the Ado and corrin ligands.\textsuperscript{179,180,222-225} This was identified based on other parameters such as the Tozer index, evaluating density and orbital characterizations.\textsuperscript{224} Based on the $D_{CT}$ index it appears that the transitions occurring at $D_{CT} \geq 6\text{Å}$ are problematic which arises due to the Ado and corrin ligand transitions.\textsuperscript{178-180} These are typically the ghost states and can be properly investigated using this index.
CHAPTER III
MECHANISM OF LIGHT-INDUCED RADICAL PAIR FORMATION IN COENZYME B₁₂-DEPENDENT ETHANOLAMINE AMMONIA-LYASE

3.1 Background

Ethanolamine ammonia-lyase (EAL) is an eliminase class of B₁₂-dependent enzyme that is found in archaea and bacteria. Carboxysome like organelle, which is encoded by Eut operon, produces EAL as the first enzyme in the ethanolamine (EA) degradation pathway as it is the source of carbon (C), nitrogen (N), and energy. However, to degrade EA, the enzyme EAL requires the AdoCbl cofactor in aerobic conditions. As discussed before, AdoCbl is a biologically active vitamin B₁₂ derivative, which is also known as coenzyme B₁₂. It is a highly complex organometallic molecule containing an Ado molecule as the upper axial ligand and the DBI as the lower axial base (Figure 1.1). The unique feature of the AdoCbl cofactor is an organometallic Co-C σ bond, which plays a crucial part in the enzymatic catalysis. It is generally believed that the Co-C bond of the enzyme bound AdoCbl is not significantly activated without the presence of a substrate. Binding of substrate EA, to the enzyme, triggers the cleavage of the Co-C bond leading to the formation of the Co(II) and Ado radical pair (Figure 3.1).
Figure 3.1. Native catalysis cycle for AdoCbl-dependent EAL.

Subsequently, the Ado radical abstracts an H-atom from the substrate to generate a stable Co(II)/Substrate radical pair which is the first detectable intermediates using electron paramagnetic resonance (EPR) spectroscopy.\(^3,5,6\) The cleavage of this Co-C bond is rapid, and the rate of the reaction is significantly enhanced by approximately \(10^{12}\)-fold inside the enzymatic environment compare to the homolysis of free AdoCbl in solution.\(^3,5,4,8,7\) The mechanism of this Co-C bond activation and the factors responsible for the observed rate acceleration remain elusive.

In contrast to native catalysis, the Co-C bond of AdoCbl cofactor inside EAL can also be cleaved homolytically by light, to generate Co(II) and Ado radical pair (Figure 3.2).\(^152,166,168,177,178\) Specifically, the AdoCbl cofactor inside EAL undergoes photolysis upon excitation at a wavelength between 530 and 300 nm. The specific mechanism of this Co-C bond photolysis inside the enzyme occurs differently than the isolated AdoCbl cofactor. Thus, the environment plays a critical role in the photodissociation mechanism of the AdoCbl cofactor. The ability to probe photolytic cleavage of the Co-C bond for enzyme
bound AdoCbl is of particular relevance in enzymatic catalysis. In EAL homolysis of the
Co-C bond is triggered by a substrate, which is kinetically coupled to the subsequent H-
atom abstraction from the substrate EA. The crucial aspect of the photolysis in EAL is that
it generates a similar pair of radicals as in the case of native enzymatic catalysis without
coupling to the H-atom abstraction. Thus, the photolysis of the Co-C bond inside EAL can
be used to mimic the native enzymatic catalysis (Figure 3.2).

**Figure 3.2.** Photolysis in Co-C bond in AdoCbl-dependent EAL. Various intermediates formed
during the reaction were shown in the figure.

The photolytic properties of AdoCbl-dependent EAL have been investigated using
a variety of experimental techniques.\(^{166,168}\) It has been shown that the protein environment
appears to influence the excited state of the cofactor rather than the ground state.
Significant changes have been observed in the rate constant for the recombination of the
post-homolysis products in the enzymatic environment when compared to the solution. It
has been suggested that the protein acts as a sort of cage to prevent the diffusion of the Ado
radical. This is evidenced by the increase in the recombination rate of the radical pair in EAL. As a result, the quantum yield of the long-lived radical pair is reduced as more tends toward geminate recombination.166

In addition to these experimental studies, significant theoretical work has been completed to explore the photolytic properties of the AdoCbl cofactor in solution.42,214 Specifically, DFT and TD-DFT were applied to elucidate the ground and excited state's details of the Cbls cofactors. Reliable structural models with various axial ligands have been used to construct PESs as a function of the axial bond lengths. Besides, semi-classical Landau-Zenner theory has been applied to explore the possibility of the involvement of the triplet states. In this chapter, a detail computational insight about the photochemical properties of AdoCbl inside EAL enzyme will be provided. Combined DFT/MM and TD-DFT/MM methods were applied to investigate the ground and low-lying properties of the EAL enzyme.

3.2 Computational details

The crystal structure of EAL was first discovered in 2010 by Toraya and co-workers.53 The EAL enzyme from E. coli was crystallized at a resolution of 2.25Å (PDB ID:3ABS) with two asymmetric units.53 The asymmetric units are identical, and the reactive parts, the cofactor, and the substrate are independent of their respective units and do not interact. One asymmetric unit of EAL is composed of α and the β subunits with adeninylpentylcobalamin (AdePeCbl) instead of AdoCbl cofactor and EA as a substrate. To prepare a reasonable initial model for the calculations, we obtained the crystal structure from the protein data bank (PDB ID: 3ABS).53 The PDB structure was then modified accordingly for further calculations (Figure 3.3).
3.2.1 QM/MM set up. The crystal structure obtained from PDB was initially modified for further calculations. One asymmetric unit was used for the calculations. In the reported crystal structure, AdePeCbl was shown as a cofactor that binds in the base-on conformations in the presence of a substrate EA. This initial structure was then protonated using Propka 3.0 software suite, and the protonation states of the titrable residues were determined based on the manual inspection. The AdePeCbl was modified by adding a ribose moiety to the purine ring to restore the realistic structural model. Finally, substrate EA was removed from the structure, and two water molecules were added to fill the empty
spaces. The substrate was removed in order to ensure that Co-C bond activation would be based on light and not by substrate triggering. The modified EAL structure was then minimized with the MM level of theory using the amber force field (FF14SB). The minimized structure was then divided into three layers. The corrin ring, imidazole (Im) part of the DBI base with the side chain replaced by H, Co ion, and Ado moiety were added to the high layer system. The high layer part of the cofactor is shown in Figure 3.4a.

![Figure 3.4](image)

**Figure 3.4.** a) Truncated structure of AdoCbl cofactor containing atom numbers, which was used in the QM region (high layer). (b) Model structure of RibCbl, with the simplified structure of the Ado ligand containing only the ribose ring, that was used in the QM region (high layer) for calculations applying different QM and MM partitions.

The remaining part of coenzyme B₁₂ with the additional water molecules was placed into the middle layer, and the rest of the protein, including the crystal water, was placed into the low layer system. Atoms within 20 Å from the Co center were kept unfrozen, and the rest of the protein was frozen. The model used in calculations contained a total of 11988
atoms, with 3550 unfrozen and 8438 frozen. All reported QM/MM calculations were performed using the Gaussian 09 software and Gaussian 16 software.228,229

3.2.2 DFT/MM calculations. The high layer part of the system, where the Co-C bond cleavage takes place, was computed using DFT with BP86 functional. The use of BP86 is crucial for the Cbl system, as already discussed in the Introduction chapter. In the case of EAL, this functional provides a good agreement with experimentally determined BDE as well as the structural parameters such as bond length, angles from the crystal structure (Table 3.1). To maintain consistency with our previous calculations, the TZVP basis set was used for H and TZVPP was used for the Co, C, N, and O in the QM layer. The middle layer of the system was computed using the semi-empirical (PM6) level of theory while the protein part of the system was treated with amber force field (FF99SB).203 Amber parameters for the AdoCbl cofactor were obtained from Marques et al.230 Ground state ONIOM calculations (DFT)/PM6/MM(Amber) were performed to optimize the structure of the EAL. Default optimization criteria were used for the calculations. Then, using the optimized geometry, the ground state (S0) potential energy curves (PECs) were constructed as a function of Co-C bond distance. The 3D S0 potential energy surfaces (PESs) for AdoCbl inside EAL and solutions, Im-[CoIII(corrin)]=Ado+, and H2O-[CoIII(corrin)]=Ado+ were constructed by optimizing the geometries using the QM (DFT)/PM6/MM(Amber) level of theory as a function of Co-C and Co-NIm bond lengths.
Table 3.1. Selected geometric parameters for optimized geometries for AdoCbl in EAL, in solution in base-on and base-off forms respectively as well as parameters for the crystal structure of EAL from (PDB id: 3ABS), and crystal structure of AdoCbl. The optimized structure of isolate AdoCbl cofactors is shown in figure A2.
3.2.3 TD-DFT/MM calculations. The optimized S₀ geometries were then used for the TD-DFT/MM calculations to obtain the low-lying singlet and triplet excited state. By using the S₀ optimized geometry, we applied single point TD-DFT calculations in constructing the S₁ PESs as a function of the axial bond lengths. The detailed orbital analysis was also performed to characterize the nature of electronic excitations. Due to the problems associated with Gaussian 09 ONIOM calculations to integrate the orbital data in the high layer, additional single-point calculations were performed using Gaussian 16 to obtain the molecular orbitals. The MLCT region of the S₁ surface was then characterized to have the excitations from the Co d orbitals to the corrin π*, whereas in the LF region, the electronic transitions occur from the Co d and corrin π orbitals to the Co dₓ² orbital. The LR-CT issues in TD-DFT methodology were considered and to ensure that TD-DFT/MM energy surfaces were not affected by the LR-CT problem, a careful analysis of corresponding electronic excitations was performed by applying several tests and diagnostics.

As a crude level, we partitioned the Ado part into two layers, ribose, and purine. To investigate with our corresponding PESs, we modified the partitioning in ONIOM calculations. Corrin ring and the Im part of the lower axial base were placed into the high layer system. The rest of the cofactor, including the purine ring, was then placed into the middle layer, and the protein part into the low layer. This structure is referred to as RibCbl and shown in figure 3.4b. This structural model was then used for the additional DFT/MM and TD-DFT/MM calculations in constructing the S₀ and S₁ PESs of RibCbl using the same level of theory, as discussed before. The corresponding PES of RibCbl revealed that the topology of the PESs in the case of RibCbl and AdoCbl is similar. The insensitivity concerning the QM/MM partition reflects the nature of electronic transitions, which are
primarily localized within Co and the corrin ring. Finally, a more rigorous analysis was also performed by applying the D_CT diagnostics. The details of this diagnostics were already discussed in Section 2.6. Gaussian 16 was applied to conduct the calculations. Based on the D_CT data, it appears that the constructed PECs and PESs from our calculations were reasonably accurate and did not suffer from the LR-CT issues.

3.3. Results and discussions

3.3.1 Ground state geometry. The modified structure of the substrate-free AdoCbl-dependent EAL has been optimized using the procedure described above (Figure 3.5). Frequency calculations confirmed that this was the optimal geometry, as no imaginary frequencies were present. The structural parameters of the optimized geometry were compared with the experimental values from the crystal structures of EAL (PDB ID: 3ABS).\textsuperscript{53,178} Besides, the results were also compared with experimental data from the high-resolution crystal structure of AdoCbl (Figure 3.5) as well as with the geometrical parameters obtained from DFT/BP86 calculations for the base-on and base-off model complexes in solution from previous computational studies (Table 3.1). When comparing the crystal structure of EAL to the optimized geometry, some structural differences were observed regarding the corrin ring and the axial bond distances. The axial Co-C and Co-N\textsubscript{Im} bond lengths in the optimized structure of AdoCbl in EAL were 2.05 and 2.37 Å, respectively (Table 3.1). The experimental values for these bond lengths in EAL crystal structure were 2.01 Å for the Co-C and 2.61 Å for the Co-N\textsubscript{Im} bonds (Table 3.1). These structural changes were most likely induced due to the replacement of the AdePeCbl moiety with AdoCbl.
Figure 3.5. (a) Crystal structure of EAL (PDB id: 3ABS) with AdePeCbl cofactor where adenine-9-yl-pentyl (AdePe) is the upper axial ligand. Substrate EA is depicted in purple. (b) Optimized structure of EAL with reconstructed AdoCbl cofactor where AdePe is replaced with the Ado ligand. Substrate EA was removed and replaced by two water molecules.

To further explore the differences associated with the lower axial base in EAL, the Co-N\textsubscript{Im} bond distance was compared to the results obtained from the calculations of the base-on and base-off AdoCbl models in solution (Table 3.1). In the case of the base-on form where the lower axial ligand is DBI, the Co-N\textsubscript{Im} bond is 2.19 Å compared to 2.37 Å in EAL. In the base-off form where DBI is replaced by water, the Co-O bond is 2.43 Å. These results indicate that the axial bond distances are sensitive to the environment, whether the cofactor is in solution or within an enzyme.

The most significant differences observed between AdoCbl inside EAL and the base-on and base-off AdoCbl models in solution were the conformations of the ribose ring of the Ado ligand. The structure of the Ado ligand consists of two units, purine and the ribose ring (Figure 3.6).
The relative orientation of the ribose with respect to the purine ring can be explained based on the glycosyl rotation angle ($\chi$), pseudorotation phase ($P$) and pseudorotation amplitude ($\Theta_m$) (the definition of $\chi$, $P$ and $\Theta_m$ is shown in figure 3.6. Note that the pseudorotation values consisting of $-90^\circ$ to $90^\circ$ correspond to the 3'-endo conformation of the ribose, whereas the $P$ with $90^\circ$ to $270^\circ$ corresponds to the 2'-endo conformation. To understand the influence of the enzymatic environment on the ribose conformation, the structural data of the ribose, extracted from the EAL optimized structure, was compared with optimized geometries of isolated base-on and base-off AdoCbl models (Table 3.2). Due to the absence of the ribose ring in the x-ray crystallographic data of EAL, it was not possible to make a direct comparison with the optimized structural model of substrate-free AdoCbl-dependent EAL. However, it was found that the ribose ring in all the structures was in 3'-endo conformation. The most significant difference observed was concerning the $\chi$ when
compared with the x-ray structure of AdoCbl and the calculated base-on and base-off forms. In the case of AdoCbl inside EAL, the $\chi = -72.46^\circ$, whereas in the crystal structure of isolated AdoCbl, $\chi = 68.00^\circ$. On the other hand, in the optimized models of base-on and base-off AdoCbl, the corresponding $\chi$ angles were $-13.69^\circ$ and $-92.33^\circ$, respectively (Table 3.2).

<table>
<thead>
<tr>
<th>Exocyclic angle (°)</th>
<th>Optimized structure of EAL</th>
<th>Crystal structure of AdoCbl</th>
<th>Calc. AdoCbl in solution base-on</th>
<th>Calc. AdoCbl in solution base-off</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1'-C2'-O2'-O2'H ($\Phi_2$)</td>
<td>-30.25</td>
<td>70.06</td>
<td>61.94</td>
<td>-36.88</td>
</tr>
<tr>
<td>C2'-C3'-O3'-O3'H ($\Phi_3$)</td>
<td>9.110</td>
<td>0.340</td>
<td>-13.16</td>
<td>-11.10</td>
</tr>
<tr>
<td>O1'-C1'-N1'-C5' ($\chi_{CN}$)</td>
<td>-72.46</td>
<td>68.10</td>
<td>-13.69</td>
<td>-92.33</td>
</tr>
<tr>
<td>Endocyclic angle (°)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1'-C2'-C3'-C4' ($\Theta_0$)</td>
<td>9.870</td>
<td>34.47</td>
<td>2.800</td>
<td>20.84</td>
</tr>
<tr>
<td>C2'-C3'-C4'-O1' ($\Theta_1$)</td>
<td>-35.86</td>
<td>-38.72</td>
<td>-29.30</td>
<td>-42.30</td>
</tr>
<tr>
<td>C3'-C4'-O1'-C1' ($\Theta_2$)</td>
<td>49.72</td>
<td>27.81</td>
<td>46.45</td>
<td>48.23</td>
</tr>
<tr>
<td>C4'-O1'-C1'-C2' ($\Theta_3$)</td>
<td>-43.30</td>
<td>-5.70</td>
<td>-45.12</td>
<td>-34.49</td>
</tr>
<tr>
<td>O1'-C1'-C2'-C3' ($\Theta_4$)</td>
<td>19.54</td>
<td>19.03</td>
<td>25.55</td>
<td>6.970</td>
</tr>
<tr>
<td>Pseudorotation Phase (P)</td>
<td>78.48</td>
<td>40.85</td>
<td>86.64</td>
<td>64.10</td>
</tr>
<tr>
<td>Pseudorotation amplitude ($\Theta_m$)</td>
<td>49.42</td>
<td>45.56</td>
<td>47.77</td>
<td>47.71</td>
</tr>
<tr>
<td>Ribose conformation</td>
<td>3'-endo</td>
<td>3'-endo</td>
<td>3'-endo</td>
<td>3'-endo</td>
</tr>
</tbody>
</table>

Table 3.2. Selected structural parameters for the optimized structure of EAL, the crystal structure of AdoCbl, and the optimized structure for AdoCbl base-on and base-off forms and the conformation of ribose ring in the Ado ligand.

Although our model does not contain the substrate, we believe that these significant differences observed in $\chi$ when comparing the structures of AdoCbl in solution to AdoCbl inside EAL can help to understand the labilization of the Co-C bond in the presence of substrate. There is still much to be understood about how the protein environment contributes to this. We feel as though the results from our electronic structure calculations corroborate the idea that the $\chi$ of the ribose moiety is an essential factor to be considered.
when elucidating the mechanism of Co−C bond homolysis in AdoCbl-dependent enzymatic catalysis.

3.3.2. Electronically excited states along the Co-C bond distance. To provide the preliminary insight into the photodissociation of AdoCbl inside EAL, the $S_0$ energy curve was generated by systematically elongating the Co-C bond distance with step size 0.05 Å within the range of 1.80 Å to 2.60 Å. The $S_0$ geometry of EAL was optimized at each point. Then, single point TD-DFT calculations were carried out to obtain the manifold of low-lying singlet and triplet excited states. For each optimized geometry, thirty excited states were calculated up to the energy value of 3.1 eV in EAL, and this was compared with the excited states obtained for the AdoCbl base-on and base-off forms (Figure 3.7 and 3.8). Although the energy curves computed inside EAL follows the similar topology as in the case of base-on and base-off AdoCbl, there are some noticeable differences (Figure 3.7 and 3.8). The most apparent similarity in all the cases is the repulsive triplet state ($\sigma\rightarrow\sigma^*$), which drops in energy, and at bond length 2.4 Å, it becomes the lowest energy state. As discussed before, this is due to the single determinant wave function used in TD-DFT methodology. This state should correctly level off in a multireference wave function approach. Inside the enzyme, the vertical excitations energies are significantly smaller when compared to the values obtained for the solution medium. In EAL the electronic transitions involved in the $S_1$ state are 59% $d_{xz}/d_{x^2} + \pi\rightarrow\pi^*$, 23% $d_{yz} + \pi\rightarrow\pi^*$ and 11% $d_{xz} + \pi\rightarrow\pi^*$ (Table 3.3).
Figure 3.7. Potential energy curves for the ground (black) and lowest, vertical singlet (red) and triplet (blue) excited states as functions of Co-C bond length for (a) AdoCbl inside EAL and (b) Im−[CoIII(corrin)]−Ado+ base-on model complex in water.

The nature of the electronic transition in the S₁ electronic state can be characterized by MLCT type because all the mentioned transitions are predominately from the metal d orbitals to the corrin π* orbital. This is consistent with the previous theoretical studies for alkylCbls. The next excited state S₂ can be characterized as 64% d<sub>yz</sub> + π → π*, 21% d<sub>x,y</sub>/d<sub>z</sub><sup>2</sup> + π→π* transitions. In excited state S₃, the major contribution is an excitation from HOMO-5 to LUMO, which has an 88% d<sub>xy</sub>→π* character. All these data were explicitly shown in table 3.3.
Figure 3.8. Potential energy curves for the ground and lowest, vertical singlet and triplet excited states as functions of Co-C bond length for (a) AdoCbl inside EAL and (b) H₂O-[CoIII(corrin)]−Ado⁺ base-off model complex.

Eight lowest singlet excited states were characterized on the basis of the orbital analysis, and the data were summarized in table 3.3. Corresponding orbital pictures of this analysis shown in figure 3.9. For the low-lying singlet excited states inside EAL, the first vertical excitation was found at 1.97 eV, whereas for the base-on and base-off AdoCbl, the calculated excitation energies did not exceed 2.30 and 2.38 eV respectively.
Table 3.3. Eight lowest vertical singlet electronic transitions and orbital characterization based on the single point TD-DFT/MM calculations of AdoCbl inside EAL

<table>
<thead>
<tr>
<th>E[eV]</th>
<th>f</th>
<th>λ[nm]</th>
<th>DCT %</th>
<th>Character</th>
<th>Exp. [nm] (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>1.97</td>
<td>.0038</td>
<td>629.4</td>
<td>1.029</td>
<td>H-1→L d_v+π → π*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>59</td>
<td>H-2→L d_v/d_y^2+π → π*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td>H-4→L d_v+π → π*</td>
</tr>
<tr>
<td>S2</td>
<td>2.12</td>
<td>.0254</td>
<td>587.5</td>
<td>1.247</td>
<td>H-1→L d_v+π → π*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>21</td>
<td>H-2→L d_v/d_y^2+π → π*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>H-2→L+3 d_v/d_y^2+π → π*</td>
</tr>
<tr>
<td>S3</td>
<td>2.26</td>
<td>.0027</td>
<td>548.1</td>
<td>0.913</td>
<td>H-1→L d_v+π → π*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td>H-4→L d_v+π → π*</td>
</tr>
<tr>
<td>S4</td>
<td>2.33</td>
<td>.0385</td>
<td>530.8</td>
<td>1.331</td>
<td>H-1→L+1 d_v+π → d_y^2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td>H-2→L d_v/d_y^2+π → π*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>65</td>
<td>H-4→L d_v+π → π*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td>H-5→L d_v+π → π*</td>
</tr>
<tr>
<td>S5</td>
<td>2.51</td>
<td>.0035</td>
<td>490.5</td>
<td>3.971</td>
<td>H-1→L+1 d_v+π → d_y^2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>67</td>
<td>H-4→L+2 π_Ado → d_y-n</td>
</tr>
<tr>
<td>S6</td>
<td>2.59</td>
<td>.0081</td>
<td>490.4</td>
<td>5.428</td>
<td>H-1→L+2 π_Ado → d_y-n</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>55</td>
<td>H-1→L+1 d_v+π → d_y^2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>H-4→L d_v+π → π*</td>
</tr>
<tr>
<td>S7</td>
<td>2.66</td>
<td>.0024</td>
<td>466.8</td>
<td>5.942</td>
<td>H-6→L d_v/d_y^2+π_Ado+π_Rib→ π*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>70</td>
<td>H-7→L d_v/d_y^2+π_Ado+π → π*</td>
</tr>
<tr>
<td>S8</td>
<td>2.68</td>
<td>.0022</td>
<td>459.2</td>
<td>3.695</td>
<td>H-1→L+1 d_v+π → d_y^2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td>H-2→L+1 d_v/d_y^2+π → d_y^2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>52</td>
<td>H-8→L d_v+π_Ado+π → π*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>H-7→L d_v+π_Ado+π → π*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td>H-6→L d_v/d_y^2+π_Ado+π_Rib→ π*</td>
</tr>
</tbody>
</table>

63
Figure 3.9. HOMO and LUMO molecular orbitals for the lowest vertical singlet electronic transitions.

The results obtained here for EAL is in good agreement with the results from the base-on isolated AdoCbl complex. In the case of the base-off model complex, antibonding character was observed in the $S_3$ state. Antibonding $\sigma^*$ character, which is an electron excitation from the HOMO-1 to LUMO+1 in EAL, was observed in $S_4$ state and up. The major component of transition contributions in the $S_4$ excited state is $7\% \, d_{yz} + \pi \rightarrow d_z^2$, $12\% \, d_{xy}/d_{xz}^2 + \pi \rightarrow \pi^*$, $65\% \, d_{xz} + \pi \rightarrow \pi^*$, and $7\% \, d_{xy} \rightarrow \pi^*$. A significant amount of $\sigma^*$ character occurs in the $S_5$ state where the transition contribution is $67\% \, \pi \text{Ado} \rightarrow d_{xy} + n + \pi^*$ and $26\% \, d_{xy} + \pi \rightarrow d_z^2$. The other significant difference between base-on and base-off AdoCbl is the absence of pure $\pi \text{Ado} \rightarrow \pi^*$ transition in the presented results of the calculations.

The energies of excited states reported here (Table 3.3) are typically lower than the energies in base-on and base-off AdoCbl in solution. This significant reduction in excitation
energies agreed well with the idea of lowering of energies induced by the enzymatic environment.

### 3.3.3. Potential energy surfaces as a function of axial bond lengths.

It was discussed before that upon excitations, the largest structural changes in Cbls occur were in the case of axial ligands, which is also the basis of our study in EAL. Because of this light sensitivity of the axial ligands, it is crucial to explore the PESs based on the Co-C and Co-N$_{Im}$ bond lengths. For EAL, the $S_0$ PES was constructed using DFT/PM6/MM by systematically elongating the Co-C and Co-N$_{Im}$ axial bonds using a step size of 0.1 Å. The manifold of low-lying singlet $S_1$-$S_n$ excited states was obtained from TD-DFT/PM6/MM calculations starting with the $S_0$ PES. The major feature of the $S_0$ PES is the single energy minimum, which is consistent with the base-on and base-off AdoCbl. The corresponding lowest energy region is around Co–C of 2.05 Å and the Co–N$_{Im}$ bond at 2.37 Å ($S_0$, Figure 3.10). On the other hand, the $S_1$ PES is characterized by two energy minima separated by a seam. Based on the analysis of the $S_0$ and $S_1$ PES it appears that the topology of the corresponding ground state PES is similar in all the cases whereas the differences is in the case of $S_1$ PES. Two minima regions were characterized in the $S_1$ PES of EAL namely MLCT and LF. The first minimum localized at Co–C equals 2.10 Å, and Co–N$_{Im}$ at 2.30 Å corresponds to an MLCT state and the second minimum, localized at the Co–C bond length of 2.50 Å and Co–N$_{Im}$ bond length of 2.90 Å, is characterized as the LF state (Figure 3.10). Comparing all the PESs, some differences and similarities were observed. The MLCT minimum in base-on AdoCbl is lower in energy than the LF minimum, whereas in the case of base-off AdoCbl, the MLCT is higher in energy than the LF minimum. In base-on AdoCbl, the seam separating the minima is more pronounced which indicates the high
barrier between MLCT and LF minima. Alternatively, the $S_1$ PES of base-off AdoCbl and the $S_1$ PES of AdoCbl inside EAL correspond to a very similar topology, where the LF region is energetically more stabilized (Figure 3.10). It appears that the enzymatic environment plays a role to stabilize the LF electronic state.

**Figure 3.10.** Potential energy surfaces for the $S_0$ with vertical projections of the $S_1$ state plotted as functions of axial bond lengths along with the corresponding $S_0$ and $S_1$ contour PESs with color codes for (a) AdoCbl inside EAL (b) $\text{Im}^-[\text{CoIII(corrin)}]^-\text{Ado}^+$ base-on model complex in water and (c) $\text{H}_2\text{O}^-[\text{CoIII(corrin)}]^-\text{Ado}^+$ base-off model complex in water.
In order to get more insight, the S₁ PES inside EAL were further scrutinized. Different minima regions were characterized based on the contributions of the molecular orbitals associated with each electronic transition. The minima regions were characterized as (I (S₁min), IIA (S₁min), and IIIB (S₁min)) and shown in figure 3.11a. The corresponding ground state geometries and detail orbital pictures for each of these minima regions (I (S₁min), IIA (S₁min), and IIIB (S₁min)) are depicted in Figure 3.12, A3, A4.

Figure 3.11. (a) Scheme of photoreaction of AdoCbl inside EAL on the S₁ PES with minima regions, separated by a seam, where MLCT minimum is denoted as I (S₁min) and the LF minimum is denoted as IIIB (S₁min) and (b) minimum energy paths (Path A and B) between MLCT and LF state plotted as a function of energy versus Q value.

The transition contributions in the minimum of the MLCT region, are 52% \( d_{xz} + \pi \rightarrow \pi^* \) and 33% \( d_{yz} + \pi \rightarrow \pi^* \) (Figure 3.12). For another MLCT minima (IIA (S₁min), Figure 3.12), the HOMO to LUMO \( (d_{xz}/d_{z^2} + \pi \rightarrow \pi^*) \) transition is dominant (Figures 3.11 and 3.12). The excitations at IIIB (MECP) are HOMO to LUMO + 1 and have been characterized as \( d_{yz} + \pi \rightarrow d_{z^2} \) (Figure 3.12). In the LF state, the HOMO to LUMO
transition is dominant. The excitation contribution for the LF state is $100\% \ d_{yz} + \pi \rightarrow d_{z^2}$ at IIIB ($S_{1\text{min}}$) (Figure 3.12). Apart from the character of the excitations, the main differences for the MLCT ($I(S_{1\text{min}})$) and LF (IIIB ($S_{1\text{min}}$)) minima are in their relative energy. Namely, the LF region is 1.2 kcal/mol lower than the MLCT, which is consistent with the experiment.

Figure 3.12. HOMO and LUMO molecular orbitals involved in electronic excitations corresponding to selected points on the $S_1$ PES (Figure 3.11a) along Path A and Path B.

3.3.4 Analysis of long-range charge-transfer excitations. As has been already pointed out in Computational Section 2.6, TD-DFT calculations often suffer from the potential underestimation of the LR-CT type excitations. Taking into account that the BP86 functional was used in calculations, such underestimation could have a potentially negative impact on the proper description of the $S_1$ PES. Specifically, for isolated AdoCbl, the LR-CT type excitations are typically associated with electronic transitions between the Ado and corrin ligands. To make sure that this is not the case, careful analysis of electronically excited states has been carried out. For the $S_1$ surface, the corresponding orbital contributions were analyzed and additional TD-DFT/MM calculations with different
QM/MM partitions were performed. The analysis was carried out for the S₁ PES based on TD-DFT/MM calculations with respect to the Ado → corrin transitions. When comparing the S₁ surface where the upper ligand was only Rib (Figure 3.13), to the S₁ surface with the full Ado ligand (Figure 3.10), it was shown that the topology of the two surfaces look very similar, although the relative energies of the MLCT and LF minima are slightly different.

![Potential energy surfaces for the singlet ground S₀ and S₁ states for RibCbl inside EAL. The structure of RibCbl is shown in figure 3.4. This partition of the high layer system was made to make sure the corresponding electronic states are not suffered from the LR-CT excitations.](image)

**Figure 3.13.** Potential energy surfaces for the singlet ground S₀ and S₁ states for RibCbl inside EAL. The structure of RibCbl is shown in figure 3.4. This partition of the high layer system was made to make sure the corresponding electronic states are not suffered from the LR-CT excitations.

For the AdoCbl in EAL S₁ PES (Figure 3.10), the sections where axial bonds were noticeably more extended than the equilibrium geometry, additional analysis of LR-CT transitions was also performed using the Dₜₜ diagnostic tool. The detail discussion about the DCT tools was discussed in the *computational detail, section 2.6*. The excited states involving charge transfer that was higher than 6.0 Å Dₜₜ length were not included in the construction of the S₁ surface. Taking all the above considerations into account we have confidence that the S₁ surface for AdoCbl-dependent EAL based on TD-DFT/MM methodology was correctly constructed.
3.3.5. Pathways for photodissociation. Photolysis of AdoCbl inside EAL to generate Co(II)/Ado radical pair should primarily occur from the LF state, similar to the base-on or base-off forms of AdoCbl (Figure 3.11). It is also expected that the photodissociation mechanism for AdoCbl in EAL should not significantly differ from free AdoCbl in solution because for both scenarios, the cleavage of the Co-C bond occurs from the LF region on the $S_1$ state. From the LF region, regardless of environment, either geminate recombination or radical pair separation is possible, although the energetics and the quantum yield of photoproduts associated with enzyme bound AdoCbl should be different from AdoCbl in solution. In contrast to base-on AdoCbl in solution, where the MLCT state is energetically more stable, inside EAL, the LF state is stabilized by 1.2 kcal/mol more than the MLCT state, which is in line with the base-off AdoCbl in solution (Figures 3.10). The main challenge in describing the photodissociation mechanism is connecting these two states via a minimum energy path. Previous theoretical studies for various Cbls have revealed that, based on energetic grounds, two possible photodissociation pathways, referred to as Path A and Path B, can be distinguished. Path A involves the initial elongation of the Co-C bond followed by Co-$N_{Im}$, while Path B involves the lengthening of Co-$N_{Im}$ bond prior to the elongation of Co-C (Figure 3.11). Likewise, the photodissociation mechanism of AdoCbl inside EAL can also be associated with two paths. Both pathways start from the energy minimum of MLCT state and proceed through the seam to the LF state, from which final photo-cleavage takes place. Path A initiates from the MLCT minima I and proceeds to the IIA of the MLCT minima and then overcomes the barrier of the seam to go to the LF state. Path B also initiates from the MLCT minima I and proceeds through the seam IIB to go into the LF minima at IIB (Figure 3.11). To determine which pathway is active, the
energetic barriers between the minima and the seam were considered. The energy barriers were plotted as a function of energy vs. the Q value

\[ Q = \sqrt{R_{Co-C}^2 + R_{Co-Nim}^2} \]

For Path A, the energy barrier is 3.48 kcal/mol at Q value 3.37, and for Pathway B it is 2.48 kcal/mol at Q value 3.47 (Figure 3.11b). Considering the energy profiles of both paths, we identified that Path B is energetically more favorable and should be considered as the active path for the photodissociation. Previous studies have shown that in the case of base-on AdoCbl in solution, Path A is active due to the low energy barrier where photodissociation initiates from MLCT state and proceeds through the elongation of Co-C bond to LF state, whereas in the base-off AdoCbl, Path B is active for photodissociation and this is more in line with our study in the enzymatic environment.42,176,214 In both cases, i.e. for base-off isolated AdoCbl as well as in the enzymatic environment, the LF state is more stabilized than MLCT state.

3.3.6. Implications of AdoCbl photochemistry inside EAL. When comparing the S1 PESs of AdoCbl inside EAL to isolated AdoCbl in solution, it is of note that the AdoCbl S1 PES adopts features from both of the S1 PESs for base-on and base-off forms of isolated AdoCbl (Figure 3.14). The topology of the S1 PES for EAL and base-on AdoCbl in solution are very similar as both contain two energy minima. However, those two surfaces are not identical from an energetic point of view. The LF state is lower in energy for AdoCbl inside EAL than in the isolated base-on AdoCbl form. When comparing the S1 PES of AdoCbl in EAL to the base-off AdoCbl form, the topologies are not as similar, but the energetics are more in line with each other. It seems as though the S1 PES of AdoCbl
in EAL can be considered as a superposition of the base-on and base-off $S_1$ PESs of the isolated cofactor.

**Figure 3.14.** Potential energy surfaces of $S_1$ state as function of axial bond lengths with photoreaction pathways depicted by red arrows for (a) AdoCbl inside EAL (b) Im–[CoIII(corrin)]–Ado base-on model complex in water and (c) H$_2$O–[CoIII(corrin)]–Ado base-off model complex in water.

The optimum energy paths for photolysis in the isolated base-on and base-off AdoCbl are different, with Path A being active for the base-on and Path B active for the base-off form (Figure 3.14). For the photolysis of AdoCbl inside EAL, the minimum energy path is a mixture of the optimum paths for the two forms of isolated AdoCbl. Dissociation ultimately occurs through the LF region, as is the case of AdoCbl in solution. Overall, based on energetics, Path B is more favorable than Path A by 1.0 kcal/mol for
AdoCbl in EAL (Figure 3.11b). Path B begins with only the elongation of the Co-N_{Im} of the axial base followed by simultaneous elongation of both axial bonds, Co-C and Co-N_{Im} until the LF minimum (IIIB) is reached. Although Path B is slightly more energetically favorable based on our calculations, we cannot rule out that Path A may also be active as well.

3.3.7. Comparison with experiment. AlkylCbls, such as CH_{3}Cbl or AdoCbl, upon excitation with visible or near UV light, generally exhibit rapid internal conversion (IC) to the lowest excited electronic state followed by cleavage of the Co-C bond. Geminate recombination limits the ultimate photolysis yield. Following the experimental work of Sension et al., and other groups including Warncke, et al., and Scrutton, et al, our current understanding of the photochemical behavior of AdoCbl can be summarized as depicted in figure 3.15. Excitation at 400 or 520 nm results in rapid IC to the lowest excited electronic state S_{1} and the nature of the S_{1} state observed in the experiments depends on the environment of AdoCbl. Our TD-DFT/MM calculations indicate that in the case of AdoCbl-dependent EAL this initial excitation state can be associated with the S_{4} state as the transition dipole moment is noticeably more substantial (i.e., order of magnitude) when compared to other energetically similar low-lying excited states (Note that the S_{2} state has a noticeable value of oscillator strength, although it has low energy versus experiment). The S_{4} state, with an excitation wavelength of 530 nm, has the most significant contribution (65%) coming from HOMO-4 → LUMO electronic excitation and can be described as a d_{xz} + \pi \rightarrow \pi^* transition (Table 3.3). Calculated excitation energy to the S_{4} state corresponds well to the maximum of experimental \alpha/\beta-band with \lambda_{max} of 527 nm.
The rapid conversion from the initial excited $S_4$ to a species that resembles the Co(II) on the $S_1$ PES involves several photoproduct intermediates, which can be readily characterized by our TD-DFT/MM calculations (figure 3.15). The first intermediate, denoted as AdoCbl(=MLCT), is the MLCT state. There is subsequent relaxation to another intermediate, AdoCbl(=LF), which is the LF state. These changes can be associated with the gradual elongation of axial bond distance. The LF state is primarily responsible for homolytic cleavage of the Co-C bond, which produces the Co(II)/Ado radical pair. At this point, according to the experiment, there is a competition between cage escape and geminate recombination with noticeable differences when these processes are compared between solvent and enzyme.$^{152,177}$ This competition dictates the quantum yield of photoproducts. When AdoCbl is in water, 75-80% of RP formed in the LF region (AdoCbl(=LF) Figure 3.15) tend toward geminate recombination on a timescale less than 1 ns while the others proceed towards diffusion and cage escape. The tendency toward geminate recombination is increased in the enzymatic environment to 90-95%. Alternatively, in solution, the geminate radical pair exists within a solvent cage, and the radical pair can diffuse apart from the jointly occupied solvation shell in a process known as cage escape. In an enzymatic environment, the quantum yield is significantly lower than the isolated cofactor in solution because of the higher recombination yield of Co(II)/Ado radical pair to Co(III). All this can be monitored experimentally by UV-vis spectroscopy, but it is difficult to model computationally without proper analysis of kinetic equations.

The additional explanation requires consideration of the excited state photoproduct relaxation to the ground state and the differences associated with whether or not the cofactor is in solution or an enzyme. For isolated AdoCbl in solution, two possible
channels for deactivation have been proposed.\textsuperscript{42,214} One involves the corrin ring distortion after detachment of axial lower ligand (base or water) while the second involves electronic de-excitation of Co(II). Based on current TD-DFT/MM calculations we postulate that only the second one is operational in the case of AdoCbl-dependent EAL. In the base-off form of AdoCbl the intermediate product is in the LF excited state (Figure 3.15). From this state, conversion to the ground state can occur by bending the corrin ring, similar as in the base-off form of CH\textsubscript{3}Cbl. This mechanism is possible when cobalt is in the Co(III) state with the axial CH\textsubscript{3} or Ado ligand attached. Deactivation takes place with Co-C shortening and the corrin ring bending. In the base-on form, the deactivation mechanism is different. Co-N and Co-C bonds are significantly elongated. On the axial axis N\textsubscript{ax}---Co---C\textsubscript{ax} system acquires the character of a radical pair Co---C, where Co(II) is in the excited state of low energy which can undergo fast deactivation to the ground state, followed by recombination of Co\textsuperscript{*} and R\textsuperscript{*} and subsequent connecting of the axial base.

![Figure 3.15. Mechanism of AdoCbl photolysis in EAL and post-homolysis photophysical events (cage escape, internal conversion), as described in Section 3.3.7](image)

To summarize, we believe there are at least three important things to take into account when considering the formation and recombination of radical pair in AdoCbl in
solution and an enzyme including, ligand field stabilization, cage effect, and the mechanism of deactivation to $S_0$. Based on energetic grounds, it is easier to generate radical pair in solution than in an enzyme because the LF is stabilized in the protein environment. However, the rate of recombination is faster in the solution medium. In solution, the cage effect is a more significant contributor to the recombination yield and the quantum yield of photoproducts. The solution medium allows more freedom for diffusive loss of the radical pairs from the solvent cage. There is less opportunity for diffusion in an enzyme. There are two mechanisms of deactivation where the geminate radical pairs can recombine to the $S_0$. The first involves corrin ring distortion that is more in line with base-off Cbls in solution. The second involves deactivation of Co(II) species to the $S_0$ from the LF by de-excitations from the $(d_{yz})^1(d_z^2)^2$ excited state returning to a Co(II) $(d_{yz})^2(d_z^2)$. For the case of AdoCbl in EAL, it would seem that de-excitations of Co(II) from the LF is the active deactivation route.

3.3.8. New perspective for photochemistry of AdoCbl inside EAL. For AdoCbl inside EAL, the bond dissociation energetic barrier of $\sim30$ kcal/mol reported for isolated AdoCbl, is overcome to form the Co(II)/Ado radical pair. Warncke et al. has suggested that this photolysis is a good mimic for cleavage of the Co-C bond in EAL.\(^{168}\) To further connect photochemistry of AdoCbl-dependent EAL in relation to the native reaction, Warncke et al. hypothesized that such connection would require the explicit introduction of the protein structural changes that can be associated with Co-C bond cleavage and RP separation.\(^{168}\) Such protein structural changes can be represented by using a protein configuration coordinate that is orthogonal to the coordinate representing Co-C bond stretch. Consequently, a simple two-dimensional free energy PES representative of $S_0$ was
proposed and characterized by two minima, Co(III)Ad and Co(II)Ad•. This model was the basis for the identification of two paths, Path 1 and 2, for radical pair formation in EAL. Path 1 is associated with photolysis and initiates in the Co(III)Ad minimum region and proceeds as Co-C lengthens while the protein configuration is unchanged. Path 2 corresponds to native thermal cleavage of the Co-C bond and is a function of two coordinates, the Co-C bond length and protein configuration. Path 2 initiates from the Co(II)Ad region until reaching the Co(II)Ad• region where dissociation takes place. Based on our calculations, we have determined that photolysis of AdoCbl inside EAL is more in line with Path 2 on the Warncke model in that photolysis involves two minima regions not one.

Taking the above information and the outcome of present TD-DFT/MM calculations into consideration we can offer a different interpretation in regard to the photochemistry of AdoCbl inside EAL. First, the photolysis of AdoCbl does not occur on the S₀ PES. It is apparent from inspection of Figure 3.7 and 3.8, that the photolysis of AdoCbl, regardless of environment, cannot be described by considering changes in only one coordinate. Upon excitation, the most significant structural changes that occur are in axial bonding whereas corrin macrocycle is only slightly perturbed. This cooperativity between the axial bond distances has also been shown experimentally with XANES measurements and with several other theoretical studies. Based on this evidence, the most critical coordinates to consider in AdoCbl photolysis are the axial bond lengths. While we agree that two coordinates are necessary to describe the formation of Co(II)/Ado Radical pair in EAL via photolysis, we suggest that the second coordinate should be the Co-Naxial bond distance instead of protein configuration.
Based on Figure 3.10, which includes $S_1$ PESs as a function of axial bond lengths, it is apparent that regardless of environment, solvent or enzyme, the $S_1$ PES of AdoCbl is always characterized by two energy minima associated with MLCT and LF electronic states where photolysis occurs from the LF state. Thus, the electronic properties associated with AdoCbl are the primary component of photolysis which can be fully described in terms of the cofactor and its axial bonding. In our opinion, protein structural changes do not play a significant role in the photolysis of EAL. This is also apparent in the Warncke model, where the protein configuration coordinated is depicted as unchanged in the photolysis pathway.\textsuperscript{168} Instead, we believe that the enzymatic environment has the important role of stabilizing the LF state on the $S_1$ surface. The other important role of EAL enzyme in terms of photochemistry may be associated with specific geometric constraints imposed on AdoCbl. In contrast, the solution does not impose these geometric constraints and upon dissociation of the Ado ligand, the radical pair can be separated freely from the corrin moiety as evidenced by the higher quantum yield of radical pair compared to the quantum yield in the enzyme.

While it is true that in the native reaction there are conformational changes that affect the reactivity of AdoCbl inside EAL, it is vital for us to re-iterate that our model is not an exact replication of the native reaction. Instead, our model is representative of photochemistry. Protein contributions were explored in our model through molecular mechanics (MM), not as a specific coordinate on the PESs. However, based on our combined QM/PM6/MM study, we have observed that the protein environment is affecting the energetics associated with photolysis as compared to AdoCbl in solution. Inside EAL, the LF, where Co-C bond cleavage occurs, is energetically lower compared to the MLCT
state. In contrast, in solution, the LF is higher in energy than the MLCT. These observations indicate that the protein environment affects the $S_1$ state and this is being accounted for in our model system.

3.4. Summary and conclusions.

Simple mechanistic details of AdoCbl photochemistry inside EAL are presented here. The photochemical data is consistent with the previous theoretical studies of isolated AdoCbl, where it was illustrated that there are significant similarities in the topologies of $S_0$ and $S_1$. The $S_1$ surface is characterized by two minima, namely MLCT and LF. The LF is 1.0 kcal/mol less in energy than the MLCT state and hence it is more stabilized, which is consistent with previous experimental work in AdoCbl-dependent enzymes. Two possible paths for photodissociation were identified to explain the mechanism of photolysis. Path A is along the Co-C bond distance and is inactive due to the higher energetic barrier. Path B is more energetically feasible as it can surpass the lower energy barrier from the MLCT to the LF state. To understand the effect of the environment on the formation of Co(II)/Ado radical pair, the photochemical data for AdoCbl in EAL was also compared with base-on and base-off AdoCbl in solution. Some similarities and differences were noted in the topology of the $S_1$ surface. The $S_1$ surface for AdoCbl in EAL was more in line with the isolated base-off cofactor which is confirmed with the experimental findings. Although the photochemistry of EAL is not physiological necessary for bacteria, these findings will shed light on the still elusive mechanistic details associated with AdoCbl-dependent enzymes. In addition to this, we expect that our study can help to understand the reactions that are induced by light.
CHAPTER IV
PHOTOLYTIC CLEAVAGE OF CO-C BOND IN COENZYME B\textsubscript{12}\-
DEPENDENT GLUTAMATE MUTASE

4.1 Introduction

Coenzyme B\textsubscript{12} also known as AdoCbl is a critical biologically active derivative of
vitamin B\textsubscript{12} (Figure 1.1).\textsuperscript{1,10,13,36} It acts as a cofactor in glutamate mutase (GLM) enzyme
to catalyze the interconversion of (S)-glutamate to (2S,3S)-methyl aspartate.\textsuperscript{54,55,81,96} GLM
belongs to the mutase class of the AdoCbl-dependent enzyme found in an anaerobic
bacterium that uses glutamate (GM) as a source of energy and carbon by degrading GM
into CO\textsubscript{2}, NH\textsubscript{4}, H\textsubscript{2}.\textsuperscript{70,96} In order to catalyze this reaction, GLM requires the AdoCbl
cofactor, which binds in the base-off mode where a histidine residue replaces the lower
axial DBI ligand.\textsuperscript{3,6,7,10,79} This is a highly unfavorable reaction without the presence of
an enzyme and has many implications in industrial processes. A detailed understanding of
this reaction catalyzed by GLM allows designing new biomimetic catalysts that could be
used for many industrial purposes.

Like other AdoCbl-dependent enzymes, homolytic cleavage of the Co-C bond
initiates the catalytic cycle, resulting in the formation of the Co(II) and Ado radical pair
(Figure 4.1a). Upon substrate (GM) binding, the rate of cleavage is accelerated by 10^{12}
-fold in the enzyme when compared to the homolysis of free AdoCbl in solution.\textsuperscript{3,4,6,49,54,87-
Figure 4.1. a) Mechanism of Co-C bond cleavage in AdoCbl-dependent GLM. The substrate is S-glutamate. The reaction catalyzed by GLM is the conversion of S-glutamate to (2S,3S)-3-methylaspartate b) Photolytic cleavage of Co-C bond in AdoCbl-dependent GLM without the binding of a substrate.

After the homolytic cleavage, Ado radical abstract the H-atom from the substrate GM and generate substrate radical, which undergoes further rearrangement to form product radical. The product radical then removes an H-atom or undergoes an elimination-addition pathway or a fragmentation recombination mechanism to generate the products CO₂, NH₄, H₂ (Figure 4.1a). This issue is still unclear and requires more studies. On the other hand, the Co-C bond can also be cleaved by light without substrate binding to generate the same radical pair as in the case of enzymatic catalysis. Upon excitation with light at 520 or 400 nm, it results in a rapid internal conversion from the initially excited high energy state to the S₁ state. Photodissociation will then take place from the S₁ state. The mechanism of Co-C bond photolysis in AdoCbl-dependent GLM has been studied by Sension and co-workers using ultrafast transient absorption spectroscopy (TAS). It has been shown that the photodissociation of the Co-C bond in AdoCbl differs significantly inside the enzyme when compared to a solution. In AdoCbl-dependent GLM, radical pair generation occurs from a low-lying excited state intermediate, which has a longer lifetime in comparison to the solution medium. The protein environment also
increases the rate of geminate recombination when compared to the rate in solution. Accordingly, the quantum yield (QY) of the photogenerated radical pair decreases in AdoCbl-dependent GLM. The enzyme appears to prevent the diffusive loss of the alkyl radical.

To understand the mechanism of the radical pair generation in more depth, a proper description of the low-lying excited states based on the theoretical analysis, in addition to experimental work, is needed. With this in mind, we employed a combined (QM/MM) method to understand the photochemical properties of the AdoCbl cofactor bound to GLM. The motivation of our work is to understand the mechanism of photo-induced cleavage of the Co-C bond inside GLM and compare the results with the mechanism for isolated AdoCbl cofactor as well as EAL. Although the Co-C bond in GLM is not physiologically cleaved by light, the photolytic rupture of the Co-C bond can serve as a probe for the activation of the Co-C bond in the native catalytic cycle. This, in turn, may have implications for understanding the \(~10^{12}\) rate enhancement of Co-C bond homolysis in AdoCbl-dependent enzymes. Furthermore, these insights can also be used to further understand the unique photochemical properties of biological photoreceptors as well as other cobalamin dependent enzymes.

4.2 Computational details

The crystal structure of AdoCbl-dependent GLM was obtained from the Protein Data Bank (PDB ID: 1I9C). The crystal structure contains two identical asymmetric units which are independent and at a distance from each other (Figure 4.2a). One asymmetric unit contains two subunits (α and β). Only one unit was used for the calculations (Figure 4.2b). In the crystal structure, two conformations of the ribose ring of the upper axial ligand
were reported and are denoted as 2'-endo and 3'-endo. The different conformations are the result of the displacement of the two ribose ring atoms, C2’ or C3’, towards the out of plane C5’ atom for the 2’-endo and 3’-endo conformations, respectively. The details regarding the orientation of ribose have been explained in chapter 3, figure 3.6

**Figure 4.2.** a) Crystal structure of AdoCbl-dependent GLM (PDB ID:1I9C) obtained from the protein data bank. The structure contains two independent asymmetric units. One asymmetric unit contains α and β subunit as shown in the figure. The α subunit depicted in light blue color ribbon and the β subunit depicted as a gray color. The cofactor binds in each of the asymmetric units in the interface of α and β subunit. The surface and the green color depict the cofactor. b) Structure of one asymmetric unit which is used for the calculations. c) Optimized structure of GLM obtained from the QM/MM calculations.

We used the 2'-endo structure to model the photolytic reaction because this conformation was reported to be present prior to the H-atom abstraction step in the native reaction. The crystal structure also contained the substrate molecule GM. In order to represent photophysical processes, which do not require substrate binding to activate the Co-C bond,
the substrate GM, was removed. Six water molecules were used to occupy the cavity left by the removal of the substrate. The structure was then protonated using the PropKa program.\textsuperscript{201,202} The axially coordinated His16 residue was protonated at N\textsubscript{δ1} position. The model system was then minimized at MM level of theory using the AMBER force field (FF14SB).\textsuperscript{203,206} The total system was minimized with the steepest descent algorithm for 5000 steps. The MM minimization was carried out using the Amber suite of programs.\textsuperscript{205} The minimized structure was used to initiate the QM/MM calculations.\textsuperscript{192,194}

In order to perform QM/MM calculations, the GLM model system was partitioned into two layers. The truncated cofactor without the side chains of the corrin, and the Im base from His16 was placed into the high layer and the rest of the cofactor such as the side chains of the corrin ring, the protein residues, and the six water molecules were placed into the low layer. The high layer part was treated with DFT/BP86 for the ground state and TD-DFT/BP86 for the excited states. The low layer part was treated with MM using the AMBER force field (FF99SB). Changes in atom positions were allowed to a radius of 20 Å from the Co (3542 atoms) and the remainder of the protein was frozen (7439 atoms). In total, the model system contained 10,981 atoms.

BP86/MM level of theory with TZVP basis set for H and the TZVPP basis set for the N, C, O and Co atoms was applied for the ground state calculations.\textsuperscript{216,217} Similar method has been used for AdoCbl-dependent EAL and isolated AdoCbl cofactors.\textsuperscript{42,176,178-180,184,232} The optimized equilibrium structure was used to construct the \textit{S}_0 potential energy curve (PEC) and the \textit{S}_0 potential energy surface (PES). The \textit{S}_0 PEC, as a function of Co-C bond length, was computed by systematically stretching the Co-C bond from 1.8 Å to 2.6 Å with a step size of 0.05 Å. On the other hand, to construct the \textit{S}_0 PES two axial bond
lengths were used, such as Co-C and Co-N\textsubscript{Im}. For low-lying excited states, TD-DFT/MM approach were applied with the same level of theory. Using TD-DFT/MM, S\textsubscript{1}-S\textsubscript{n} PEC were constructed that were based on the geometries obtained from the S\textsubscript{0} PEC. A similar approach was applied to construct the S\textsubscript{1} PES. Various points on the S\textsubscript{1} PES were characterized by thorough orbital analysis to ensure that all states were properly selected. In order to make sure that our calculations were not affected by LR-CT states, a detailed orbital analysis was performed and only the proper transitions were selected. The D\textsubscript{CT} diagnostic test was also applied to confirm the selections and will be discussed in the results and discussions section.\textsuperscript{225,226}

4.3 Results and discussions

4.3.1. Ground state geometry. Geometry optimizations were carried out for the structural model where six water molecules replaced the substrate. To confirm that the geometry was optimized to a stationary point, frequency calculations were also conducted. No imaginary frequencies were present. The optimized structures were then compared with the crystal structure of GLM (Figure 4.3, Table 4.1). In the case of the GLM optimized structure, the Co-C and Co-N\textsubscript{Im} bond lengths are 2.08 Å and 2.27 Å, respectively, whereas in the GLM crystal structure the bond lengths are 3.17 Å and 2.22 Å, respectively. The Co-C bond distance is longer in the crystal structure because it is representative of the geometry just prior to the hydrogen abstraction step. The optimized structure of GLM was then compared with the optimized structure of EAL. In EAL, the Co-C and Co-N\textsubscript{Im} bond lengths are 2.05 Å and 2.37 Å, respectively (Figure 4.3, Table 4.1).\textsuperscript{178,179} The difference in the Co-N\textsubscript{Im} lengths between the GLM and EAL optimized structures has to do with the binding mode
of the cofactor. In EAL, the axial base is DBI from nucleotide loop of B\textsubscript{12} system, whereas in GLM, the lower axial ligand is replaced by an Im group from His\textsubscript{16} residue.

Figure 4.3. a) Crystal structure of GLM obtained from the Protein Data Bank (PDB ID:1I9C). The structure contains two conformers, A and B, both are depicted. In conformer A, the ribose orientation is 3′-endo, and it represents the position after H-abstraction. The Co-C bond distance in conformer A is 4.19 Å. In conformer B, the ribose orientation is 2′-endo, and it represents the position before H-abstraction. The Co-C bond distance in conformer B is 3.17 Å. b) Optimized structure of GLM without the substrate. Coordinates of conformer B were used in the model system and the substrate molecule was replaced with six water molecules to fill the empty space left from the removal of the substrate molecule.

The observed differences in Co-N\textsubscript{Im} bond length is the result of His being more basic than DBI. The substrate was removed from the GLM model structure in our calculations to represent photoinduced cleavage. There is no crystal structure for substrate-free GLM, so calculations must be relied on. That being said, it was important to consider any effect that the removal of the substrate might have on the structure of AdoCbl inside
GLM. Accordingly, we optimized the model structures of GLM with and without the substrate. Differences were observed between the optimized structures for the substrate-free and substrate-bound models with respect to the Ado moiety (Figure 4.3, Table 4.2). The ribose orientation was found to be 3′-endo in the substrate-free model (Table 4.2, Figure 4.3 and A1).

<table>
<thead>
<tr>
<th>Structural parameters</th>
<th>Crystal structure of GLM (Conformer B, PDB ID-1I9C)</th>
<th>Optimized structure of GLM without substrate</th>
<th>Optimized structure of EAL without substrate</th>
<th>Optimized structure of GLM with substrate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bond length (Å)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-C</td>
<td>3.17</td>
<td>2.08</td>
<td>2.053</td>
<td>2.11</td>
</tr>
<tr>
<td>Co-Nim</td>
<td>2.22</td>
<td>2.27</td>
<td>2.372</td>
<td>2.31</td>
</tr>
<tr>
<td>Co-N21</td>
<td>1.88</td>
<td>1.92</td>
<td>1.934</td>
<td>1.92</td>
</tr>
<tr>
<td>Co-N22</td>
<td>2.02</td>
<td>1.94</td>
<td>1.962</td>
<td>1.94</td>
</tr>
<tr>
<td>Co-N23</td>
<td>1.93</td>
<td>1.94</td>
<td>1.942</td>
<td>1.95</td>
</tr>
<tr>
<td>Co-N24</td>
<td>1.90</td>
<td>1.90</td>
<td>1.914</td>
<td>1.89</td>
</tr>
<tr>
<td>Bond Angle (°)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-Co-Nim</td>
<td>170.5</td>
<td>163.8</td>
<td>171.5</td>
<td>164.3</td>
</tr>
<tr>
<td>N21-Co-C</td>
<td>90.8</td>
<td>101.9</td>
<td>92.6</td>
<td>101.9</td>
</tr>
<tr>
<td>N22-Co-C</td>
<td>80.4</td>
<td>88.3</td>
<td>94.7</td>
<td>89.5</td>
</tr>
<tr>
<td>N23-Co-C</td>
<td>82.5</td>
<td>81.8</td>
<td>93.1</td>
<td>83.8</td>
</tr>
<tr>
<td>N24-Co-C</td>
<td>92.4</td>
<td>92.0</td>
<td>84.4</td>
<td>93.2</td>
</tr>
<tr>
<td>Torsional Angle (°)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N21-N22-N23-Co</td>
<td>5.27</td>
<td>-2.52</td>
<td>-4.4</td>
<td>-3.94</td>
</tr>
<tr>
<td>N21-N22-N23-N24</td>
<td>-0.09</td>
<td>-3.06</td>
<td>-3.3</td>
<td>-2.63</td>
</tr>
<tr>
<td>Co-N22-C6-C10</td>
<td>-9.44</td>
<td>-10.19</td>
<td>-3.9</td>
<td>-15.72</td>
</tr>
<tr>
<td>Co-N22-C6-C5</td>
<td>8.69</td>
<td>3.29</td>
<td>-1.0</td>
<td>4.36</td>
</tr>
</tbody>
</table>

Table 4.1. Selected geometric parameters like bond length, bond angles, torsional angles for the optimized structure of AdoCbl in GLM without the substrate, optimized structure of GLM with the substrate, the crystal structure of GLM (PDB ID:1I9C) and the optimized structure of EAL without the substrate. The optimized structures are shown in Figure 4.3 and Figure A5.
This finding is also consistent with previous studies where the ribose conformation was found to be 3'-endo in a substrate-free model of EAL and for a model of free AdoCbl cofactor in solution. In the case of the substrate-bound model of GLM, the orientation was 2'-endo. Although the substrate is not necessary for photo-cleavage of the Co-C bond, it is essential to note that these differences in the orientation of the ribose, based on substrate binding, may have implications for the native catalytic mechanism. However, future studies are required to explore this further.

<table>
<thead>
<tr>
<th>Structural parameters</th>
<th>Crystal structure of GLM (Conformer B, PDB ID-1I9C)</th>
<th>Optimized structure of GLM without substrate</th>
<th>Optimized structure of EAL without substrate</th>
<th>Optimized structure of GLM with substrate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exocyclic angle (°)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1'-C2'-O2'-O2'H (Φ2)</td>
<td>170.8</td>
<td>167.33</td>
<td>-30.25</td>
<td>51.65</td>
</tr>
<tr>
<td>C2'-C3'-O3'-O3'H (Φ3)</td>
<td>173.5</td>
<td>-6.68</td>
<td>9.110</td>
<td>39.43</td>
</tr>
<tr>
<td>O1'-C1'-N9'-C8' (ψCN)</td>
<td>17.04</td>
<td>-11.57</td>
<td>-72.46</td>
<td>-29.46</td>
</tr>
<tr>
<td>Endocyclic angle (°)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1'-C2'-C3'-C4' (θ0)</td>
<td>-45.21</td>
<td>1.64</td>
<td>9.87</td>
<td>-34.04</td>
</tr>
<tr>
<td>C2'-C3'-C4'-O1' (θ1)</td>
<td>29.28</td>
<td>-27.05</td>
<td>-35.86</td>
<td>4.15</td>
</tr>
<tr>
<td>C3'-C4'-O1'-C1' (θ2)</td>
<td>0.73</td>
<td>44.01</td>
<td>49.72</td>
<td>29.97</td>
</tr>
<tr>
<td>C4'-O1'-C1'-C2' (θ3)</td>
<td>-29.83</td>
<td>-42.76</td>
<td>-43.30</td>
<td>-52.47</td>
</tr>
<tr>
<td>O1'-C1'-C2'-C3' (θ4)</td>
<td>46.65</td>
<td>23.88</td>
<td>19.54</td>
<td>53.46</td>
</tr>
<tr>
<td>Pseudorotation phase (P)</td>
<td>152.7</td>
<td>87.91</td>
<td>78.48</td>
<td>129.2</td>
</tr>
<tr>
<td>Pseudorotation amplitude (θm)</td>
<td>51.37</td>
<td>52.90</td>
<td>49.42</td>
<td>53.60</td>
</tr>
<tr>
<td>Ribose conformation</td>
<td>2'-endo</td>
<td>3'-endo</td>
<td>3'-endo</td>
<td>2'-endo</td>
</tr>
</tbody>
</table>

Table 4.2. Selected geometric parameters like exocyclic angles and endocyclic angle of the Ado ligand to determine the ribose orientation using pseudorotation phase (P) and pseudorotation angle (θm) as shown in Figure 3.6, for the optimized structure of AdoCbl in GLM without the substrate, optimized structure of GLM with the substrate, crystal structure of GLM (PDB id-1I9C) and the optimized structure of EAL without the substrate. The optimized structures are shown in Figure 4.3 and Figure A5.

88
4.3.2 PECs as a function of Co-C bond length. To simulate the photodissociation mechanism of the Co-C bond in AdoCbl-dependent GLM, vertical excitation energies of the low-lying singlet and triplet states were calculated using TD-DFT/MM along the Co-C bond length. PECs were constructed by systematically elongating the Co-C bond from 1.8 Å to 2.6 Å using the step size of 0.05 Å. PECs were then compared with AdoCbl-dependent EAL. Some similarities and differences were noted (Figure 4.4).

![Figure 4.4](image)

Figure 4.4. Potential energy curve (PEC) of the ground state (S₀) as a function of Co-C bond length with the vertical projection of low-lying triplet states (blue color) and singlet states (red color) for (a) AdoCbl-dependent GLM and b) AdoCbl-dependent EAL. The repulsive triplet state is indicated as $3\pi \rightarrow \pi^*$. 

The most obvious commonality in both the enzymes is the repulsive triplet state which is marked by a sharp drop in energy. For GLM and EAL, this repulsive triplet state is lowest in energy after the Co-C bond length is 2.2 Å and 2.4 Å, respectively. The repulsive nature of this triplet state is a result of the $\sigma \rightarrow \sigma^*$ character. This triplet state does not have the proper dissociation limit because of the single-determinant approach in TD-DFT. The RHF
wavefunction becomes unstable at relatively short Co-C bond distances. This state would correctly level-off by applying a UHF wavefunction. The eight lowest vertical singlet excited states were characterized by the orbital analysis and shown in Table 4.3 and Figure 4.4.

<table>
<thead>
<tr>
<th></th>
<th>E[eV]</th>
<th>f</th>
<th>λ[nm]</th>
<th>D$_{CT}$</th>
<th>%</th>
<th>Character</th>
</tr>
</thead>
<tbody>
<tr>
<td>S$_1$</td>
<td>2.13</td>
<td>.0063</td>
<td>581</td>
<td>1.13</td>
<td>52</td>
<td>$d_{xy}/d_{z^2}-\pi \rightarrow \pi^*$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$d_{xy}^+ \pi \rightarrow \pi^*$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>35</td>
<td>$d_{yz}^+ \pi \rightarrow \pi^*$</td>
</tr>
<tr>
<td>S$_2$</td>
<td>2.31</td>
<td>.0213</td>
<td>537</td>
<td>0.44</td>
<td>44</td>
<td>$d_{xy}^+ \pi \rightarrow \pi^*$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>34</td>
<td>$d_{xy}/d_{z^2}^+\pi \rightarrow \pi^*$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>05</td>
<td>$d_{xy}/d_{z^2}^+\pi \rightarrow d_{z^2}$</td>
</tr>
<tr>
<td>S$_3$</td>
<td>2.46</td>
<td>.0233</td>
<td>504</td>
<td>0.58</td>
<td>43</td>
<td>$d_{xy}^+ \pi \rightarrow \pi^*$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>33</td>
<td>$d_{xy}/d_{z^2}^+\pi \rightarrow d_{z^2}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>07</td>
<td>$d_{xy}^+ \pi \rightarrow d_{z^2}$</td>
</tr>
<tr>
<td>S$_4$</td>
<td>2.51</td>
<td>.0032</td>
<td>493</td>
<td>0.94</td>
<td>88</td>
<td>$d_{xy}^+ \pi \rightarrow \pi^*$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>03</td>
<td>$d_{xy}^+ \pi \rightarrow \pi^*$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>03</td>
<td>$d_{xy}^+ \pi \rightarrow d_{z^2}$</td>
</tr>
<tr>
<td>S$_5$</td>
<td>2.56</td>
<td>.0346</td>
<td>483</td>
<td>1.08</td>
<td>35</td>
<td>$d_{xy}/d_{z^2}^+\pi \rightarrow d_{z^2}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>27</td>
<td>$d_{xy}^+ \pi \rightarrow \pi^*$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td>$d_{xy}^+ \pi \rightarrow d_{z^2}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>06</td>
<td>Ado $\pi \rightarrow \pi^*$</td>
</tr>
<tr>
<td>S$_6$</td>
<td>2.59</td>
<td>.0079</td>
<td>477</td>
<td>5.40</td>
<td>90</td>
<td>Ado $\pi +$Rib $\pi \rightarrow \pi^*$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>04</td>
<td>$d_{xy}/d_{z^2}^+\pi \rightarrow d_{z^2}$</td>
</tr>
<tr>
<td>S$_7$</td>
<td>2.65</td>
<td>.0075</td>
<td>466</td>
<td>0.21</td>
<td>58</td>
<td>$d_{xy}^+ \pi \rightarrow d_{z^2}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td>$d_{xy}/d_{z^2}^+\pi \rightarrow d_{xy}+n$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>09</td>
<td>$d_{xy}/d_{z^2}^+\pi \rightarrow d_{z^2}$</td>
</tr>
<tr>
<td>S$_8$</td>
<td>2.75</td>
<td>.0084</td>
<td>449</td>
<td>4.56</td>
<td>86</td>
<td>Ado $\pi +$Rib $\pi \rightarrow \pi^*$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>03</td>
<td>$d_{xy}^+ \pi \rightarrow d_{z^2}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>02</td>
<td>$d_{xy}^+ \pi \rightarrow \pi^*$</td>
</tr>
</tbody>
</table>

*Table 4.3.* Eight lowest vertical singlet excited states characterized by the detail orbital analysis. The orbital picture is shown in Figure 4.6 and the vertical singlet excited states along the Co-C bond length is shown in Figure 4.4.
The excitation energies to the $S_1$ and $S_2$ states are 2.13 eV and 2.31 eV, respectively, which are higher than the excitation energetics to lowest singlet states in EAL. The orbitals involved in $S_0 \rightarrow S_1$ and $S_0 \rightarrow S_2$ transitions are metal d orbitals to the corrin $\pi^*$, also named as MLCT transitions. This is consistent with the assignment of the electronic transitions in $S_1$ state using TAS. In higher excited states like $S_5$, $\sigma^*$ character transitions from metal d + corrin $\pi \rightarrow \sigma^*(d_{z^2})$ is more prominent than MLCT. Thus, the LF state is higher in energy than the MLCT state at equilibrium geometry. The vertical excitations for the $S_1$-$S_8$ states were also confirmed by $D_{CT}$ analysis (Table 4.3). All excited states considered have appropriate values from $D_{CT}$ analysis and the corresponding excitation energies are not underestimated.

4.3.3 $S_1$ PES as a function of axial bond lengths. It is evident from previous theoretical studies and also from XANES measurements for cobalamins that upon excitation, the most significant structural changes occur in axial bond distances, whereas the corrin macrocycle does not undergo significant structural changes. Thus, to characterize the intermediates associated with photodissociation, it is crucial to explore the excited state PES as a function of two axial bonds Co-C and Co-Nin. It has been shown theoretically and experimentally that photodissociation of the Co-C bond generates the RP from the $S_1$ state. The construction of the $S_1$ PES and the methodology applied were discussed in Computational Details, section 4.2.

The corresponding $S_0$ and $S_1$ PESs constructed for GLM were compared with the PESs calculated for EAL (Figure 4.5). The topology of the $S_0$ and $S_1$ PESs for GLM are comparable to EAL however, the energetics are different. In both cases, the $S_0$ PES contains a single energy minimum whereas the $S_1$ PES contains two minima regions. For
the S₁ PES, the two minima regions are characterized as the MLCT and LF states which are separated by a seam (Figure 4.5).

**Figure 4.5.** Potential energy surface (PES) of ground state (S₀) and first excited singlet (S₁) electronic states as a function of axial bond lengths, Co-C and Co-Nᵦ with the contour representation of the S₀ and S₁ electronic states along with color scales for energetics for a) AdoCbl-dependent GLM and b) AdoCbl-dependent EAL.

For the GLM S₁ PES, the MLCT minimum is localized at approximately Co-C 2.2 Å and Co-Nᵦ 2.1 Å bond lengths (Figure 4.5). The electronic transition associated with the MLCT is 87% dₓz,dₓ² + π → π*. Overall in the MLCT region, the dominant contribution are from metal d orbital + corrin π → corrin π* (Figure 4.5 and 4.6). The LF minimum is localized at bond distances of approximately Co-C 2.5 Å and Co-Nᵦ 2.7 Å. The electronic contribution in this region is characterized as 100% dᵧ + π → σ*(dₓ²). This minima region has a dominant metal d orbital → σ* contributions (Figure 4.6 and A6). Dissociation from this state is more likely because the excitation causes the weakening of the axial Co-C bond. Moreover, the LF state is lower in energy compared to the MLCT state.
As was discussed in the *Computational Details* section, the selection of the correct electronic state is crucial to generate a reliable \( S_1 \) PES for describing the photodissociation mechanism. It is crucial to ensure that the \( S_0 \rightarrow S_1 \) selected transitions did not suffer from potential LR-CT issues. Accordingly, thorough orbital analysis and \( D_{CT} \) calculations were performed to ensure that the \( S_1 \) PES was properly constructed. This will also ensure that the mechanism of photodissociation is properly determined. We have characterized several points along the potential pathways for photodissociation and performed \( D_{CT} \) calculations for these points to further ensure that the selected transitions were not the result of LR-CT.
The D<sub>CT</sub> results are shown in Table 4.4. The calculated negative t values in all the cases indicate the proper overlap between the occupied and virtual orbitals in all the transitions. Besides, the D<sub>CT</sub> value is also well below the threshold values reported in previous studies.

### Table 4.4

<table>
<thead>
<tr>
<th>Selected Points</th>
<th>f</th>
<th>λ [nm]</th>
<th>%</th>
<th>Character</th>
<th>D&lt;sub&gt;CT&lt;/sub&gt;</th>
<th>q</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Co-C</td>
<td>Co-N&lt;sub&gt;Im&lt;/sub&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLCT 2.20</td>
<td>2.10</td>
<td>.0142</td>
<td>606</td>
<td>H-2→L</td>
<td>87</td>
<td>d&lt;sub&gt;xz&lt;/sub&gt;/d&lt;sub&gt;y&lt;/sub&gt;&lt;sup&gt;2&lt;/sup&gt; + π → π*</td>
<td>1.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H-4→L</td>
<td>03</td>
<td>d&lt;sub&gt;y&lt;/sub&gt; + π → π*</td>
<td></td>
</tr>
<tr>
<td>MLCT 2.20</td>
<td>2.30</td>
<td>.0102</td>
<td>583</td>
<td>H-2→L</td>
<td>62</td>
<td>d&lt;sub&gt;xz&lt;/sub&gt;/d&lt;sub&gt;y&lt;/sub&gt;&lt;sup&gt;2&lt;/sup&gt; + π → π*</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H-3→L</td>
<td>25</td>
<td>d&lt;sub&gt;xz&lt;/sub&gt; + π → π*</td>
<td></td>
</tr>
<tr>
<td>MLCT 2.30</td>
<td>2.10</td>
<td>.0137</td>
<td>638</td>
<td>H-2→L</td>
<td>32</td>
<td>d&lt;sub&gt;xy&lt;/sub&gt; + π + ado π → π*</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H-1→L</td>
<td>58</td>
<td>d&lt;sub&gt;xz&lt;/sub&gt;/d&lt;sub&gt;y&lt;/sub&gt;&lt;sup&gt;2&lt;/sup&gt; + π → π*</td>
<td></td>
</tr>
<tr>
<td>LF 2.50</td>
<td>2.50</td>
<td>.0017</td>
<td>752</td>
<td>H-3→L</td>
<td>15</td>
<td>d&lt;sub&gt;xz&lt;/sub&gt; + π → d&lt;sub&gt;y&lt;/sub&gt;&lt;sup&gt;2&lt;/sup&gt;</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H-1→L</td>
<td>82</td>
<td>d&lt;sub&gt;y&lt;/sub&gt; + π → d&lt;sub&gt;x&lt;/sub&gt;&lt;sup&gt;2&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>LF 2.50</td>
<td>2.70</td>
<td>.0017</td>
<td>813</td>
<td>H-1→L</td>
<td>100</td>
<td>d&lt;sub&gt;y&lt;/sub&gt; + π → d&lt;sub&gt;x&lt;/sub&gt;&lt;sup&gt;2&lt;/sup&gt;</td>
<td>0.40</td>
</tr>
</tbody>
</table>

Table 4.4. Orbital characterization and D<sub>CT</sub> diagnostics for several points from the MLCT and LF regions of the S<sub>1</sub> PES.

#### 4.3.4 Photodissociation mechanism based on S<sub>1</sub> PES.

Radical pair generates from the LF state on the S<sub>1</sub> PES, where the axial bonds are partially elongated. It is more energetically feasible to photo-dissociate from the LF state than MLCT state. For GLM, the LF state is 1.87 kcal/mol lower in energy than the MLCT state. Based on previous studies, two pathways have been identified namely, Path A and Path B, to connect the MLCT and LF state (Figure 4.7).
Figure 4.7. a) Photodissociation scheme for AdoCbl-dependent GLM based on the $S_1$ PES. Two minima regions were identified namely, MLCT and LF, separated by a seam. Characteristic points were located on the $S_1$ PES and indicated as I ($S_{1\text{min}}$) in the MLCT region, IIB ($S_{1\text{min}}$) at the MECP, and IIIB ($S_{1\text{min}}$) in the LF region. Path A and Path B were identified to connect the MLCT and LF minima. b) The energetics associated with Path A and Path B were plotted as a function of $Q$ vs energy.

From I ($S_{1\text{min}}$), Path A begins by the elongation of the Co-C bond, followed by the elongation of the Co-N$_{Im}$ bond, whereas Path B involves the elongation of the Co-N$_{Im}$ bond followed by the elongation of the Co-C bond. The energetic profile of the pathways was plotted as a function of $Q$ versus energy where $Q$ is defined as,

$$ Q = \sqrt{R_{Co-C}^2 + R_{Co-N}^2} $$

The $Q$ coordinate represents the distance of a point along either Path A or Path B from the origin of the surface. This has been conducted to represent a one-dimensional minimum energy profile from two bond lengths, Co-C and Co-N$_{Im}$. This allows for a one-dimensional quantitative comparison of the energetic differences of Path A and Path B. A similar analysis was conducted in our previous work that described the photodissociation of
AdoCbl bound to EAL. Based on the energy plot as a function of energy versus Q, it appears that two pathways are energetically feasible in GLM. In Path A, the barrier to cross the seam is 1.88 kcal/mol whereas in Path B it is 1.91 kcal/mol (Figure 4.7). Both pathways lead to the LF state, with a comparable energy barrier, from where photodissociation will take place. In previous studies of various Cbls in solution, there is typically an active pathway and an inactive pathway that can be identified based on energetic grounds. It is not this straightforward in the enzymatic environment as it would appear that both Path A and B are energetically feasible.

It is also important to describe the deactivation mechanism while describing photodissociation. After the radical pair generation from the LF state, the alkyl radical can either escape the ‘cage’ provided by the enzyme or recombine with the Co(II)corrin system, resulting in the return to the ground state. There are two possible channels for deactivation of the isolated cofactor AdoCbl that have been proposed. The first one involves corrin ring distortion after the axial base has been detached. N-Co-N valence angle puckers inward and the Co-C bond is shortened at the point where the $S_0$ and $S_1$ PESs cross. Another involves de-excitations of Co(II) from the LF state to the ground state by changing the electronic configuration from $(d_{yz})^1 (d_z^2)^2$ to $(d_{yz})^2 (d_z^2)^1$. Based on our calculations, it appears that for GLM or EAL, the first channel of deactivation is inactive because the enzymatic environment imposes constraints on the corrin macrocycle that prevent distortions. The second channel of de-activation appears to be more feasible in the enzymatic environment.

4.3.5 Comparison of the photolytic mechanism in GLM with EAL. GLM and EAL represent two major classes of AdoCbl-dependent enzymes. Understanding the mechanism
in both cases allows one to gain a comprehensive insight into the AdoCbl-dependent enzymes within these groups. In both enzymes, activation of the Co-C bond by light generates the same radical pair from the S1 state. However, it was not automatically manifested if the mechanisms of photodissociation would be comparable especially considering the differing binding modes of AdoCbl in GLM and EAL. Recall that the AdoCbl cofactor in GLM is in the base-off/His-on form whereas in EAL the base-on form is present. To compare the mechanisms, the same level of theory has been applied to construct the PESs. The topology of the S0 surface in both enzymes are indistinguishable (Figure 4.8).

Figure 4.8. S1 PES as a function of axial bond lengths Co-C and Co-Nim with the photodissociation pathways depicted by red arrows for a) AdoCbl-dependent GLM and b) AdoCbl-dependent EAL.

However, differences can be noted for the S1 surfaces. In GLM, the LF minimum is located where the Co-C bond is ~2.50 Å and the Co-Nim is ~2.70 Å. In EAL, the axial bonds for
the LF minimum are ~2.50 Å and 2.90 Å for Co-C and Co-N_{Im}, respectively. A potential explanation for this is that the higher basicity of the lower axial base shifts the LF minimum IIIB (S_{1min}) to a shorter Co-N_{Im} distance. There is also a difference in the energetics between the LF and MLCT states compared to GLM with EAL. In GLM the LF minimum IIIB(S_{1min}) is 1.87 kcal/mol lower in energy than the MLCT minimum 1(S_{1min}) (Figure 4.9). In EAL, the LF minimum is 1.20 kcal/mol lower than the MLCT minimum (Figure 4.9). In addition, the barrier to cross the seam in GLM is 1.91 kcal/mol which is 2.48 kcal/mol in EAL (Figure 4.9). Thus, it appears the energetic barrier and the relative energy differences in GLM is smaller compared to EAL. These differences are rather small; however, the origin of these differences can be associated with the presence of different lower axial ligands in these two enzymes. It was also shown that among the two pathways, Path B is more feasible than Path A in EAL, whereas in GLM both pathways seem to be equally possible (Figure 4.8).

**Figure 4.9.** Relative energetics of the MLCT and LF state in AdoCbl cofactor for a) AdoCbl-dependent GLM and b) AdoCbl-dependent EAL along the coordinate Q.
4.3.6 Comparison with experiments

**4.3.6.1 TAS and EPR studies:** The photolysis of AdoCbl in solution and bound to enzymes was investigated using TAS by Sension and coworkers.\textsuperscript{152,177} It was shown that for AdoCbl in solution and bound to the enzyme, excitation with light at 400 nm or 520 nm results in rapid IC from the initially excited high energy state to the $S_1$ state. It is here, in the $S_1$ state, where homolytic cleavage of the Co-C bond generates the RPs and this is followed by rapid IC to the ground state. The sequence of events that occurs in the $S_1$ state can be described qualitatively based on our TD-DFT calculations. Our calculations for the AdoCbl-dependent GLM model, indicate that the $S_3$ and $S_5$ states have higher oscillator strengths at 504 nm and 483 nm. The significant contribution for the electronic transitions in $S_3$ state are from 43% $d_{yz} + \pi \to \pi^*$ and 33% $d_{xz}/d_{z^2} + \pi \to d_{z^2}$ and for the $S_5$ state 35% $d_{xz}/d_{z^2} + \pi \to \sigma^*(d_{z^2})$ and 27% $d_{yz} + \pi \to \pi^*$ (Table 4.3). These states have a mixing of MLCT and LF character. On the other hand, the $S_2$ state has a comparable oscillator strength (0.0213) at 537 nm. The dominant contribution is from 44% $d_{xz} + \pi \to \pi^*$, 34% $d_{xz}/d_{z^2} + \pi \to \pi^*$ in this state. This is consistent with the MLCT assignment based on TAS study and this corresponds well with the excitation energy of the UV-visible absorption spectrum for the $\alpha/\beta$-band maximum. Based on these comparisons with experiments, we have hypothesized that the initial excitation is associated with $S_2$ state from which IC takes place to the $S_1$ MLCT state (Figure 4.10). This is followed by the subsequent relaxation to the LF state through Path A or Path B (Figure 4.10). The energy difference between $S_1$ MLCT and $S_1$ Franck-Condon point ($S_{1\text{vert}} - S_{1\text{MLCT}}$) is 0.92 kcal/mol. The barrier from the MLCT state to the LF state is 1.38 kcal/mol at IIB (MECP). The overall barrier which is the energy difference between the $S_0$ minimum and the IIB ($S_{1\text{min}}$) is 47.0 kcal/mol (Figure 99).
4.9). This is within the activation energy of light, which is 49 kcal/mol based on our calculations.

Figure 4.10. The photodissociation mechanism in AdoCbl-dependent GLM. Both pathways are simultaneously active in connecting the MLCT and LF states. The deactivation mechanism is also shown. Deactivation by the distortion of the coordination sphere is inactive inside GLM.

According to our calculated $S_1$ PES, it appears that the lower axial bond Co-N$_{Im}$ is significantly lengthened in the LF minimum for AdoCbl in solution. In the LF state of the isolated cofactor, the minimum is localized at longer Co-N$_{Im}$ bond distances (Figure 4.11). From our calculations, it also appears that LF minimum on the $S_1$ PES is localized at shorter Co-N$_{Im}$ bond distance for AdoCbl in GLM when it is compared with the isolated cofactor in solution (Figure 4.11). When comparing the environment from solution to the enzyme, localization of the LF minimum is shifted to shorter Co-N$_{Im}$ bond distances. The excited state intermediates in AdoCbl bound to GLM are longer-lived compared to those in the isolated cofactor. This can be explained using the calculated $S_1$ PES. In GLM the LF region is broader and stabilized. In other words, the LF state is energetically lower than the MLCT state. The electronic contributions in the LF region are from Co d orbital + corrin π → Co d$_{z^2}$. Thus, the enzymatic environment influences the $S_1$ state by stabilizing the Co d$_{z^2}$ orbital. This stabilization of the molecular orbitals is consistent with TAS studies.
Figure 4.11. PES of $S_0$ and $S_1$ electronic states plotted as a function of axial bond lengths along with the contour representation of $S_0$ and $S_1$ states as a function of axial bond lengths to make the comparison of $S_1$ PES and the photodissociation mechanism in a different environment. a) AdoCbl-dependent GLM b) AdoCbl-dependent EAL c) Isolated cofactor AdoCbl in base-on form where the lower axial ligand is DBI. d) Isolated cofactor AdoCbl in the base-off form where the lower axial ligand is detached and replaced by water. The color codes are shown along with the figures. Panel b, c, d was plotted using the same color scales whereas different color scale used for panel a as shown in the figure.
These differences between the isolated cofactor and the cofactor bound to the enzyme can be further explained based on electron EPR studies. Based on EPR studies conducted by Marsh et al., it was shown that there is a difference in the EPR spectra of free Co(II) and Co(II) bound to GLM. The EPR experiment is focusing on the ground state, represented in our study by the $S_0$ PES, and thus we need to focus on the g and A-tensors associated with the Co $d_{z^2}$ center. Since both tensors are axially symmetric, the differences in the super hyperfine coupling pattern are mainly modulated by the local environment of Co, in other words, axial bonding. Indeed, when we compare the isolated cofactor containing Co(II) to our QM/MM calculations of GLM, the outcome of our studies are consistent with Marsh’s work. Both calculations and experiment point out that the histidine ligand of the protein coordinates more strongly to the Co than does the DBI base in the isolated cofactor.

4.3.6.2 Resonance Raman spectroscopy: Resonance Raman (RR) spectroscopy was used to characterize the electronic transitions and to assign the vibrational modes of the AdoCbl cofactor. RR studies of B$_{12}$ dependent systems were interpreted by Spiro and coworkers indicated that the $S_0\rightarrow S_1$ electronic transition is associated with the corrin $\pi\rightarrow\pi^*$ character. However, the original assignment for AdoCbl-dependent systems have been revised by Kozlowski, Spiro, and co-workers. It was noticed that for the Ado derivative there is a mixing of the Co-C stretching with the ribose deformation coordinate. According to our study, the $S_1$ state has primarily MLCT character. Our results are consistent with RR data reported by Spiro taking into consideration that: low-lying excited states ($S_1 - S_5$) of AdoCbl, computed via TD-DFT, have mixed character of corrin $\pi\rightarrow\pi^*$ + Co $d$ orbitals transition (see Table 4.3). In RR spectroscopy the excitation wavelength is carefully chosen.
to overlap with (or be close to) an electronic transition that has a large transition dipole moment. RR spectra of AdoCbl in solution was acquired using 530.9 nm excitation, while for AdoCbl-dependent GLM with 514.5 nm.\textsuperscript{55} Consequently, the computed S\textsubscript{2} and S\textsubscript{3} excitations are much closer energetically to laser excitation than the computed S\textsubscript{1}. Besides, both states have noticeably larger, at least by one order of magnitude, transition dipole moments. Both states contain significant contributions from the corrin \(\pi\rightarrow\pi^*\) transition. Thus, it appears that the S\textsubscript{1} state cannot be explained simply as containing only corrin \(\pi\rightarrow\pi^*\) transition because the low-lying electronic transitions have more complex character and have a mixing with Co. Considering all of the above, the S\textsubscript{1} state is primarily MLCT in character which is also consistent with Circular Dichroism (CD) and Magnetic Circular Dichroism (MCD) and TAS studies.\textsuperscript{124,155,235,236}

4.3.6.3 Oxidation states of Co ion in MLCT state: While it appears that the oxidation state of the Co in the MLCT state can be described as Co(IV), further detailed analysis of the electronic properties indicate that the formal oxidation state cannot be assigned. To explore whether Co inside corrin can sustain such high oxidation, we performed additional DFT calculations on one-electron oxidized (1e-Ox) form of AdoCbl cofactor. Electronically, Co(IV) inside corrin macrocycle should be comparable to the one-electron oxidized analog, in a similar way that has been reported for metalloporphyrins.\textsuperscript{237} The electronic structure of the 1e-Ox AdoCbl cofactor was analyzed using BP86/TZVPP and B3LYP/TZVPP level of theory. An analysis of orbitals was performed to characterize the Kohn-Sham orbitals (KSO) (Figure 4.12). The KSO has Co-C bonding character with a mixed contribution from the corrin \(\pi\) and Co \(d_{z^2}\) orbital. Spin density distribution of the 1e-Ox models was also analyzed using BP86 and B3LYP functionals and shown in Figure
4.12 and A7. The spin density distribution was found to be Co 30%, corrin 51%, and Ado 12% using BP86 functional (Figure 4.13 and A8). For B3LYP, the spin distribution is Co 6%, corrin 69%, Ado 18% (Figure 4.13). Thus, based on the spin polarization, the existence of Co(IV) is not justifiable. Additional complications may potentially result from the pseudo-Jahn-Teller effect.\textsuperscript{238} The three highest occupied molecular orbitals for Co(III) are \(d_{xz}, d_{yz}, d_z^2\) and the two unoccupied orbitals are \(d_{xy}\) and \(d_{x^2-y^2}\). As the \(d_{xz}\) and \(d_{yz}\) orbitals are nearly degenerate, the removal of an electron makes the electronic configuration unstable and the system experiences the pseudo-Jahn-Teller effect.\textsuperscript{238} Consequently, the best description would be consistent with a mixture of Co(IV)/Co(III) states.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{energy_diagram.png}
\caption{Energy diagram of frontier KSOs for Im-[CoIV(corrin)]-Ado\textsuperscript{2+} (labeled 1e-Ox) model complex at BP86/TZVPP level of theory.}
\end{figure}
The question naturally arises if such excitation should be formally described as a state having MLCT character. To answer this question, we computed the electron density difference between the $S_0$ and $S_1$ states for AdoCbl in GLM (Figure 4.14). Additional TD-DFT calculations only partially support that the $S_1$ intermediate state is of MLCT character. Closer inspection of electron density changes indicates that it is not solely shifted from the Co but from the axial $N_{\text{Im}}$-Co-$C_{\text{Ado}}$ bonding to corrin ligand. These types of electronic excitations have been described as $\sigma$ bond-to-ligand CT transitions (SBLCT)\textsuperscript{172}. Thus, probably the most appropriate way to describe this state would be having mixed MLCT/SBLCT character. This has been noticed in our previous studies for MeCbl but for consistency with experimental data, we simply use MLCT.

Figure 4.13. (a) Isosurface of spin density obtained from BP86/TZVPP calculations for the Im-[CoIV(corrin)]-Ado$^{2+}$ model complex (labeled 1e-Ox). (b) Spin density distribution on fragments of the complex based on Mulliken population analysis.
Figure 4.14. Electron density differences between $S_0$ and $S_1$ states for the AdoCbl-dependent GLM at $S_{(1\text{min})}$ MLCT point. Results obtained from TD-DFT calculations. Isosurface plot value of 0.002 was used.

4.3.6.4 Quantum yield and relaxation rates: QY of photolysis is another experimental entity that needs to be considered. Based on our $S_1$ PES it would seem that in the enzymatic environment, it is easier to generate the radical pair in the LF intermediate state compared to the solution environment. However, once the radical pair are generated in the LF state, it is more difficult to escape the ‘cage’ provided by the enzyme and the radicals is less likely to diffuse. Based on experiments by Sension and coworkers and Warncke and coworkers, the QY for the long-lived radical pairs in the isolated cofactor AdoCbl is relatively higher when compared with AdoCbl bound to an enzyme. In GLM, the QY reported for photodissociation is $0.05 \pm 0.03$, whereas for the isolated cofactor in solution it is $0.28 \pm 0.02$.\textsuperscript{119,152,155,177} Thus, the generation of the long-lived radical pairs occur more easily in the isolated cofactor compared to AdoCbl bound to GLM.
In GLM 90-95% of the radical pair tend towards geminate recombination to the ground state after light-induced cleavage. A similar case was also observed for AdoCbl-dependent EAL. Based on the experimental studies, the enzymatic environment increases the rate of geminate recombination to the ground state, hence reducing the loss of radical pairs from the LF state. Sension and co-workers showed that the protein environment decreases the rate constant for geminate recombination by ~30% inside AdoCbl-dependent GLM compared to the isolated cofactor. Computationally, geminate recombination rate can be estimated by considering the radiative decay from the S$_1$ LF state to the S$_0$ state by using the deactivation rate constant (k$_{fl}$) which is also known as Einstein coefficient $A$. Specifically, k$_{fl}$ can be expressed as,

$$ k_{fl} = A = \frac{2\pi v^2 e^2}{c^3 m_e \varepsilon_0} * f $$

where, $v$ is the emission wavenumber, $c$ is the speed of light, $m_e$ is the mass of an electron, $f$ is the oscillator strength, $\varepsilon_0$ is the vacuum dielectric constant. Based on our S$_1$ PES, we
have calculated the $k_{fr}$ for AdoCbl-bound to GLM and compared this with the isolated cofactor in solution. According to our calculations, the rate constant for the geminate recombination is reduced by 45% inside GLM when compared to the isolated cofactor. Thus, our results are qualitatively consistent with the experiment considering that the protein decreases the rate constant by a modest amount which in turn decreases the QY of the long-lived radical pairs.

4.3.7 Implications of GLM photochemistry. This study has revealed some fundamental insights about the photochemical cleavage of the Co-C bond of AdoCbl bound to GLM. Photochemical and photophysical properties of GLM were studied using ultrafast TAS and it was shown that inside an enzyme the cleavage of Co-C bond by light differs than the isolated cofactor of AdoCbl.\textsuperscript{152,177} It is also clear from our study that the topology of the $S_1$ PES and the pathways of photodissociation differ between the two mediums. The $S_1$ PES was characterized as having two electronic states namely MLCT and LF. Inside enzymes like GLM and EAL, the MLCT is higher in energy than the LF (Figure 4.9). In the base-on form of the isolated cofactor, the MLCT minimum is lower in energy than LF minimum. Moreover, the LF minimum is also located at shorter Co-N\textsubscript{Im} distances inside GLM than the isolated cofactor. It is apparent that the enzymatic environment influences the LF state. This finding can be reconciled with the previous studies that suggest that the role of enzymatic environment is to influence the excited state rather than the ground state.

4.4. Conclusions

The photodissociation mechanism of AdoCbl inside GLM can be clearly understood by scrutinizing the $S_1$ PES with respect to the axial bonds. The topology of the $S_1$ PES in GLM is consistent with the topology of the $S_1$ PES in EAL. The $S_1$ PES contains
two minima regions namely MLCT and LF. In GLM the LF minimum is located at a shorter Co-N_{Im} distance than in the isolated cofactor which is consistent with the TAS spectrum observed for the GLM. Inside enzymes like GLM or EAL, the energy associated with the LF is lower than the energy in MLCT. Thus, the enzymatic environment appears to lower the energy of the LF state. Inside enzymes radical pairs are more easily generated in LF state but the QY for the long-lived radical pairs is lower in enzyme due to the preventive loss of radical pairs by the cage provided by the enzymatic environment. There are also some differences noted between the two classes of enzymes considered here. In the case of GLM, the LF minimum is energetically more stabilized compared to the EAL. The LF minimum is also located at a shorter Co-N_{Im} distance than in EAL. Thus, it appears that the binding mode of the lower axial ligand affects the S_{1} state in AdoCbl-dependent enzymes.

There are two pathways, Path A and Path B that were identified to connect the MLCT to LF state that describe the photodissociation mechanism. In GLM there is no obvious distinction between the two pathways energetically, whereas in the case of EAL, one pathway is energetically slightly more favorable than the other one. In the case of isolated cofactor, there is a clear distinction between the pathways and only Path A is active for the base-on AdoCbl and Path B is active for the base-of AdoCbl. Considering all of the above, this study is a crucial step to clearly understand the photochemical activation of Co-C bond in AdoCbl-dependent enzymes.
CHAPTER V

A COMPUTATIONAL INVESTIGATION OF NATIVE CATALYSIS MECHANISM IN ADOCBL-DEPENDENT ETHANOLAMINE AMMONIA-LYASE.

5.1. Background

Enzymes are nature’s most ubiquitous biomolecules that not only enhance the rate of chemical reactions but also prevent unwanted side reactions. Different types of enzymes employ different techniques in order to accelerate the rate of reactions to a significant extent compare to the uncatalyzed reactions.\textsuperscript{240} The enormous catalytic efficiency of enzymes has been described as either electrostatic stabilization, entropic contribution or strain hypothesis to catalyze thermodynamically unfavorable reactions.\textsuperscript{5,6,108,109,241} One of the most well-known classes of enzyme groups is vitamin B\textsubscript{12}-dependent enzymes that participate in numerous reactions such as hetero atom elimination, 1,2-migration reactions, and deamination reactions.\textsuperscript{3,5-7,60,185,242} These reactions are essential for every form of life including mammals and bacteria for many purposes such as DNA regulation as well as energy generation.\textsuperscript{10,57-59} Although vitamin B\textsubscript{12} is biologically inactive and not used as a cofactor however it's active biological analog known as coenzyme B\textsubscript{12} plays the role of a cofactor in these classes of enzymes.\textsuperscript{3,113} As discussed earlier, there are two biologically active B\textsubscript{12} derivatives, CH\textsubscript{3}Cbl and AdoCbl.\textsuperscript{1,13} Ethanolamine ammonia-lyase (EAL) is the AdoCbl-dependent enzyme that catalyzes the deamination of amino alcohols.\textsuperscript{53} One of the most commonly used substrates in this type of reaction is ethanolamine (EA) which is
being utilized as a sole source of carbon and energy for different types of organisms.\textsuperscript{227,243}

Enzymes such as EAL rely on AdoCbl to catalyze a wide variety of reactions utilizing the cofactor AdoCbl as a radical reservoir.\textsuperscript{3,53,70,72,101,185} It is a naturally occurring complex organometallic molecule that possess a unique Co-C $\sigma$ bond.\textsuperscript{1,36} The detail structural formula of AdoCbl has been shown in figure 1.1 and discussed in detail in section 1.2. The x-ray structure of EAL was also resolved and showed that the EAL structure is a trimer of the $(\alpha/\beta)_2$ dimer.\textsuperscript{53} AdoCbl cofactor bind with the enzyme in the interface of the $\alpha$ and $\beta$ subunits. The substrate EA binds with the enzyme in the C-terminal side of the barrel which forms a hollow space. Upon substrate binding, with the enzyme, the initial step of the enzymatic catalysis in EAL is to homolytically cleave the Co-C bond to generate Co(II) and Ado radical. This Ado radical species then abstracts an H-atom from the substrate EA to generate adenosine and substrate radical, which is followed by the rearrangement to generate product radical. This mechanism is shown in detail in Figure 3.1. The most important aspect of this reaction is that the rate of the Co-C bond is a trillion-fold accelerated inside the enzyme compared to the dissociation of Co-C bond in solution.\textsuperscript{3,53,61,87,185} Also, the cleavage of the Co-C bond and the H-atom abstraction steps are coupled as evidenced by the kinetic isotope effect.\textsuperscript{105,244}

In order to understand the detailed mechanism of coenzyme B$_{12}$-dependent enzymes, it is paramount to look at the factors responsible for this enormous catalytic efficiency. Several studies have been conducted to explore the detailed mechanism. The proposed mechanism showed enthalpic factors such as steric and electronic contributions as well as entropic factors are responsible for the rate enhancement of Co-C bond cleavage.\textsuperscript{108,109,185,241} However, a consensus was not reached based on those mechanisms.
For example, it has been suggested that the substrate-binding deforms the corrin ring structure to weakens the upper axial Co-C bond. Some other critical hypotheses based on the experiments are bending of the upper axial ligand, steric strain induced by the binding of the substrate. However, based on the x-ray structure of EAL, it appears that the substrate-binding does not switch the structure of the corrin ring. Besides, recent UV-visible spectroscopic studies also showed that the Co-C bond is not significantly distorted upon substrate binding. This was also further explored in the photolysis study of the Co-C bond for the substrate-bound and substrate-free complex, where both similarly undergoes the photolysis. The stabilization of the transition state was proposed for methyl-malonyl Co-A mutase (MCM). This has been validated using different spectroscopic and computational techniques. However, this needs to be investigated for many other B_{12}-dependent enzymes.

There is also another hypothesis that is less explored in understanding the mechanism of B_{12}-dependent enzymes, which is the effect of the reduced cofactor in understanding the rate enhancement. In the 1980s, Lexa and Savant demonstrated that the addition of an electron to the Cbls reduced the BDE of Co-C bond by a factor of the half. This reduction in BDE may lead to the significant enhancement of the reaction rate inside the enzyme. In the case of the enzyme, the protein residue acts as the source of the electron which initially goes to the corrin ring otherwise can be called as π*corrin state. Upon stretching of the Co-C bond, this electron moves to the axial σ*Co-C state and weakens the bond. This was shown to be useful in many other biological systems such as cytochrome C, photosystem I and II. This needs to be investigated for B_{12}-dependent enzymes.
In addition to these, the catalytic origin of AdoCbl-dependent enzymes is also associated with the fate of generated Ado radicals. Upon cleavage of the Co-C bond, the Ado radical must migrate to the substrate to abstract the H-atom. In that case, the AdoCbl enzymes elegantly control the Ado radical to prevent unwanted side reactions. This is known as negative catalysis based on the Retey concept. The migration of Ado radical occurs from 6-10 Å depending on the type of enzyme. For mutase, it was shown that the Ado radical migrates 4-6 Å whereas in eliminases it travels 7-9 Å to abstract the H-atom.\textsuperscript{3,5,70,92} Based on different studies it appears that in order to facilitate the H abstraction the Ado radical undergoes some conformational changes such as adenine ribose rotation in MCM\textsuperscript{254}, ribose puckering in GLM \textsuperscript{81}, pseudo-rotation of glycosyl moiety in DDH \textsuperscript{95}. These are important understanding and requires further investigation for many other AdoCbl-dependent enzymes.

Taking all these into account, it appears that the understanding of AdoCbl-dependent enzymatic reactions is difficult to probe using the experimental techniques alone. Mainly the reaction proceeds through several intermediates which are transient and highly reactive species. In order to understand the mechanism completely, several groups employed different computational methods for different B\textsubscript{12}-dependent enzymes. Specifically, Jensen and Ryde performed QM/MM calculations to understand the cleavage of the Co-C bond in AdoCbl-dependent GLM.\textsuperscript{111} It was shown that, the catalytic effect in GLM is induced by several factors such as geometrical distortions of the cofactor by enzyme and the stabilization of the Co(II) state due to the electrostatic interactions of protein residue with the cofactor. Morukuma and co-workers employed ONIOM method to understand the mechanism of the Co-C bond dissociation in MCM and ornithine
aminomutase (OAM). In both cases, the homolysis of AdoCbl is shown to be stepwise which was argued as a concerted mechanism in some other previous studies. According to some studies of model calculations, it was suggested that the concerted mechanism is favorable over the stepwise. This is not the case inside the enzyme as evidenced by the studies of Morukuma and co-workers. Ursula and co-workers conducted a detail QM/MM study of MCM enzyme. Based on their study it appears that the enzyme acts as a mechanochemical switch to control the Ado radical intermediate. This specific capability to control the radical intermediates is an inherent factor of B<sub>12</sub>-dependent enzymes that allows for this unusual reactivity.

Despite these valuable contributions of different scientists and research groups, there is still much to be learned about the mechanism of AdoCbl-dependent enzymatic catalysis. Specifically, the role of the specific residues, the effect of the reduced cofactor inside enzymes, are important issues yet to be considered. Herein this chapter, we employed the QM/MM method to understand the mechanism of the AdoCbl-dependent EAL. In this regard, we investigated the total catalytic cycle of EAL and different intermediated were located and studied. BDE was computed for the AdoCbl cofactor in the native and reduced condition and compared with the cofactor in solution. Different factors that contributed to the lowering of BDE were also identified and discussed.

5.2. Computational details

5.2.1. System preparation. The initial structure of the EAL enzyme was obtained from the protein data bank (PDB Id: 3ABS). The structure of EAL was complexed with Adeninylpentylcobalanin (AdePeCbl) at crystallographic 2.10Å resolution. The EAL enzyme is a hexamer of α/β subunit where one minimal unit is composed of α and β
subunits, AdePeCbl cofactor and the EA substrate. There are two minimal asymmetric reactive units in one EAL structure which are independent and far from each other (Figure 3.3). The structure was initially protonated with PDB2PQR webserver at pH 7.0. The protonation states of the titratable residues were determined carefully and also with manual inspection. The protonated structure was then further screened and processed with chimera program.\textsuperscript{255} Using chimera one asymmetric unit was used for further calculations. To prepare a realistic structural model, adeninylpentyl (AdePe) group was replaced with the adenosine (Ado) part.

**5.2.2. Molecular dynamics simulation.** The active site of the EAL structure contains AdoCbl cofactor, and the EA substrate. The AdoCbl cofactor was parameterized using metal center parameter builder (MCPB) implemented in Amber software suite.\textsuperscript{204} It is very popular and can be used for the parameterization of a variety of metalloproteins. Based on the MCPB.py a smaller model system was prepared which is then subject to further quantum mechanical (QM) calculations. Gaussian software was used for the QM calculations to obtain the partial charges and the missing atomic parameters.\textsuperscript{229} The partial atomic charges and the missing atomic parameters were calculated using BP86/6-31g(d) level of theory. In order to neutralize the system, 17 sodium ions were added to the protein surface. Finally, the total system was solvated in a rectangular box of TIP3P water using 18 Å cutoff from the protein boundary. In total 20644 water molecules were added in the system and altogether, the system contained 33000 atoms.

The resulting system was then initially minimized with the steepest descent and conjugate gradient methods for several steps. Initially, the position of the water and ions were minimized while keeping the enzyme, cofactor and the substrate restrained with a
harmonic force constraint of about 500 kcal. mol\(^{-1}\) Å\(^{-2}\). This was done to fix the position of the nonstandard residues which is then followed by the minimization of the total system without any restraint. After the initial minimization, the system is gently heated from 0 K to 300 K for 20 ps at the canonical ensemble (NVT) MD simulation with a weak restraint of 5 kcal. mol\(^{-1}\) Å\(^{-2}\). Afterward, the system was equilibrated at isothermal-isobaric ensemble (NPT) for 20 ps with a weak restraint of 5 kcal. mol. Å\(^{-2}\). This is then followed by the 20 ps NPT MD simulation without any restraint. This short MD equilibration was chosen to avoid the bubbles due to the added tleap water molecules in the system. After this initial heating and equilibrium, a stable density was achieved for the system. This is then followed by a 200 ns production NVT MD simulation, which is a reliable sampling method implemented in amber.\(^{205}\) For all the MD simulation shake algorithm was used to constraint the bond lengths involving hydrogens. Periodic boundary conditions were applied, and 2fs time step was chosen for the MD simulation. The calculations were performed using Amber force field (FF14SB)\(^{206}\) implemented in Amber 14 software package.\(^{205}\)

5.2.3. QM/MM calculations. Prior to the QM/MM calculations, it is important to begin with the moderate structure from the MD simulation. One of the challenges is to find a suitable MD snapshot. We analyzed the root mean square deviation (rmsd) from the initial structure in MD simulation. The rmsd value based on our MD simulation reached a plateau after 4 ns and after that, it indicates inconsiderable structural changes in the structure. We used a snapshot at 50 ns to start our QM/MM calculations where the rmsd differences are very negligible. This snapshot is then further evaluated with MolUp program implemented in VMD software package.\(^{208,209}\) Using Molup initially the solvent water and the ions were
removed from the structure. It was shown in other studies that keeping the ions in the QM/MM calculations has negligible effect in the calculations. Protein residues 20 Å away from the Co center were frozen whereas the rest of the system was allowed to move. In total, the system contained 3158 unfreeze atoms and 7834 frozen atoms.

The system is then partitioned into three layers. The high layer part contains the substrate EA, the corrin ring which is partially truncated and the Ado group. The middle layer part contains the rest of the corrin ring, and the protein part were included in the low layer. The high layer part of the system was treated with DFT where the BP86 functional and TZVP/P basis set used for the calculations. The middle layer part was treated with semi-empirical PM6 and the molecular mechanics (MM) were used for the low layer part using Amber force field (FF99SB). DFT/PM6/MM calculations were used to construct the BDE curve and to identify different intermediates in the catalytic cycle. Gaussian 16 was used for the QM/MM calculations. In addition, the isolated AdoCbl cofactor was also used for the model system calculations. DFT with BP86 functional and TZVP/P basis set was used for the calculations of isolated AdoCbl cofactor in polarizable continuum model (PCM). Gaussian 09 was used for these calculations.

5.3. Results and discussions.

5.3.1. Ground state geometry. Initially, the AdoCbl cofactor in solution and the EAL model system were optimized using the level of theory discussed in the computational details section. Table 5.1 shows the detailed comparison of optimized structures using these two levels of theory along with the crystal structures.
Table 5.1. Structural parameters of the AdoCbl cofactor in different environments.

Based on the structural data, it appears that upon binding with the enzyme there are certain geometrical parameters that are susceptible to the environmental effect. This is also evident from the previous studies conducted by the Kozlowski group and others.\(^{178-180}\) Though the protein binding does not influence the overall geometry of the cofactor, however, there are few minor changes that occur in the case of ground-state structure. For instance, the Co-N\textsubscript{Pyrrolic} bond distance is not perturbed in different environments; however, the changes occur in the case of axial ligand distances. For isolated cofactor, the Co-C and Co-N\textsubscript{Im} bond length were found to be 2.01 Å and 2.19 Å respectively whereas in EAL optimized structure it is 2.07 Å and 2.40 Å. These bond lengths also vary in the case of the EAL
crystal structure (Table 5.1, Figure 5.1). Thus, the axial ligands are sensitive to the environment which is very crucial for photolysis and also in the native catalysis.

There are other issues that also need to be considered such as the torsional angle and the fold angle of the corrin ring. Though it appears that the substrate-free and substrate-bound EAL has similar structural features based on our previous studies however differences were observed with the isolated AdoCbl cofactor in solution. Glycosidic angle ($\chi$) which is defined as the torsional angle between $O1$-$C1$-$N9$-$C8$ atoms in adenosyl moiety differs in these two mediums (Figure 3.4). In the case of an isolated cofactor, the $\chi$ angle was found to be $-82^\circ$ which is $-71^\circ$ inside the enzyme. This degree of differences allows the adenosyl moiety to be more flattened inside the enzymatic environment and allows the H bonding network with the vicinal protein residue. The fold angle of the corrin ring is another factor in determining the structural features in corrin. It is defined by the angle between Co and the planes of the conjugated ring system such as (N21, C4, C5, C6, N22, C9, C10) and (N23, C11, C10, N24, C16, C15, C14) (Figure 1.1, Table 5.1). The fold angle is lower inside EAL compare to the isolated cofactor.
Figure 5.1. The structure of EAL. The active site such as cofactor and the substrate are shown in ball and stick model and the protein residue is shown in ribbons. a) Crystal structure of EAL obtained from the protein data bank (PDB id- 3ABS). b) The optimized structure of EAL.

5.3.2. Cleavage of Co-C bond. The most crucial step of the Cbl-dependent enzymatic catalysis is the cleavage of the Co-C bond. The Co-C BDE is ~31.0 kcal/mol for AdoCbl in solution which reduces to 15-20 kcal/mol inside enzyme.\(^6\,^9\) This reduction in BDE inside enzymatic environment could be achieved because of the steric crowding, angular distortions of the corrin ring, interaction of the AdoCbl moiety with the surrounding environment. There are many other issues contributing to the reduction of BDE and also the rate enhancement. In order to understand this, the Co-C BDE was calculated using a potential energy curve (PEC) constructed by the elongation of the Co-C bond from 1.8 Å to 3.6 Å with a step size of 0.1 Å. The PEC of the EAL, isolated cofactor is shown in figure 5.2 which shows an identical pattern of the PEC in all the cases.
Figure 5.2. a) Bond dissociation energy curve of the Co-C obtained using different level of theory and compared b) Spin density distribution of the Co-C bond throughout the cleavage. c) MM energy contributions throughout the cleavage of the Co-C bond. At elongated bond length the contributions are lowest which is higher in the equilibrium state.

Spin density distribution indicates that the cleavage of the Co-C is homolytic. According to our calculations, the spin polarizations appear at ~2.6 Å and at bond length 3.4 Å, the cleavage is complete. In that case, the Co ion has +1 spin and the C atom has -1 spin indicates the diradical state of the system (Figure 5.2b). The BDE is thus the differences of the energetics of this diradical state with the equilibrium Co(III) state. In addition, the BDE is also calculated by optimizing the different intermediates in the reaction pathway such as \( ([\text{Co(II)}+\text{Ado radical}]-\text{Co(III)}\text{AdoCbl}) \). For isolated cofactor, the BDE was found to be 30 kcal/mol using BP86 level of theory. This is consistent with experimentally obtained BDE
and with other ab-initio studies. Thus, the use of this DFT functional is crucial in our QM/MM studies also. Based on the PEC in figure 5.2a, it appears that the BDE differs significantly inside the enzyme compared to the isolated cofactor.

In ONIOM calculations, the calculated BDE was considered in two ways. The QM energy where the energy of the model system was considered and the ONIOM energy where the energy of the total system including the QM, MM and PM6 energy was used. In the case of QM energy contributions, the calculated BDE from the PEC is 20 kcal/mol which is 15 kcal/mol based on the ONIOM contributions. To explain this in more details, there are certain things that needs to be contemplated carefully. The difference in QM, ONIOM and the solution energetics indicates that the steric and electronic contributions of the environment to the cofactor. Upon binding with the substrate EA in EAL, the ribose moiety of the AdoCbl cofactor forms a hydrogen bond with the substrate EA. The adenosyl moiety also undergoes the H-bonding network with the surrounding protein residue such as N193, S247 and G289, E287, Y404, Y285 (Figure A9). This was also depicted in a study conducted by Warshel and co-workers.\textsuperscript{108,109} Thus, the role of the Ado moiety is to labilize the Co-C bond through these H bonds inside the protein environment. This could be understood easily by replacing the Ado moiety with a less bulky group as Cbl upper axial ligand. The MM environment also plays a role in the labilization of energy. From figure 5.2c it appears that the energy of the MM layer is very high at the equilibrium Co(III) state. This energy follows exponential changes at the Co(II) state. The MM energy contributions are mostly coming from the Vander-Waals interaction which is very high at Co(III) state and decreases at Co(II) state. This is due to the electrostatic interactions of Co \(d_{z^2}\) with the surrounding protein environment.
5.3.3. Migration of Ado radical. The result of the Co-C homolytic cleavage is the generation of Co(II) and Ado radicals. Ado radical is a primary radical which is very reactive and short-lived. Inside enzymatic environment, the existence of this species has been shown by ultrafast spectroscopic technique for GLM and EAL while photolyzing of the Co-C bond.\(^{166,177}\) However, in the case of native catalysis the detection of Ado radical is not straightforward. Upon generation of the radical through the homolytic cleavage it migrates to the vicinity of the substrate to abstract the hydrogen atom. In EAL the migration is longer compare to the mutases such as GLM and MCM.\(^{53,81}\) Different factors play the role to facilitate the migration of the radical such as pseudo-rotation of the Ado as well as ribose puckering.

In order to understand this radical shuttling process, we modeled the reaction based on different coordinates. The \(\chi\) angle was systematically changed from a state where it represents the lowest energy state of the radical pairs [Co(II) + Ado]. The topology of this PEC is shown in figure 5.3. Our finding suggests that rotation of \(\chi\) around 70° from its initial position brings the Ado radical closer to the substrate atom. The conformation of Ado, in that case, is 3’-endo and the energy barrier to abstract the H atom from the substrate is 18 kcal/mol based on the ONIOM calculations. In the case of QM energy, the barrier was found to be 10 kcal/mol. Thus, this stabilization of the QM energy indicates the role of steric interactions of the enzymatic environment to the Ado radical. This is consistent with the Retey concept of negative catalysis. In that case, observation leads that the ONIOM energy is higher than the QM energy which indicates the strict control of the protein environment to guide the Ado radical for the H abstraction.
Figure 5.3. a) Potential energy curve of EAL along the dihedral angle (χ) to depict the hydrogen abstraction by Ado radical from the substrate. The χ angle is described as the torsion between O1-C1-N9-C8 atoms of adenosine molecule. b) The rotation of χ angle is shown.

Following the H abstraction from the substrate, generates substrate radical and the adenosine. The structure of these two catalytic steps [Co(II) + Ado radical] and [Co(II) + Subs radical] is shown in figure 5.4. Co(II) and the Subs radical is the first detectable intermediates in this catalytic cycle. EPR spectroscopic technique have been used to detect these intermediates in the reaction. The distance of the Co(II) and Subs radical were also measured by EPR which is around 8.2-9.0 Å.168,241 This is consistent with our studies which is 8.56 Å.
5.3.4. **Mechanism of EAL catalysis.** As discussed earlier the first detectable intermediates in the native catalytic cycle is Co(II) and Subs radical. Another important intermediate in this catalytic step is adenosine. The formation of this intermediate [Co(II)+ Subs rad] required to overcome 18 kcal/mol energy barrier which was discussed in the previous section. This intermediate was optimized using our QM/MM level of theory and shown in figure 5.4. [Co(II)+ Subs rad] is energetically favorable by 5.5 kcal/mol compare to the [Co(II) and Ado] radical step. In order to understand this clearly it is important to look at figure 3.1 about the scheme of native catalysis in EAL. After the formation of the substrate radical, it instantly rearranges to the product radical as shown by [Co(II) + Prod rad]. This is known as 1,2-rearrangement. Although the structure of product radicals is unknown in EAL but the existence of these types of radicals was demonstrated in different AdoCbl-
dependent enzymes.$^{102,165}$ Upon generation of product radical from the substrate radical, it abstracts the hydrogen atom from the adenosine molecules which is known as re-abstraction of hydrogen. The hydrogen re-abstraction step is depicted in figure 5.5.

**Figure 5.5.** a) The optimized structure of intermediate as shown in figure 5.7. The notation of this intermediate is $[\text{Ade} + \text{Prod}^\bullet]$. This intermediate forms after the rearrangement of the substrate radical in the catalytic cycle. b) The energy diagram of the hydrogen re-abstraction as a function $\chi$ angle with energy. The ONIOM energy and QM energy were shown separately to show the effect of the environment on the hydrogen re-abstraction.

This step is also shown by a PEC which is constructed by the elongation of the $\chi$ angle of adenosine. In that case after a $70^\circ$ rotation of the $\chi$ angle, the adenosine molecule reaches the product radical. The product radical re-abstract the hydrogen atom from the adenosine and form back the Ado radical. The energetic barrier to re-abstract the hydrogen atom is shown in figure 5.5b. The ONIOM energy is 16.5 kcal/mol and the QM energy is 9 kcal/mol. The role of the protein environment is to provide a strict control on the adenosine for the H re-abstraction. After the hydrogen abstraction step by the product radical, the next step is to generate the Ado radical and product species. The Ado radical
then recombine with Co(II) and form back the AdoCbl and close the catalytic cycle. These two intermediates were optimized using our ONIOM calculations and depicted in figure 5.6.

![Figure 5.6](image)

**Figure 5.6.** The optimized structure of intermediates species in EAL catalytic cycle. a) Optimized structure of diradical after the hydrogen re-abstraction. b) Optimized structure of the final product formed in the EAL catalytic cycle. The energetics of this species is shown in figure 5.7.

Total catalytic cycle in EAL was studied by splitting into a few individual steps and optimize the structures of each intermediates. The native catalytic cycle in EAL can be divided into these individual steps, step 1a: cleavage of the Co-C bond, step 1b: Formation of Ado and Co(II) radical, step 1c: hydrogen abstraction step, step 1d: formation of adenosine and substrate radical, step 1e: re-arrangement to product radical, step 1f: hydrogen re-abstraction step from the adenosine, step 1g: recombination to AdoCbl and the formation of the final product. The structures of each intermediate is shown in different figures 5.4, 5.5, 5.6 and the energetics is depicted in figures 5.7. The energetic contributions are depicted in figure 5.7 from both ONIOM and the QM energy part.
Figure 5.7. The overall energetics of the EAL catalytic cycle. The individual intermediates is depicted by 1a, 1b, etc. The blue color indicates the ONIOM energy whereas the red color is for the QM energy. QM energy is the energy of only the model system which includes the active site of the cofactor. This is discussed in the computational details section. The notation is following, \([\text{Subs}]\) is for substrate molecule, \([\text{Ado}^\bullet]\) is for adenosine radical also called as Ado radical, \([\text{H-abs}]\) is hydrogen abstraction, \([\text{Ade}]\) is for adenosine molecule, \([\text{Subs}^\bullet]\) is substrate radical, \([\text{Prod}]\) is for the final product, \([\text{Prod}^\bullet]\) is for the product radical which is formed by the rearrangement of the substrate radical, \([\text{H re-abs}]\) is for the hydrogen re-abstraction.

Based on figure 5.7 the energetic distributions of the EAL catalytic cycle can be expressed as follows. Upon the binding of the substrate, the initial step of the catalysis is the homolytic cleavage of the Co-C bond to generate Co(II) and Ado radical which requires 14.7 kcal/mol energy. This is shown as intermediate 1a. The conformation of the ribose moiety in Ado radical is 3'-endo. In order to abstract, the Ado radical undergoes conformational changes. In the case of mutases, the ribose pseudo-rotation occurs which facilitates the hydrogen abstraction.\(^{81}\) In GLM and MCM it was shown that the ribose changes the conformation from 2'-endo to 3'-endo to abstract the hydrogen abstraction.\(^{81,254}\)

Based on our studies, it appears that the migration of Ado radical to the substrate is
facilitated by the changes in $\chi$ angle whereas ribose conformation is preserved. To understand this, we scanned the $\chi$ angle of Ado radical from -70° to +20° which is shown in figure 5.3. In order to abstract the hydrogen from the substrate the Ado radical migrated to the proximity of the substrate. In that situation, the distance of Ado radical and the hydrogen atom from the substrate is 3.01 Å. The barrier to abstract the hydrogen atom was found to be 18.3 kcal/mol from the ONIOM energy and 9.8 kcal/mol based on the QM energy. Thus, the rotation of $\chi$ angle around 70° brings the Ado radical close to the substrate. The enzymatic environment seems to provide strict control of the radical intermediates in this case. The Ado conformation was found to be 3'-endo.

The next step after the hydrogen abstraction is the formation of the adenosine and substrate radical which is denoted as [Ade + Subs•] in figure 5.7 and identified as intermediate 1d. This intermediate could easily be detected using EPR spectroscopy. The optimized structure of this intermediate is shown in figure 5.4b. The formation of this intermediate is energetically favorable. Intermediate 1d is 6.8 kcal/mol lower in energy than intermediate 1b. In the next catalytic step, the adenosine and substrate radical react in the forward direction for the formation of the product radical. The structure of this product radical is unknown however it was shown that 1-amono-1-hydroxyethane-2-yl radical is energetically favorable. We optimized the structure using this radical intermediate. The energetics of this intermediate is shown in figure 5.7. The QM energy of intermediate 1e is highly favorable. Based on our calculations it appears that intermediate 1d is ~8.5 kcal/mol energetically favorable than intermediate 1d. Accordingly, this highly favorable energetics facilitate the formation of this intermediate. In the case of ONIOM energetics, intermediate 1e is energetically 3 kcal/mol higher in energy than 1d. This is consistent with
the previous discussion, as it indicates the influence of enzymes in the presence of primary radical pair.

The final detected intermediates in this reaction profile are the 1f and 1g from figure 5.7. 1f is the hydrogen re-abstraction step. After the formation of the product radical, which is shown as int 1e in figure 5.7, the hydrogen atom is re-abstracted from the adenosine molecule. In order to depict this, we have shown the reaction profile in figure 5.5b. The energy barrier to re-abstract the hydrogen atom is 8.5 kcal/mol. The $\chi$ angle rotation is around 70 ° which follows a similar pattern as in the case of hydrogen abstraction step. This is followed by the formation of the final product 1g as shown in figure 5.7. There are two possible pathways were discussed in the literature to identify the pathways of the product formation. In this current study, we optimized the final product aldehyde and ammonia and the AdoCbl structure. This is shown in figure 5.6b. It appears that the relative energy of this 5.6 kcal/mol higher in energy compared to 1a intermediate. However, the QM energy of this 1g intermediate is 5.2 kcal/mol lower in energy compared to 1a. Though the ONIOM energy is higher for this step but based on the analysis of QM energy, it appears that the formation is highly likely.

5.3.5. Implications of the study. As discussed earlier, the enzymology of B$_{12}$ is one of the most critical studies for experimentalists and computational chemists. Different groups used computational methods to understand the mechanism of B$_{12}$-dependent enzymes. Jensen and Ryde, employed QM/MM method to understand the mechanism of GLM$^{111}$, Morukuma and co-workers studies B$_{12}$-dependent enzymes MCM and OAM$^{85,254}$, Ursula and co-workers studied QM/MM method for the MCM$^{113}$, Warshel and co-workers employed empirical valence bond (EVB) method for EAL and MCM$^{109}$, Brunold and co-
workers studied the MCM enzyme using different computational approach\textsuperscript{52}, Kastner and co-workers studied the mechanism for GLM\textsuperscript{231} and many others. In addition to these Radom and co-workers applied different computational methods to understand the radical chemistry employed by the Cbl-dependent enzymes.\textsuperscript{185} Undoubtedly these are great contributions towards the understanding of the mechanism in B\textsubscript{12}-dependent enzymes; however, a significant consensus has not been reached yet.

In this study, we used the QM/MM method to understand the detailed mechanism by identifying each step. The associated intermediates were optimized along with the reaction path. According to our study, the initial step of the AdoCbl enzymatic catalysis inside the enzyme is energetically favorable by a factor of half compared to the bond fission in isolated cofactor. This is consistent with the idea that in order for the trillion-fold rate acceleration bond dissociation should be reduced by a factor of the half. Moreover, the cleavage of the Co-C bond in solution is also occurred in a similar manner as in the case of enzymatic catalysis; however, the energetics are different. Our study is consistent energetically with the studies conducted by Ursula, Warshel and co-workers. Based on the study by Warshel and co-workers, it appears that inside the enzyme the activation barrier to cleave the Co-C bond is around 13 kcal/mol which is reduced by half compared to the solution. This reduction of the Co-C BDE is due to the large-scale catalytic effect of enzymes. The origin of this effect is primarily attributed as electrostatic which is induced by the interaction of Ado with the surrounding environment. In the case, of EAL, the residue contributing to this is discussed in section 5.3.2 and shown in figure A9. The idea of this catalytic effect was referred to as entropic by Warncke and co-workers\textsuperscript{241} which has been argued and challenged by Warshel.\textsuperscript{109} Warshel and co-workers showed that the
The entropic effect has eventually played a role for the catalysis; however, the origin of this is associated with the electrostatic effect. This is consistent with our ideas that the electrostatic interactions with this surrounding protein residue play a role in facilitating the cleavage of the Co-C bond.

The stepwise and concerted mechanism is another argument in B\textsubscript{12}-enzymatic catalysis. For EAL based on the experiment,\textsuperscript{241} it was shown that the small kinetic isotope effect which led us to conduct our calculations considering the stepwise mechanism. Though it was shown in several studies that the topology of the PEC for Co-C bond cleavage occurs in a different way in enzyme than the solution\textsuperscript{110,254}, we did not see this in EAL. The PEC of Co-C bond is similar inside the enzyme and also in the solution, however, the BDE is much less compared to the solution. The formation and the reactivity of the Ado radical is another aspect of B\textsubscript{12}-catalysis. In MCM, GLM, the Ado radical undergoes conformational changes to facilitate the abstraction. The conformation of the Ado is 2'-endo which changes to 3'-endo for the hydrogen abstraction. In the case of EAL, we have shown that the trajectory of the Ado radical is somewhat different compared to other reported enzymes. Upon the formation of the Ado radical, the rotation of the \( \chi \) angle brings the Ado radical close to the substrate hydrogen atom. The effect of the enzymes, in that case, is to exert strict control for this migration. The conformation of Ado remains the same in that migration process whereas the barrier to abstract the hydrogen atom is 18 kcal/mol. On the other hand, the barrier for hydrogen re-abstraction is also studied in our studies. From figure 5.7 it appears that the hydrogen re-abstraction requires 8 kcal/mol energy which is much less compared to the hydrogen abstraction.
5.4. Summary and conclusions

This study showed the quantitative analysis of the associated intermediates in the EAL reaction mechanism. There are many arguments about the mechanistic insights of B$_{12}$-dependent enzymes. To that line, several studies were conducted which identified different factors responsible for the catalytic activity of B$_{12}$-enzymes. However, a consensus about a common and widely accepted mechanistic insight is still missing. In this study, we have shown the BDE energy of the Co-C bond by employing the QM/MM method. The BDE inside is lowered by a factor of half compare to the BDE of the isolated cofactor in solution. This is consistent with many other previous studies conducted for MCM, GLM, OAM, and EAL using QM/MM, EVB methods. The most apparent commonality with these studies is the absence of ground-state destabilization. The structural parameters of the AdoCbl cofactor inside EAL with the substrate and without the substrate appear to be similar. This indicates that the ground state destabilization upon substrate binding is not the case. However, the reduced BDE inside EAL is due to the electrostatic interactions of the surrounding protein residue with the adenosine ligand of the AdoCbl cofactor. The pseudo-rotation of Ado is also another factor that was considered in our study. We have shown that the hydrogen abstraction from the substrate radical is facilitated by the rotation of the $\chi$ angle of Ado. Inside EAL the long migration of the Ado radical is being possible due to this rotation of Ado ligand. This is controlled by the enzyme active site. The Ado radical is inside the enzyme which appears to prevent the formation of unwanted side reactions. This is also shown by the increase in the energetic barrier during the hydrogen abstraction. The formation of the product and its energetics is also shown in our calculations. Overall each individual step of the native catalysis was studied,
and the associated intermediates were optimized. This is an initial step towards a complete understanding of the EAL catalysis mechanism. Further study, in this case, requires considering the reduction of the AdoCbl cofactor with extensive sampling of the free energy surface. The changes in the energetics by reduction may shed light to consider an alternative approach to elucidate the detailed mechanism.
CHAPTER VI

CAN PHOTOLYSIS OF THE CO-C BOND IN COENZYME B\textsubscript{12}-DEPENDENT ENZYMES BE USED TO MIMIC THE NATIVE REACTION?

6.1 Introduction

Coenzyme B\textsubscript{12}-dependent enzymes catalyze complex molecular transformations where cleavage of the Co-C bond initiates the catalytic cycle.\textsuperscript{1,3} In native catalysis of AdoCbl-dependent enzymes, the formation of the Ado radical is subsequently coupled with H-atom abstraction from the substrate to generate a more stable substrate radical, which together with Co(II), forms the first detectable intermediates via EPR spectroscopy (Scheme 6.1a).\textsuperscript{3,6,7,71,72} However, the Ado radical itself has never been observed experimentally. Alternatively, the Co-C bond can also be cleaved with light.\textsuperscript{152,168,178,241} In both cases, the rupture of the Co-C bond results in the formation of Co(II)/Ado radical pair. The photoinduced cleavage does not require the substrate but will still generate the same radical pair as in the case of native enzymatic catalysis (Scheme 6.1b). The generation of radical pair by light has been investigated using different experimental and computational techniques both for isolated AdoCbl in solution and for several coenzyme B\textsubscript{12}-dependent enzymes.\textsuperscript{124,155,177-180} Based on these studies it appears that the mechanism of photodissociation depends on several factors. Specifically, the environment of the cofactor
plays a decisive role. Inside the enzyme, radical pair generation is energetically favored which contrasts with the isolated cofactor in solution.

Scheme 6.1. a) The homolytic cleavage of the Co-C bond in the presence of a substrate for AdoCbl-dependent enzymes. b) The homolytic cleavage of the Co-C bond to generate Co(II) and Ado radical in the presence of light without the binding of a substrate molecule.

However, the rate of recombination increases in the enzyme. As a result, the QY of the radical pairs in enzymes is lower than the QY of radical pairs in the isolated cofactor. Regardless of the mechanism of Co-C bond activation, the generated radical pairs are the same in both solution and the enzymatic environment.

The observation that light-induced cleavage and the native catalytic cycle results in the same radical pair formation has led to a simple intuitive question: can photolysis be used to mimic the native reaction that occurs inside an enzyme? While the photochemistry of cobalamins has been used to elucidate the electronic properties of B_{12} cofactors, the formal connection between the native reaction and photo-cleavage of the Co-C bond has
been lacking. In order to establish this connection using a computational approach, it is important to consider another less obvious question: are the topologies of the PESs similar? which can be used to describe the reaction pathways associated with radical pair formations. For this to be determined, it is necessary to understand the PESs associated with both photolytic cleavage and native catalysis. It is rather straightforward to pinpoint the coordinates for PESs corresponding to photolytic cleavage and these include both axial ligands. However, it is not immediately apparent how to construct PESs that are representative of the native catalytic mechanism.

Herein, the possible connections between the native reaction and photolysis of AdoCbl-dependent enzymes, using ethanolamine ammonia-lyase (EAL) as a test case, will be presented. EAL belongs to the eliminase sub-class of AdoCbl-dependent enzymes. It utilizes ethanolamine (EA) as a substrate and AdoCbl as a cofactor. There are several reasons why AdoCbl-dependent EAL was selected to demonstrate the possible connection. First, the high-resolution x-ray structure of EAL is available and deposited in the Protein Data Bank (PDB). Second, the detailed photochemical study of AdoCbl-dependent EAL has been conducted experimentally using transient absorption spectroscopy (TAS) and these results have been compared with insights from computational studies in our previous section. Third, it has been suggested binding of a substrate in EAL does not lead to significant conformational changes. Finally, a hypothesis has been proposed by Warncke and co-workers that aims to provide the connection between photolysis and native catalysis, using EAL as a model, which can be compared with our current findings. In addition, establishing the connection between photolysis and catalysis in EAL can help to elucidate the mechanisms of B_{12} photoreceptors, such as CarH. The structure of EAL with
the substrate-bound and substrate-free model were considered in our investigations as these represent catalysis and photolysis, respectively. A combined QM/MM framework was applied to construct the corresponding PESs for the model systems as a function of axial bond lengths.

6.2 Computational details

The crystal structure of AdoCbl-dependent EAL was obtained from the protein data bank (PDB ID: 3ABS).\textsuperscript{53} It is comprised of two asymmetric units and each contains EA as the substrate and AdePeCbl as a cofactor. One asymmetric unit was used for the calculations as they both contain the same cofactor and the substrate and also the reactive unit is independent and far from each other. In order to build the realistic structural model, the AdePeCbl was replaced with the AdoCbl cofactor. The structure was then protonated to pH 7.0 using PropKa with the PDB2PQR web server.\textsuperscript{201,202} The protonation states of the histidine were determined based on the protonation at either the delta or alpha N position, such as HID or HIE. Thus, the protonation of the histidine are as follows, HIE 227A, HID317A, HID349A, HID368A, HID410A, HIE62B, HIE263B. For the photolysis, the substrate was removed and replaced with two water molecules to make sure the activation is occurring through light and for the catalysis study, the substrate was preserved. Both of these structural models were then minimized with 5000 steps steepest descent algorithms in the Amber suite of programs using the Amber force field (FF14SB).\textsuperscript{205,206} The Amber parameters for the AdoCbl cofactor was obtained from a MM study by Marques et al.\textsuperscript{230}

The minimized structures were then used for the QM/MM calculation set up. For the substrate-free model, the model system was divided into three layers. The high layer part contains the truncated corrin ring, Co ion, the Im part of the DBI base and the Ado
moiety. The rest of the cofactor along with the two water molecules was placed into the middle layer and the protein part together with the crystal water molecules was placed into the low layer part of the system. Residues within 20 Å from the cofactor were allowed to move whereas the rest of the protein part was frozen. The substrate-free model used in the calculations contained a total of 11998 atoms with 3550 unfrozen and 8438 frozen atoms. In the case of substrate-bound model, the partitioning of the ONIOM layers is as follows, the truncated corrin ring, the Ado moiety, the Im portion of the DBI base and the substrate EA was placed into the high layer, rest of the cofactor was in the middle layer and the protein part were included into the low layer portion. In total, the substrate-bound model contained 3555 unfrozen atoms and 8438 frozen atoms altogether 11993 atoms in the prepared model. The substrate-free model was used to construct the PECs, the S₀ PES and the S₁ PES. The substrate-bound model was used for the construction of D₁ PES which involved the addition of an extra electron to the AdoCbl cofactor for the reduced model.

The geometries of EAL substrate-free and bound models were optimized with the DFT/PM6/MM level of theory. For the high layer part of the system BP86 functional was used with the TZVP basis set for H and TZVPP basis set for Co, C, N, and O.²¹⁶,²¹⁷ The details regarding the use of this level of theory has already been discussed in the Computational details' section 2.1-2.6. To construct the S₀ and D₁ PESs, the axial bonds were systematically elongated using a step size of 0.1 Å. The upper axial Co-C bond was systematically elongated from 1.8 Å to 2.6 Å. The lower axial bond Co-N_{Im} was elongated from 1.9 Å to 3.5 Å. For the substrate-free model, the S₁ PES was constructed based on vertical excitations and was used to describe the photodissociation mechanism. This was compared with the D₁ PES based on the substrate-bound model of EAL. TD-DFT/MM
calculations were used to construct the $S_1$ PES. The ground state optimized geometries were used to conduct the single point TD-DFT/MM calculations. The detailed orbital analysis was performed to verify that the selected transitions in TD-DFT/MM calculations were not underestimated. In addition, density-based matrix $D_{CT}$ analysis was performed to further ensure that the state selection was accurate.\textsuperscript{225}

6.3 Results and discussions

6.3.1. Potential energy curves as a function of Co-C bond lengths. The initial step of photolysis or catalysis in AdoCbl-dependent enzymes is the homolytic cleavage of the Co-C bond. In the case of photolysis, the activation of the Co-C bond is induced by light whereas in the case of enzymatic catalysis the activation through substrate triggering.\textsuperscript{178} The straightforward way to elucidate the reaction mechanism is to investigate the PECs as a function of Co-C bond distance. To describe the photodissociation mechanism, such PECs were constructed by systematically elongating the Co-C bond using the step size of 0.1 Å for the substrate-free model of AdoCbl-dependent EAL (Figure 6.1). The DFT/MM method was applied to optimize the ground state geometries at each point. The singlet excited states for the substrate-free model of EAL were computed using the TD-DFT/MM method for each optimized point from the $S_0$ PEC. The singlet states using the Co-C coordinate were found to be non-repulsive (Figure 6.1). The $S_1$ state is characterized as having dominant transitions from metal d orbitals to the corrin $\pi^*$ and can be characterized as metal-to-ligand charge transfer (MLCT).
Figure 6.1. Potential energy curves (PECs) of lowest singlet state along with the low-lying singlet excited state as a function of Co-C bond length for AdoCbl-dependent EAL without the substrate.

A similar approach to describe the mechanism of photodissociation has been applied to other systems where the Fe-CO bond for the CO-ligated models of myoglobin and Cr-CO bond for phosphine-substituted transition metal carbonyl complexes \( \text{Cr(CO)}_5 \text{PH}_3 \) was used as the coordinate to construct PECs.\(^{129,257}\) While the use of PECs as a function of only Co-C bond distance has provided a vital undertaking to describe the photodissociation mechanism; it is apparent from more extensive investigations that another coordinate need to be considered. Based on computational studies photodissociation can be fully described by considering both axial ligands of the AdoCbl
cofactor.\textsuperscript{178,179} This is because photodissociation takes place from the ligand field (LF) electronic state with an elongated lower axial bond or even a partially detached axial base. Also, it has been verified through XANES measurements of CNCbl that upon excitation, the largest structural changes that occur were in the case of axial ligands.\textsuperscript{161} Thus, it is crucial to explore the PES as a function of axial bond distances in order to understand the mechanism clearly.

6.3.2. Low-lying excited state PES as a function of axial ligands. Upon excitation with light, a higher-order excited state is initially populated, and this is followed by radiationless decay through several metastable states to the lowest excited state $S_1$. The $S_1$ state is the most crucial excited state to consider in terms of photolytic properties as both photodissociation and geminate recombination occurs from here. These characteristics features of the $S_1$ state can be depicted clearly by a PES as a function of the axial ligand distances including the Co-C and Co-N$_{im}$.\textsuperscript{178,179} In order to construct such a PES, the axial bond distances were systematically elongated using a step size of 0.1 Å and the ground state geometry was optimized at each point using DFT/MM. The $S_1$ PES was constructed using the TD-DFT/MM framework and it was based on the vertical excitations from the optimized $S_0$ state surface. Figure 6.2 represents the $S_0$ and $S_1$ surfaces along with the vertical projections of both surfaces.

The $S_0$ surface was characterized as having a single energy minimum, whereas the $S_1$ PES has two energy minima regions which result from the crossing of two electronic states namely MLCT and LF. The MLCT minima regions arise from the electronic transitions associated with Co metal d orbitals to the corrin $\pi^*$ orbitals. The bond length of
Figure 6.2. Potential energy surface along with the vertical projections of $S_0$ and $S_1$ surface for the substrate-free model of AdoCbl-dependent EAL. a) Characteristics points were located in the $S_1$ PES and shown in the figure. Radical pair generation occurs from the LF state. The radical pair is shown in figure as Co(II) and Ado radical. b) The relative energy differences between different states and the characteristics points were shown.

the Co-C and Co-N$_{Im}$ are essentially unchanged in that case. In the case of LF state, the axial bond distances were elongated, and the electronic transitions are the result of metal d orbitals to the d$_{z^2}$ orbitals. This is shown in figure 6.3. These two regions are separated by a seam that we identified as the minimum energy crossing point (MECP) and the barrier height to cross the MECP varies with varying environmental conditions of the cofactor. In the case of the enzymatic environment, the energetic barrier is lower compared to the solution medium. Thus, it appears that the radical pair formations in the LF region is easier inside an enzyme in comparison to the isolated cofactor. The generation of radical pairs occurs from the LF state. Before this, excitations with light results in a population of the MLCT region (Figure 6.2). The MLCT minimum is 1.2 kcal/mol higher than the LF region. In order to determine the photodissociation mechanism, we identified two different pathways to connect the MLCT and LF state through an MECP shown in figure 6.2b. In
the case of EAL, the barrier to cross the MECP is 2.5 kcal/mol which is much lower than the isolated cofactor.

![Diagram](image)

**Figure 6.3.** The 2D contour representation of the first singlet excited state ($S_1$). The PES was represented as a function of two axial bond lengths. The $S_1$ surface is characterized as having two minima regions namely metal-to-ligand charge transfer (MLCT) and the ligand field (LF) region for the isolated cofactor as well as AdoCbl inside enzyme. The radical pair generate from the LF state. Two pathways, Path A and Path B, were identified that connect the MLCT to the LF state.

### 6.3.3. Ground state PES as a function of two axial coordinates.

The native catalytic reaction of AdoCbl-dependent EAL involves the homolytic cleavage of the Co-C bond as the initial step of the catalytic cycle.\(^3\) The cleavage of the Co-C bond results in the formation of the Co(II) and Ado radical which is followed by the subsequent H-atom abstraction from the substrate. This event occurs in the presence of the substrate involving the ground electronic state. Thus, in order to understand this reaction, it is crucial to scrutinize the ground electronic state. With this mind, we initially optimized the ground state of EAL using DFT/MM methods. The optimized structure was then compared with the crystal structure and also the optimized structure obtained from the substrate-free model. The noted structural changes were observed in the cases of two axial ligands of the
cofactor whereas the corrin structure is slightly perturbed. This is shown in figure 6.4.

**Figure 6.4.** Active site figures of AdoCbl-dependent EAL. AdoCbl cofactor is depicted as ball and stick model and the protein residue using ribbon. a) Crystal structure of AdoCbl-dependent EAL. The crystal structure was obtained from the Protein Data Bank (PDB ID-3ABS) b) Optimized structure of AdoCbl-dependent EAL with the substrate molecule. The substrate EA depicted as ball and stick model along with the cofactor. The $S_0$ PES of the substrate-bound model was initiated from this optimized structure.

Based on this observation we used Co-C and Co-N$_{Im}$ as the corresponding coordinates for constructing the $S_0$ PES for depicting the ground state reaction. Using that idea, we constructed the $S_0$ PES for EAL with the substrate-bound model and compare it with the PES for the substrate-free model. The topology of the $S_0$ PES for both the substrate-bound and substrate-free models are indistinguishable (Figure 6.5). The $S_0$ PES both contain a single energy minimum corresponding to the Co(II) state. By increasing the Co-C and Co-N$_{Im}$ bond distance from the energy minimum, the energetics of the $S_0$ PES increase simultaneously.
Figure 6.5. The 2D contour representation of the PESs. The relative energies were shown above all the figures along with the color code. The relative energy unit is in eV a) The $S_0$ PES of the model system without the substrate molecule. b) The $S_0$ PES of the substrate-bound model system.

6.3.4. Assessment of our hypothesis with literature. Warncke and co-workers established one of the best intuitive approaches to connect photolysis and native catalysis.\textsuperscript{168} In order to depict that, a hypothetical PES was constructed where photolysis was explained with respect to the Co-C bond. On the other hand, to explain the native reaction they invoked a second orthogonal coordinate, namely protein configuration. The topology of the PES was shown to have two minima regions, namely Co(III)/Ado and Co(II)/Ado (Scheme 6.2). Furthermore, it was suggested that the actual coordinate for the native reaction would be the result of cooperativity between the Co-C stretch and protein configuration. Although this was an important step towards understanding the connection between photolysis and native catalysis, based on our study it is not representative of the full picture.
In our assessment, there are several issues that should be considered when photochemistry of B$_{12}$-dependent enzymes is used to imitate the native reaction. First, the photochemical cleavage of the Co-C bond and the thermal Co-C bond cleavage involves different electronic states and timescales. Second, the lower axial ligand is a crucial coordinate for representing the mechanism of photolysis. Third, the protein configuration component plays a role, but this cannot be easily considered as a coordinate because it cannot be readily quantified. Besides, it has been demonstrated that large conformational changes do not occur upon substrate binding in EAL.$^{53,72,168}$ This is not to preclude that other protein contributions may play a role in the native catalytic cycle. To address the activation of the Co-C bond in AdoCbl-dependent systems, a number of multiscale
quantum chemical calculations have been performed including QM/MM and empirical valence bond (EVB). While it is apparent that the Co-C bond is necessary for the construction of the PES, it is not immediately clear based on these studies how to account for the role of the protein. Taking all of this into consideration, it seems that using protein configuration as a coordinate is not the best representation of either photolysis or the native reaction in EAL. Still, the role of protein configurational changes can be seen in other ways and our QM/MM studies have confirmed this. In our calculations, the protein residues could respond freely while constructing the PES as a function of axial bond lengths Co-C and Co-N_{Im} instead of protein configuration. We have observed that the protein environment is affecting the energetics associated with photolysis. The S_{1} PES for the isolated cofactor and the S_{1} PES for EAL differs in terms of energetics. Inside the enzyme, the LF minimum is energetically lower than the MLCT minimum, whereas the opposite scenario is observed for the isolated AdoCbl cofactor. The role of the protein is, in our opinion, to stabilize the LF state and possibly, to keep the AdoCbl cofactor in a specific configuration where the corrin ring and the adenine ring are fixed in space and the only free rotation is with the ribose moiety.

Photolysis of the Co-C bond occurs in the excited state, this is well understood both experimentally and computationally. Two coordinates are needed to explain photolysis of AdoCbl inside EAL, not just one (Figure 6.3). The first step in photolysis would be the displacement of the axial base and then Co-C elongation in a diagonal along the PES until the second energy minima LF is reached (Figure 6.3). Considering one coordinate to describe photodissociation is simply not enough because the cleavage of the Co-C bond by light involves the cooperativity between both Co-axial bonds. As discussed earlier, both S_{0}
PESs for the substrate-bound and substrate-free models contain a single energy minimum which corresponds to the Co(III) state. Based on our calculations, there is no obvious explanation of how the $S_0$ PES can have two minima. Should a second minimum be present, it would be the result of a different electronic state. In this regard, the hypothesis proposed by Warncke, *et al* and the connection between photolysis and native catalysis cannot be justified simply by our DFT/MM calculations.\(^{168}\)

6.3.5. **The connection between photolysis and catalysis.** The possible connection of photolysis and catalysis can be obtained clearly by scrutinizing the PES associated with photodissociation and the PES representative of the native reaction. According to our study, the $S_0$ PES associated with the native reaction cannot be directly compared with the $S_1$ PES for the native reaction. However, we can offer an alternative explanation. The $S_1$ surface from TD-DFT/MM calculations contains two electronic states, namely the MLCT and LF states (Figure 6.3). The MLCT is a result of electron excitation from Co to the corrin ring and does not involve significant geometrical changes. On the other hand, the LF state involves excitations to $N_{Im}$-Co-C axial bonding. It requires geometrical changes, mainly elongation along with the axial bonding. Alternatively, a similar electronic structure, which would result in a similar topology of the $S_1$ PES, can be generated when one electron is added to the ground state. This corresponds to the one-electron reduced PES. To demonstrate this, we constructed a PES ($D_1$) for AdoCbl-dependent EAL by adding one electron, as shown in Figure 6.6. When comparing this side-by-side to the Warncke *et al.* PES, the topological features are the same. The $D_1$ surface from the DFT/MM calculations, where an electron was added to the cofactor, contains two minima regions which are also the result of two distinct electronic states.
Figure 6.6. The PESs of S₁ and D₁ state along with the contour representation. a) The excited state (S₁) PES of the model system without the substrate molecule. TD-DFT/PM6/MM method was applied to construct the S₁ surface. Characteristic points MLCT and LF has shown in the surface plot along with the spin densities. The spin density of these two states were shown right above the minimum regions. b) The doublet ground state (D₁) PES of the model system with the substrate molecule calculated using DFT/PM6/MM. One extra electron was added to reduce the cofactor of that model. The D₁ PES was also characterized as having two minima regions.

It is obvious from Figure 6.6 that the D₁ PES is a result of the crossing of two electronic states. The two minima regions in the PES represents two distinct electronic states. One minimum I(D₁min) is localized at Co-C bond distance of ~2.05 Å and Co-Nₐmin ~2.90 Å. Spin density distribution is observed on the corrin ring. This indicates that the electron is delocalized onto the corrin ring after addition. The structure of the minimum was further optimized without constraining the Co-C and Co-Nₐmin bond (Figure 6.7). The
Co-N$_{Im}$ bond is lengthened at the minimum of the reduced model which is also observed in the x-ray crystal structures of AdoCbl-dependent EAL. Thus, the enzymatic catalytic efficiency may be induced by a one-electron reduction of the cofactor. Upon elongation of the Co-C and Co-N$_{Im}$ bond, the localization of the electron shifts from the corrin $\pi^*$ to the $\sigma^*$. The second minimum III(D$_{1\text{min}}$) is localized around the Co-C bond length of $\sim$2.30 Å and Co-N$_{Im}$ length of $\sim$3.10 Å. The geometry of this minimum was optimized without constraining the axial bond lengths (Figure 6.7). This corresponds to the $\sigma^*$ state which is characterized as $(\sigma)^2(\sigma^*)^1$. The cleavage of the Co-C bond, in that case, proceeds via the transfer from a $\pi^*$ corrin type state to the $\sigma^*_{\text{Co-C}}$ state. This is seeming to be consistent with the $S_1$ PES for the photodissociation of the Co-C bond.

**Figure 6.7.** The optimized structures of the two minima. DFT/PM6/MM were applied to optimize the structures a) Optimized structure of AdoCbl-dependent EAL model system with the substrate at $\pi^*$ minimum (I (D$_{1\text{min}}$), Figure 6.6b) in D$_1$ PES. b) Optimized structure of the model system in $\sigma^*$ minimum (III (D$_{1\text{min}}$), Figure 6.6b).
6.3.6. Potential role of the reduced cofactor in native catalysis. A natural question that may arise based on our previous discussion is, ‘is the reduction of the cofactor physiologically relevant?’ This concept for the reduction of the cofactor has precedence in the literature. Lexa and Savant proposed that an electrochemical addition of an electron to isolated vitamin B\textsubscript{12} derivatives reduces the bond dissociation energy by a factor of one half.\textsuperscript{248,249} Thus, it was hypothesized that the activation of the Co-C bond in coenzyme B\textsubscript{12}-dependent enzymes can be induced by the reduction of the cofactor. In the enzymatic environment, different protein residues may serve as internal redox centers to facilitate the transfer of an electron to the cofactor. Based on previous studies of AdoCbl-dependent methylmalonyl Co-A mutase (MCM) and glutamate mutase (GLM), it was shown that a tyrosine (Y) residue in the vicinity of the substrate may play a role in the electron transfer (ET) process.\textsuperscript{258,259} In order to be energetically feasible, ET has to be coupled with proton transfer (PT) otherwise known as proton-coupled electron transfer (PCET). In other biological systems such as photosystem II, cytochrome c oxidase, and ribonucleotide reductase, to name a few, Y has been proposed to play a role in the PCET process.\textsuperscript{252,253,260} In the case of GLM and MCM, it has been proposed that the Y residue initiates PCET by transferring phenolic hydrogen to the substrate. This is followed by ET to the cofactor. The extra electron in the cofactor is initially delocalized on the corrin ring. Further elongation of the axial bond’s Co-C and Co-N\textsubscript{Im} leads the electron to the antibonding $\sigma_{\text{Co-C}}$ orbitals. The cleavage of the Co-C bond occurs from the three electrons $(\sigma)^2(\sigma^*)^1$ state as the bond strength is significantly reduced in this case.

Based on these studies, the role of the Y residue is to initiate the reduction of the cofactor by one electron. Accordingly, we considered the reduced cofactor and modeled
our system in an enzymatic environment. It is important to mention that the native catalysis reactions occur in the ground state which follows $S_0$ chemistry rather than $D_1$ chemistry. The crucial thing to be noted here is that the potential involvement of Y to initiate cofactor reduction which results in an $S_0$ PES. Transfer of an electron from Y residue to the cofactor would generate a $[\text{AdoCbl(II)}\cdots\cdots\text{Y•}]$ diradical complex where the overall multiplicity is a singlet. Thus, it is true that the native reaction follows $S_0$ chemistry, whether or not PCET is occurring. On the other hand, the addition of the extra electron, without considering the diradical nature, will lead to doublet multiplicity. Hence the PES constructed based on that model will be the $D_1$ PES. The intriguing outcome of our studies is to observe the similarities of the PES for the reduced cofactor and the excited state PES. The $S_1$ and $D_1$ PESs based on electronic excitations via light and addition of an extra electron, respectively, result in similar topologies. The $S_1$ and the $D_1$ PESs are the results of the crossing of two electronic states and for both cases, these states are similarly characterized (Figure 6.6). It would appear that the photolytic cleavage of the Co-C bond can be used to mimic native catalysis reactions in a way that involves a reduced cofactor. This conclusion is an important step towards understanding the connection between photolysis and native catalysis. Additional studies are required to scrutinize this possibility further in AdoCbl-dependent enzymes.

6.4. Concluding remarks

Our calculations have shown that the mechanism of photolysis of AdoCbl-dependent EAL and the corresponding $S_1$ PES cannot be used as a direct comparison to the neutral PES ($S_0$) from the native reaction. Regardless of the nature of the second coordinate orthogonal to Co-C, there is no distinct mechanism to stabilize the Co(II)/Ado• radical
state, even inside an enzyme. However, the photolysis of the Co-C bond can indeed generate radical pairs, and this can potentially be useful to study the reactions which involve the cleavage of the Co-C bond of AdoCbl in ground state processes. While this is an intuitive connection between photolysis and native catalysis, based on our calculations, alternative explanations should be considered. As we have shown, photolysis can be used to make comparisons between the $S_1$ PES and the $D_1$ surface. In other words, the $S_1$ PES based on photochemical cleavage of the Co-C bond and the $D_1$ PES based on the one-electron reduced AdoCbl (Figure 6.6) are very similar in terms of topology and thus are similar in terms of the reaction mechanism. The connection between photolysis and the native reaction of AdoCbl-dependent enzymes is more unique than previously thought. As our results indicate, it involves considering the native reaction in terms of the one-electron reduced form. Although it does not appear that a one-to-one connection between photochemistry and catalysis in an enzyme such as EAL is not plausible, this sort of connection may be found in the case AdoCbl-dependent CarH which directly uses light to regulate DNA transcription.
The chemistry of vitamin B\textsubscript{12} is exceptionally challenging and as well as interesting. Since the discovery of vitamin B\textsubscript{12} in 1926, it remains an important issue of study. There are many questions of B\textsubscript{12} chemistry left unanswered. Moreover, with the advancement of technology, it is becoming clear that vitamin B\textsubscript{12} and its derivatives can be utilized for many different applications. The scopes of utilizing these molecules are in the drug delivery, harnessing optogenetic tools, photodynamic therapy, catalyst design, and many other purposes. In order to apply to these fields, one of the crucial factors is to understand the underlying mechanism of the reaction that is occurred in vitamin B\textsubscript{12} derivatives and its dependent enzymes. The two most important reactions that are commonly known and studied in B\textsubscript{12} chemistry, are photolysis and catalysis reactions. Among many derivatives of vitamin B\textsubscript{12}, two are biologically active and plays a crucial role in enzymatic catalysis. The biologically active derivatives are AdoCbl and CH\textsubscript{3}Cbl and act as a cofactor in numerous enzymatic reactions. This study focused on the understanding of AdoCbl-dependent enzymes. Two critical reactions, photolysis, and catalysis for two different AdoCbl enzymes were studied such as EAL and GLM.

EAL and GLM both are AdoCbl dependent enzymes belongs to two different classes. EAL is eliminases class of AdoCbl enzymes whereas GLM is a mutase class of enzymes. EAL catalyzes the conversion of ethanolamine to aldehyde and
ammonia whereas GLM catalyzes the conversion of S-glutamate to (2S-3S) methyl aspartate. Although these enzymes are bacterial enzymes and the reactions occur in bacteria, but the reactions itself are chemically and industrially important. Chemically speaking the enzymes utilize the AdoCbl radical chemistry to catalyze these thermodynamically unfavorable reactions. AdoCbl is a complex organometallic molecule with a unique Co-C σ bond. Inside the enzyme, the initial step of the catalysis is the homolytic cleavage of the Co-C bond to generate the radical pairs. The rate of the Co-C bond cleavage is accelerated by a trillion-fold inside enzyme compare to the isolated factor. There are many factors responsible for this high rate acceleration. On the other hand, the Co-C bond can also be activated by light inside enzyme and in the isolated cofactor. The mechanism of photodissociation varies depending on the environment of the cofactor and also the axial ligand. Thus, photochemistry of CH$_3$Cbl and AdoCbl are largely different as well as the photochemistry of isolated cofactor and inside enzymes.

In this dissertation, the photochemical mechanism of Co-C bond cleavage in EAL and GLM were studied using the different computational methods. In order to understand the effect of the environment, the isolated AdoCbl cofactor was studied and compared with the enzyme. The motivation for studying these two enzymes is that they catalyze two different reactions and belong to different classes of AdoCbl enzymes. Understanding these reactions will allow gaining a complete insight into the AdoCbl photochemistry in enzymatic environment. Moreover, the high-resolution x-ray crystal structures are available for these two enzymes which are the fundamental basis for our computational modeling. Besides, TAS and spectroscopic studies were conducted for EAL and GLM
which can be directly compared with our computation. These are the basic concept behind our studies.

For studying the photochemical mechanism, we have shown that the low-lying electronic states are responsible for the cleavage of the Co-C bond. To that end, we constructed the PEC along the Co-C bond. Based on the PEC, it appears that the cleavage of the Co-C bond cannot be explained using one-dimensional curve. Based on the analysis of the electronic transitions it appears that at equilibrium geometry the $S_1$ state electronic transitions are predominately from the metal d orbitals to the corrin $\pi^*$ orbitals. The electronic transitions in the $S_1$ state thus described as MLCT transition. This is consistent with the results of MLCT assignment for the $S_1$ state based on the TAS experiments. Thus, using BP86 functional is important to use for studying photochemistry. On the other hybrid functional underestimates, the MLCT assignment and the electronic transitions based on the B3LYP are $\pi$ to $\pi^*$. This inconsistency with the experiment further motivates us not to use hybrid functional for studying the photochemistry of AdoCbl enzymes.

To further understand the effect of the environment on the excitation energy, many folds of low-lying singlet excited states were analyzed in EAL and GLM and compared with the isolated cofactor. In the case of enzymes, the excitation energy is lower compared to the isolated cofactor. Thus, the role of the enzyme is to stabilize the electronic states. The orbital analysis was performed to characterize the nature of electronic states. In the case of equilibrium Co-C bond length, the lowest energy electronic states are MLCT in nature whereas at elongated bond length the electronic transitions in the lowest energy electronic states are predominately from the metal d orbitals to metal $d_{z^2}$ orbital. The electronic transition associated with metal $d \rightarrow d_{z^2}$ ascribed as LF state. The dissociation of
the bond occurs from the LF state as it represents three electrons weaker bond and lowest energy. This orbital characterization is consistent with experiment and previous theoretical studies.\textsuperscript{124,155} The motivation of our study is to find the detailed mechanism of the photodissociation and using this one-dimensional coordinate is not enough to fully describe the mechanism. In order to understand this detail mechanism, two-dimensional PES were constructed along the Co-C and Co-N\textsubscript{Im} bond length. This was done because of the sensitivity of the cofactor axial ligands upon binding with the enzyme. In addition, TAS studies also showed the light sensitivity of the axial ligands.\textsuperscript{161} Upon shining with light, the largest structural changes occur in the cofactor were in the case of axial ligands whereas the structure of the corrin ring is unperturbed. TD-DFT analysis of isolated AdoCbl cofactor also indicated that the optimized geometry of the \textit{S}_1 states also undergoes the largest structural changes in the axial ligands. Thus, it is crucial to explore the photolysis mechanism with respect to the axial ligands.

The analysis of the TAS and TD-DFT data proved that the \textit{S}_1 state is responsible for the photolytic cleavage.\textsuperscript{123,124,155} The radical pair generation and the recombination occurs from the \textit{S}_1 state. With that in mind, the \textit{S}_1 PES of the AdoCbl cofactor inside enzyme and also isolated base-on and base-off were constructed, analyzed and compared. The topology of the \textit{S}_1 PES appears to contain two energy minima. At the equilibrium bond length, the nature of the electronic transition is MLCT type. At elongated bond length, the minima region is of LF type. A seam separates the two minima regions. This topology is similar inside the enzyme and also in the isolated cofactor. Inside the enzyme, the energetics are lower compared to the isolated cofactor. The role of the enzyme is to stabilize the LF state to facilitate the radical pair formation. The active site of the enzyme contains
protein residue such as glutamate, asparagine, serine which is in contact with the Ado ligand. The electrostatic interactions with the Ado ribose moiety play the role for the LF stabilization. This is also evident from the spectroscopic study where it is shown that the mutation of the E287 residue which Vander Waals contact with the Ado moiety effects the photolysis of the Co-C bond in EAL.\(^1\)

Based on the analysis of the \(S_1\) state it is clear that the \(S_1\) state is comprised of two different electronic states, namely MLCT and LF. The most important aspect to describe the photodissociation is to connect the MLCT and LF state. We have developed a method to establish a connection between these two states. There are two pathways were identified. Path A and Path B, where both Paths initiates from the MLCT state and ends at the LF. In the case of MLCT state, the initial elongation is the Co-C bond which is then followed by both axial ligands to cross the seam to go into the LF state. In Path B, the initial elongation is the Co-\(N_{Im}\) bond which is then followed by the elongation of both axial bonds. In LF state where both axial bonds are elongated, the radical pairs can escape the cage or recombine to the ground state. Our calculations and analysis have shown that in enzyme EAL path B is more favorable than Path A, and in the case of GLM, both pathways are equally possible. This is shown based on the analysis of the corresponding energetics in each path. In the solution environment, there is only one possible path that the radical pair can undertake. For base-on AdoCbl the only possible path is Path A whereas for base-off it is Path B. On the other hand, the enzymatic environment promotes the radical pair generation in both the pathways.

From LF there are two possibilities, radical pair separation or the recombination to the ground state. The competition between these two processes dictates the QY of the
radical pair. Based on our studies, we have seen that the energy of the LF state inside the enzyme is lower which indicates that the generation of radical pair in LF is more natural. However, it is difficult to escape the cage from the enzyme. That is why the rate of recombination increased inside the enzyme. Whereas in solution, the QY is much higher compare to the enzyme. Thus, it is challenging to generate radical pairs in LF but once it generated, it can easily escape the cage. Nature controls the AdoCbl enzymes in a way that prevents the formation of this free radical. The cage effect of the enzyme controls the radical pair diffusion. Though it was initially described that the unusual activity of the B\textsubscript{12} catalysis is controlled by the cage effect of the enzyme. This idea was not accepted by both experimentalists and computational chemists. Our study also suggests that the idea of cage effect is not active for the cleavage of the Co-C bond inside enzyme; however, upon the generation of the radical pairs, the trajectory is controlled by that. That is why photolysis is not the physiological reaction for AdoCbl enzymes. But in the case of CarH, an AdoCbl based photoreceptor utilize the photochemistry of Cbl to control the transcription of the DNA. In the CarH, the photolysis of the Co-C bond is associated with the large-scale conformation changes. The role of the protein, in that case, is to facilitate the radical pair generation which undergoes into different chemistry.

This idea and insights from photochemistry are very critical and have implications while studying the catalysis mechanism of B\textsubscript{12}-enzymes. Here in this study, the native catalysis mechanism of AdoCbl enzyme EAL was studied using a combined QM/MM method. As discussed earlier, AdoCbl acts as a cofactor and a potential source of radicals for EAL catalysis. The catalysis mechanism in EAL consists of various intermediates and unstable species. The initial step is the fission of the Co-C bond to generate Ado radical
and Co(II). We have shown that the BDE for Co-C fission is reduced by a half compared to the isolated cofactor. The trajectory of the Ado radical for the hydrogen abstraction was also shown. The barrier for the hydrogen abstraction is 18 kcal/mol. The reduction of BDE by the AdoCbl enzyme is one of the crucial aspects of B$_{12}$ catalysis. Based on this study it was shown that the ground state destabilization as hypothesized before is not the case for this rate enhancement. This is also established previously with some other great works.$^{109,111,113,254}$ In the case of EAL, the reactions occur in a stepwise manner which is established by some pioneering works.$^{109,241}$ Based on the kinetic isotope study it was found that in EAL the kinetic isotope effect is very less. Thus, the reactions may proceed in a stepwise manner compared to other AdoCbl-dependent enzymes where a concerted mechanism is shown to be effective.$^{109}$ Overall the catalytic effect in EAL simply cannot be described by one factor. The electrostatic effect of the enzyme induces the dissociation of the Co-C bond. This is due to the electrostatic interaction of Ado ligand with the surrounding protein residue. However, upon the generation of Ado radical, the enzyme environment exerts strict control over the trajectory of Ado radical for the hydrogen abstraction. This is known as negative catalysis by Retey concept where it is shown to be effective to remove unwanted side reactions. In our photochemical study of EAL and GLM we have shown that the cage effect is crucial for the AdoCbl photodissociation. Consistent with that we believe that the cleavage of the Co-C bond is induced by the electrostatic interactions.

Although it appears based on some other studies, that the main contributions in the catalysis is coming from the entropic effect, however, the origin of this effect was shown to be electrostatic. Our results and discussion are consistent with this concept. However,
once cleavage of the Co-C bond is complete, the next hydrogen abstraction step is happened to be steric and electrostatic together. Due to the cage effect exerted by the enzyme the Ado radical undergoes by strict control. In that case, the trajectory of the Ado radical is controlled by the enzyme active site residue. In the case of EAL, the rotation of χ angle about 70° brings the C5 atom of Ado radical in a close distance to the hydrogen atom of the substrate. The orientation of the Ado, in that case, was found to be 3’-endo. The relative energetics of all the intermediates formed in the catalysis were identified and the energetics were shown (Figure 5.7). The hydrogen re-abstraction by the substrate radical was shown and the energy barrier to abstract the hydrogen atom from the adenosine was found to be 8 kcal/mol. This is reduced by a factor of half compare to the hydrogen abstraction process. Thus, the migration of radical inside enzyme environment require a higher energy barrier compare to the neutral molecule. Finally, the product structure was optimized, and the relative energetics was shown. Overall the EAL study is important step towards the complete understanding of catalysis.

Based on the study of photolysis and catalysis, it appears that the initial step of reactions in both cases is the homolytic cleavage of the Co-C bond to generate the Co(II) and Ado radical pairs. It is apparent that in both mechanisms, similar radical pairs are generated in different time scales and electronic states. For instance, in the case of photolysis, the radical pairs generate is in the excited state whereas in catalysis ground state is responsible for the cleavage of the Co-C bond. In our studies, we have shown the connection between photolysis and catalysis. Photolysis can be explained by properly describing the \( S_1 \) electronic state and by constructing an \( S_1 \) PES. The \( S_1 \) PES consists of two energy minima regions, MLCT and LF. Now to describe catalysis it is crucial to
investigate the S\textsubscript{0} PES. However, the S\textsubscript{0} PES contains a single energy minimum which cannot be directly compared with the photolysis. In order to alleviate this problem, we reduce the cofactor and constructed a ground state double PES which is called as D\textsubscript{1} PES. The topology of the S\textsubscript{1} PES and D\textsubscript{1} PES are similar. The mechanism of the Co-C bond fission utilizing the two PESs is also similar. In the case of photolysis, the electronic excitation is induced by light and weakens the bond whereas in the case of reduction, the extra electron is from the surrounding protein residue to weaken the Co-C bond. Upon addition of an extra electron, it delocalizes in the corrin ring at equilibrium Co-C bond distance which is shown as $\pi^{*}_{\text{Co-C}}$ state. With the elongation of the Co-C bond, the extra electron moves to the $\sigma^{*}_{\text{Co-C}}$ state and weakens the bond. This mechanism is similar to the mechanism of photodissociation. In order for the reduction, the PCET was shown to be an effective mechanism.\textsuperscript{258,259} In order for the PCET to occur, a Y residue in the vicinity of the substrate plays the role. The Y residue initiates the PCET by transferring the phenolic hydrogen to the substrate, which is then followed by electron transfer to the corrin ring. This is also shown to be effective in many biological systems.\textsuperscript{252,253,260} Finally, the study of the connection between the photolysis and catalysis is more unique than previously thought. According to our results, the consideration of native catalysis reaction in one-electron reduced form thus establishes a fair connection between photolysis and native catalysis.

Overall these studies of photolysis and catalysis is a major step towards the understanding of B\textsubscript{12} chemistry. These contributions provide many novel insights and open the path for future investigations. The photolysis of AdoCbl inside enzymes is well understood now. The quantum yield of photolysis is shown to be dependent on the
magnetic field. We have constructed the PEC for the enzyme EAL, GLM and also for the isolated cofactor.

Further studies required to use the low-lying PEC to calculate spin-orbit coupling, which can be used for the magnetic field effect on photolysis. The magnetic field effect in EAL was shown a long time ago. However, the concept is still debated. Understanding of the magnetic field in EAL may shed light to obtain a bigger picture of radical pair generation. Our studies indicate that the photolysis inside enzymes is inhibited by the cage effect of the enzyme. In order to obtain the quantitative picture, certain residues of the enzymes need to be considered and should be included in the calculations. Inclusions of E287 residue while doing QM/MM calculations can be used to identify the effect of mutations in the photolysis rate. The cage effect can also be investigated to understand its effect on photolysis. This will also give insights into the native catalysis mechanism. We have also shown the connection between photolysis and native catalysis in this study. The catalysis can be connected with the photolysis in a way that involved a reduced cofactor. Thus, the idea of reduction of B_{12} cofactor can be used for further understanding of catalysis. Future studies to understand this effect requires the mutation of N193, E287, Y404 to get insight into the details of the photolysis and the cage effect of enzymes to control the Ado radical.
REFERENCES


(47) Krautler B Chemistry 2015, 21, 11280-11287.


(49) Marsh EN, Patterson DP, Li L Chembiochem 2010, 11, 604-621.


(100) Bender G, Poyner RR, Reed GH *Biochemistry* **2008**, 47, 11360-11366.


171


APPENDIX

Figure A1. Structure of Ado ligand. The exocyclic and endocyclic angles were shown in detail. Glycosidic χ angle is defined as O1′-C1′-N9′-C8′

\[
P = \arctan\left(\frac{\theta_2 + \theta_4 - \theta_1 - \theta_3}{2\theta_0(\sin 36^\circ + \sin 72^\circ)}\right) + X
\]

\[
\theta_m = \frac{\theta_0}{\cos P}
\]

\begin{align*}
\theta_0 &= C1′ - C2′ - C3′ - C4′ \\
\theta_1 &= C2′ - C3′ - C4′ - O1′ \\
\theta_2 &= C3′ - C4′ - O1′ - C1′ \\
\theta_3 &= C4′ - O1′ - C1′ - C2′ \\
\theta_4 &= O1′ - C1′ - C2′ - C3′
\end{align*}

Figure A2. (a) Crystal structure of AdoCbl with corrin ring, the side chains were removed and the axial base is simplified. Crystal structure obtained from Randaccio et al. (b) Optimized geometry of AdoCbl base-on model complex denoted as (Im-[CoIII(corrin)]=Ado). (c) Optimized model of AdoCbl base-off complex denoted as (H2O-[CoIII(corrin)]=Ado)}
**Figure A3.** HOMO and LUMO molecular orbitals for selected points in the $S_1$ PES (Figure 3.13) in addition to the molecular orbitals displayed in figure 3.11

**Figure A4.** The ground state optimized geometries for the selected points in $S_1$ PES shown in figure 3.13
**Figure A5.** (a) Optimized structure of EAL without the substrate. (b) Optimized structure of GLM with the substrate. The structural analysis is shown in Table 4.1

**Figure A6.** Detail molecular orbital pictures to characterize the several points in the $S_1$ PES as shown in figure 4.7 in addition to the orbital analysis in figure 4.6
Figure A7. Energy diagram of frontier KSO for Im-[CoIV(corrin)]-Ado$^2+$ model complex (labeled 1e-OX) at B3LYP/TZVPP level of theory.
**Figure A8.** (a) Isosurface of spin density obtained from B3LYP/TZVPP calculations for the Im-[CoIV(corrin)]-Ado$^{2+}$ (labeled 1e-Ox) model complex. (b) Spin density distribution on fragments of complex based on Mülliken population analysis.

**Figure A9.** The active site figure of EAL. Ado ligand is shown in the figure along with the H-bonding network with the surrounding protein residue.
LIST OF ABBREVIATIONS

Vitamin $B_{12}$  CNCbl
Coenzyme $B_{12}$  AdoCbl
CNCbl  Cyanocobalamin
$\text{CH}_3\text{Cbl}$  Methylcobalamin
AdoCbl  5'-deoxy 5'-adenosylcobalamin
$\text{H}_2\text{OCbl}$  Aquacobalamin
OHCbl  Hydroxycobalamin
$\text{SO}_2\text{Cbl}$  Sulphitocobalamin
Cbl  Cobalamin
HC  Haptocorrin
TC  Transcobalamin
IC  Intrinsic factor
DBI  dimethybenzimidazole
Im  Imidazole
GLM  Glutamate mutase
EAL  Ethanolamine ammonia lyase
DDH  Dioldehydratse
GDH  Glycerol dehydratase
MCM  Methylmalonyl Co-C mutase
OAM  Ornithine amino mutase
LAM  Lysine amino mutase
MetH  Methionine synthase
CH₃-H₄folate  Methyltetrahydrofolate
BDE  Bond dissociation energy
H-atom  Hydrogen atom
KIE  Kinetic isotope effect
Met  Methionine
Hcy  Homocysteine
H₄-folate  Tetrahydrofolate
His  Histidine
E  Glutamate
Y  Tyrosine
N  Arginine
EPR  Electron paramagnetic resonance
rR  Resonance Raman
CD  Circular dichroism
MCD  Magnetic circular dichroism
TAS  Transient absorption spectroscopy
QM  Quantum mechanics
MD  Molecular dynamics
QM/MM  Quantum mechanics/Molecular mechanics
DFT  Density functional theory
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TD-DFT</td>
<td>Time-dependent density functional theory</td>
</tr>
<tr>
<td>CC</td>
<td>Coupled cluster</td>
</tr>
<tr>
<td>MRCI</td>
<td>Multireference configuration interaction</td>
</tr>
<tr>
<td>MP2</td>
<td>Meller plesset perturbation</td>
</tr>
<tr>
<td>ONIOM</td>
<td>Our own-n layered integrated molecular modeling and molecular mechanics</td>
</tr>
<tr>
<td>CASSCF</td>
<td>Complete active space self-consistent field</td>
</tr>
<tr>
<td>LMCT</td>
<td>Ligand-to-metal charge transfer</td>
</tr>
<tr>
<td>MC-XQDPT2</td>
<td>Second-order multiconfigurational quasi degenerate perturbation theory</td>
</tr>
<tr>
<td>LR-CT</td>
<td>Long range-charge transfer</td>
</tr>
<tr>
<td>LDA</td>
<td>Local density approximation</td>
</tr>
<tr>
<td>GGA</td>
<td>Generalized gradient approximation</td>
</tr>
<tr>
<td>EA</td>
<td>Ethanolamine</td>
</tr>
<tr>
<td>QY</td>
<td>Quantum yield</td>
</tr>
<tr>
<td>AdePeCbl</td>
<td>Adeninylpentylcobalamin</td>
</tr>
<tr>
<td>PDB</td>
<td>Protein data bank</td>
</tr>
<tr>
<td>MLCT</td>
<td>Metal-to-ligand charge transfer</td>
</tr>
<tr>
<td>LF</td>
<td>Ligand field</td>
</tr>
<tr>
<td>MECP</td>
<td>Minimum energy crossing point</td>
</tr>
<tr>
<td>alkylCbls</td>
<td>AdoCbl and CH₃Cbl</td>
</tr>
<tr>
<td>non-alkylCbls</td>
<td>H₂OCbl OHCbl CNCbl</td>
</tr>
<tr>
<td>HOMO</td>
<td>Highest occupied molecular orbital</td>
</tr>
</tbody>
</table>
CURRICULUM VITAE

Abdullah Al Mamun
a0mamu01@louisville.edu/mamunche.du@gmail.com

ACADEMIC BACKGROUND

Ph.D. Candidate in Chemistry (Spring 2020)
Department of Chemistry, University of Louisville, KY

M.S. Physical Chemistry
Department of Chemistry, University of Louisville, KY
CGPA (3.89/4.00) 2018

M.Sc. Inorganic Chemistry
University of Dhaka, Bangladesh
CGPA (3.39/4.00) 2014

B.Sc. Chemistry
University of Dhaka, Bangladesh
CGPA (3.20/4.00) 2013

RESEARCH HIGHLIGHTS

Spring 2015-Spring 2020, Advisor- Pawel M. Kozlowski, University of Louisville, Louisville, KY, USA

❖ Photochemistry of Metalloenzymes

▪ Photochemical studies of AdoCbl-dependent ethanolamine ammonia lyase (EAL) and glutamate mutase (GLM) enzymes using QM/MM method.
▪ Photodissociation mechanism of AdoCbl in the biological photoreceptor CarH enzyme using QM/MM method.
▪ Photochemical studies of MeCbl-dependent methionine synthase (MetH) and the effect of mutation in the cap domain using QM/MM and MD simulation method.

❖ Enzymatic Catalysis in Vitamin B12-Dependent Enzymes

▪ Connection between photolysis and native catalysis in the activation of Co-C bond for AdoCbl-dependent enzymes using QM/MM method.
- Modelling the enzymatic reactions catalyzed by AdoCbl-dependent EAL and AdoCbl-dependent GLM using QM/MM method and MD simulation method.

- **Chemistry of Vitamin B₁₂ Derivatives**
  - Resonance Raman study of cobalamin cofactors using DFT.
  - Electronic structure of hydridocobalamin (HCbl) using DFT and CASSCF.

- **Hydrogen evolution reactions**
  - Metal centered hydrogen evolution reaction (HER) for Ni catalyst using DFT.
  - Ligand centered hydrogen evolution reaction (HER) for Zn catalyst using DFT.

**Aug 2013- Dec 2014, Advisor-Dr. Abdus Salam, University of Dhaka, Bangladesh.**

- **Identifications and Characterizations of Toxic Materials in the Air**
  - Identification and characterization of trace metals in the air of Dhaka city.
  - Identification and source characterization of particulate materials (PM2.5 and PM10) and carbon monoxide (CO) in the air of Dhaka city.

**TECHNICAL SKILLS**

- **Technical Expertise:** Quantum chemical calculations, Density Functional Theory (DFT), Time-dependent Density Functional Theory (TD-DFT), Multireference Self-Consistent Field Calculations (CASSCF-MCQDPT2), Couple cluster singles and doubles (CCSD) method, Multiscale protein/ligand modelling (QM/MM), Molecular Dynamics (MD) simulation, Parallel/HPC computing.
- **Quantum Chemistry codes:** Gaussian, NwChem, Orca, VASP, Gamess-US
- **Molecular Dynamics:** Amber, Gromacs
- **Visualization software:** Chimera, Gaussview, VMD, Chemcraft, Avagadro, Material Studio
- **Plotting software and data analysis:** Origin, Gnuplot, Matlab
- **Programming languages:** Basic knowledge of Python and linux shell scripting
- **Operating systems:** Linux, Windows and Mac
- **Experimental techniques:** GC, GC-MS

**PUBLICATIONS**


10. Higa, Kaoru M.; Mamun A. A.; Ogura, T; Kitagawa, T; and Kozlowski, P. M.; Resonance Raman investigation of dithionite reduced cobalamin. *Under review, Journal of Raman Spectroscopy*. 2019. (First two authors have equal contributions.)

**Manuscript in preparation:**

1. Toda, M. J.; Mamun, A. A.; Lodowski, P.; Kozlowski, P. M., Why is the CarH photoreceptor photolytically active? *Manuscript in preparation*


POSTER AND ORAL PRESENTATIONS

1. **Abdullah Al Mamun**, Pawel M. Kozlowski, “Photochemistry of AdoCbl-dependent ethanolamine ammonia-lyase” From Computational Biophysics to Systems Biology (CBSB), University of Cincinnati, May 2017 (Poster presentation)


5. **Abdullah Al Mamun**, Pawel M. Kozlowski, “Can photolysis of the Co-C bond be used to mimic the native enzymatic catalysis of the coenzyme B12?” Graduate Student Regional Research Conference (GSRRRC), University of Louisville, February 27-28,2019, (Oral Talk)

6. **Abdullah Al Mamun**, Pawel M. Kozlowski, “Can photolysis of the Co-C bond in coenzyme B12-dependent enzymes be used to mimic the native enzymatic reactions”. Biophysical Society Annual Meeting, Baltimore, Maryland, March 2-6, 2019 (Poster presentation)

7. **Abdullah Al Mamun**, Pawel M. Kozlowski, “Connection between photolysis and native catalysis in AdoCbl-dependent enzymes” 51st Midwest Theoretical Chemistry Conference, June 6-8, 2019 (Poster presentation)


10. Arghya Pratim Ghosh, **Abdullah Al Mamun**, Pawel M. Kozlowski, “Light-Induced Activation of Organo-Metallic Co-C Bond in MeCbl-Dependent Methionine Synthase-QM/MM Study” Biophysical Society Annual Meeting, Baltimore, Maryland, March 2-6, 2019 (Poster presentation)

11. **Abdullah Al Mamun**, Computational Modeling of Vitamin B12-Dependent Enzymatic Reactions: Mechanistic Insights, invited talk, Vanderbilt University, September, 21,2019

AWARDS

- Doctoral Dissertation Completion Award, University of Louisville, Spring 2020
- Graduate Teaching Assistant, University of Louisville, Spring 2015-Fall 2019
- Graduate Student Council Travel Award, ($350) Spring -2019
- International Center Tuition Award, University of Louisville ($200) Fall-2018
▪ International Center Tuition Award, University of Louisville ($100) Spring-2018
▪ GNAS Travel Award, University of Louisville ($250) 2018
▪ Graduate Student Council Travel Award, ($350) 2018
▪ Graduate Student Council Travel Award, ($ 350) 2017
▪ Department of Chemistry Travel Award, ($200) 2017
▪ Bangabandhu Fellowship, Bangladesh (60000 Tk) 2014

TEACHING EXPERIENCES

Graduate Teaching assistant Spring 2015 – Fall 2020
Department of Chemistry, University of Louisville, KY, USA

▪ Intro Chemistry Laboratory (Chem 103)
▪ Chemical Analysis Laboratory I (Chem 207)
▪ Chemical Analysis Laboratory II (Chem 208)
▪ Chemical analysis Laboratory III (Chem 209)

PROFESSIONAL MEMBERSHIP

▪ American Chemical Society 2017-present
▪ Biophysical Society 2018-present