Computational modelling of chemical reactions catalyzed by cobalamin (B12)-dependent enzymes: mechanistic insights.

Neeraj Kumar
University of Louisville

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

Recommended Citation
https://doi.org/10.18297/etd/782

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 CHEMICAL REACTIONS CATALYZED BY COBALAMIN (B12)-DEPENDENT ENZYMES: MECHANISTIC INSIGHTS

By

Neeraj Kumar
B.S., Panjab University, Chandigarh, India, 2005
M.S., Panjab University, Chandigarh, India, 2007

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

Department of Chemistry
University of Louisville
Louisville, Kentucky

August, 2013
COMPUTATIONAL MODELLING OF CHEMICAL REACTIONS CATALYZED BY COBALAMIN (B12)-DEPENDENT ENZYMES: MECHANISTIC INSIGHTS

By

Neeraj Kumar
B.S., Panjab University, Chandigarh, India, 2005
M.S., Panjab University, Chandigarh, India, 2007

A Dissertation Approved On:
25th July, 2013

by the following Dissertation Committee:

Dissertation Director: Dr. Pawel M. Kozlowski

Dr. Craig A. Grapperhaus

Dr. Frederick A. Luzzio

Dr. Chakram. S. Jayanthi

Dr. Jinjun Liu
DEDICATION

This dissertation is dedicated to my respected father ‘Sh. Subash Chand’, late mother ‘Smt. Usha Rani’ and my grandparents.
ACKNOWLEDGEMENTS

I am very grateful to my mentor, Dr. Pawel M. Kozlowski for his invaluable guidance, support, and encouragement that he has given me throughout my graduate life. I would cherish our healthy discussions on science as well as on life forever. Also, I am truly thankful for the financial assistanship he has provided me during the course of the graduate research.

I want to express my deep appreciation to my committee members at University of Louisville: Dr. Craig. A. Grapperhaus, Dr. Frederick A. Luzzio, Dr. Jinjun Liu, and Dr. Chakaram P. Jayanthi, for their assistance with my progress towards my doctorate. In addition, I express my gratitude to Dr. George Pack for the invaluable guidance towards my research and he always had encouraging words for me; I hope he recovers soon. I am also extremely grateful to our experimental collaborators, Prof. James R. Kincaid from Marquette University. Special thanks is also extended to my summer intern mentor Dr. Felice Lightstone, the directorate of Bioscience division at Lawrence Livermore National Laboratory (LLNL) for giving me the opportunity to work in a collaborative rich environment and Dr. Eric Schwegler at LLNL for giving me the guidance regarding the postdoctoral research in National Laboratories. I would also like to thank my Master's adviser Dr. Neetu Goel at Panjab University, Chandigarh, India and Dr. Harjinder Singh at IIT, Hyderabad, India for encouraging me to pursue higher studies in the field of theoretical chemistry.

I would also like to thank the Research Computing Governance Council for providing us the high priority access to Cardinal Research Cluster (CRC) during the course of my graduate studies. I would also like thank to James Simrall for his great help related to the technical issues and solving the problems regarding the CRC. Last but not the least, I would like to thank awesome office staff in Chemistry Department; Sherry Nalley, Syble Bullock, Aaron and Sabrina for being so supportive and helpful.
I would like to thank my colleagues and lab mates Dr. Jadwiga Kuta, Dr. Manoj Kumar, Dr. Karina Kornobis and Hui Lui for their friendship and support. Also, I would like to thank my peer fellows and buddies Davinder Kumar, Manish Kumar, Rajat, Shrinivas, Radhika, Sadakat Ali Gori, and Rahul Jain for their friendship and unforgettable memories we shared together.

I would like to express my special appreciation and thanks to my father and late mother for giving exceptional educational opportunities and bringing the best out of me. I would sincerely like to thank my family members including, sisters, brother, uncle and aunts for the encouragement and support throughout my educational path. This journey would not have been possible without their keen guidance and prayers. I would also like to thank my supportive wife Deepika for being there with me through all the thick and thin.
The biologically active derivatives of vitamin B$_{12}$ (cobalamin) possess the only examples containing organometallic Co-C bonds in living systems where the cleavage of Co-C bond initiates the catalytic cycle of the enzymatic reactions. In this dissertation, the electronic structure of various intermediates and the mechanistic details of the methyl transfer reactions involved in the catalytic cycle of the methionine synthase (MetH) have been computationally investigated using hybrid quantum mechanics/molecular mechanics (QM/MM) simulations, density functional theory (DFT), and complete active space self-consistent field with second order perturbation theory (CASSCF/QDPT2) computational methodologies. QM/MM calculations in particular reveal that the cob(I)alamin intermediate is not axially coordinated inside the enzyme which is consistent with free cob(I)alamin in solution, thus indicating an unprecedented role of enzyme-bound cob(I)alamin. In addition, DFT and high-level ab initio CASSCF/QDPT2 calculations further elucidate that the ground state of the cob(I)alamin is multiconfigurational where the diradical Co(II)-corrin$^+$ radical configuration contributes to the electronic structure of the cob(I)alamin intermediate, indicating a noninnocent behavior of the corrin ring. Furthermore, our QM/MM study reveals the traditionally assumed S$_{N2}$ mechanism for the formation CH$_3$-cob(III)alamin and the activation energy barrier for S$_{N2}$ reaction is
found to be comparable with respect to the experimental rate constant. However, the possibility of an alternative ET-based radical mechanism consistent with the close-lying diradical states has also been suggested, where an electron transfer (ET) from His-on cob(I)alamin to pterin ring of the protonated CH$_3$-H$_4$Folate takes place, indicating Co$^{II}(d^7)$-pterin radical ($\pi^*$)$^I$ state, followed by a methyl radical transfer. The similar mechanistic details of $S_N2$ and ET-based radical pathway have also been investigated in the second-half catalytic cycle of MetH as indicated by ionic and diradical states of the MeCbl:Homocysteine reaction complex. The mechanistic details of the MeH could also help in understanding some of the critical aspects related to the enzymatic methyl transfer reactions in other methyltransferases.

In addition, the role of tyrosine active site in AdoCbl-dependent enzyme methylmalonyl CoA mutase (MCM) has been studied using density functional reactivity theory in which the charge separation propensity of electron transfer site is often distant from the proton-acceptor site, indicating that one electron reduced form AdoCb may be involved in the initial step of AdoCbl-dependent catalysis. As a result, the QM/MM optimized intermediates and transition states along with their calculated energetic profiles indicate that the reaction consisting of Co-C bond cleavage and subsequent hydrogen abstraction occurs in a concerted fashion when the reduced form of AdoCbl cofactor is involved in the reaction, whereas it takes place in a stepwise manner in a neutral form. The concerted pathway was further supported by the free energy metadynamics simulations.
TABLE OF CONTENT

ACKNOWLEDGEMENTS..............................................................................................................vi
ABSTRACT...................................................................................................................................vi
LIST OF TABLES..........................................................................................................................xiii
LIST OF FIGURES........................................................................................................................xv
LIST OF SCHEMES....................................................................................................................xxii

CHAPTER Page
1 INTRODUCTION 1
1.1 Computational Methodology...............................................................................................9
1.1.1 QM/MM Calculations........................................................................................................10
1.1.2 DFT Calculations................................................................................................................11
1.1.3 CASSCF Calculations.........................................................................................................13
1.2 Mechanistic Details of Methyl Transfer Reactions in Methionine Synthase (MetH) ..............14
1.3 Catalysis in Adenosylcobalamin-Dependent Mutases: Methylmalonyl Coenzyme A Mutase (MCM)..................................................................................................................21

2 ROLE OF THE AXIAL BASE IN THE MODULATION OF THE COB(I)ALAMIN ELECTRONIC PROPERTIES 28
2.1 Introduction..........................................................................................................................28
2.2 Computational Details..........................................................................................................33
2.2.1 QM(DFT)/MM Calculations............................................................................................34
2.2.2 Gas Phase DFT Calculations...........................................................................................35
2.2.3 CASSCF/MC-XQDPT2 Calculations................................................................................36
2.3 Result and Discussion..........................................................................................................36
2.3.1 Structure of Cob(I)alamin Inside MetH ........................................36
2.3.2 Structural and Electronic Analysis Based on DFT Calculations .................................................................40
   2.3.2.1 His-off Cob(I)alamin Intermediate ........................................40
   2.3.2.2 Influence of the Axial Base His .............................................43
2.3.3 CASSCF/MC-XQDPT2 Analysis .........................................................46
2.3.4 Energy Changes as Function of Co–N̸_Im Distance .................51
2.3.5 Implications for the Remethylation of the Cob-(I)alamin Intermediate in MetH ......................................................52
2.4 Conclusion .................................................................................53

3 MECHANISTIC INSIGHTS FOR THE METHYLATION OF THE COB(I)ALAMIN FROM CH₃-H₄FOLATE CATALYZED BY METHIONINE SYNTHASE 56
3.1 Introduction ..................................................................................56
3.2 Computational Methods .................................................................62
   3.2.1 System Preparation and ONIOM-based QM/MM Calculations 62
   3.2.2 Enzyme-Substrate Reaction Complex ......................................65
3.3 Results and Discussion ................................................................66
   3.3.1 Chemical Models and Enzyme-Substrate Reaction Complex ......66
   3.3.2 SN2-type Nucleophilic Displacement (His-off) .......................77
   3.3.3 ET-based Methyl Radical Transfer Mechanism (His-on) ..........82
   3.3.4 Understanding the protonation process of the CH₃-H₄Folate Substrate .........................................................88
   3.3.5 Implications for the Remethylation of the cob(I)alamin/ cob(I)amide intermediate to other Methyltransferases ..............88
4 ELECTRONIC STRUCTURE OF COFACTOR-SUBSTRATE REACTANT COMPLEX OF METHIONINE SYNTHASE

4.1 Introduction........................................................................................................................................96
4.2 Computational Methods and Model System..................................................................................101
  4.2.1 Structural Models.......................................................................................................................101
  4.2.2 CASSCF/MR-PT Calculations....................................................................................................103
4.3 Results and Discussion....................................................................................................................104
  4.3.1 Electronic Properties of Reactant Complex Based on DFT Calculations..................................104
  4.3.2 Selection of the Active Space.....................................................................................................105
  4.3.3 Analysis of Ionic and Diradical States.......................................................................................111
  4.3.4 ET versus S_N2-Type Mechanism..............................................................................................118
4.4 Conclusion and Summary................................................................................................................119

5 ELECTRONIC STRUCTURE OF ONE-ELECTRON OXIDIZED FORM OF THE METHYLCOBALAMIN COFACTOR

5.1 Introduction........................................................................................................................................121
5.2 Computational Details.......................................................................................................................126
  5.2.1 Structural Models.......................................................................................................................126
  5.2.2 DFT Calculations........................................................................................................................126
  5.2.3 CASSCF/MC-XQDPT2 Calculations..........................................................................................127
5.3 Results and Discussion.....................................................................................................................127
  5.3.1 Geometry of the 1e-Ox Form of MeCbl and Related Porphine Complexes..............................127
  5.3.2 Electronic Structure of the 1e-Ox MeCbl..................................................................................130
5.3.3 Spin Density Profiles in the 1e-Ox Models of MeCbl and Related Porphine Complexes ......................................................... 130
5.3.4 Dissociation of the Co–C Bond in the 1e-Ox MeCbl ............ 136
5.3.5 Spin Density Changes along the Co–C Bond Stretch ............. 140
5.3.6 CASSCF/MC-XQDPT2 Analysis of the 1e-Ox MeCbl .......... 141
5.4 Conclusion ........................................................................... 147

6 CHARGE SEPARATION PROPENSITY OF THE COENZYME B12–TYROSINE COMPLEX IN ADENOSYLCOBALAMIN-DEPENDENT METHYLMALONYL–COA MUTASE ENZYME 150
6.1 Introduction ........................................................................... 150
6.2 Computational Details ................................................................. 153
6.2.1 Gas Phase DFT Calculations and Prepared Models .......... 153
6.2.2 ONIOM-Based QM/MM Calculations ................................. 155
6.3 Results and Discussion ............................................................... 155
6.3 Conclusion ............................................................................. 162

7 REDUCTIVE Co-C BOND CLEAVAGE MECHANISM IN B12− DEPENDENT METHYLMALONYL CoA MUTASE: ENERGETIC IMPLICATIONS 163
7.1 Introduction ........................................................................... 163
7.2 Computational Methods and Model Systems .......................... 170
7.2.1 Initial Set up for the System and Molecular Mechanics Calculations ................................................................. 171
7.2.2 QM(DFT)/MM calculations .................................................. 171
7.2.3 QM/MM CPMD Simulations ............................................... 172
7.3 Results and Discussion ............................................................... 173
7.3.1 Mechanistic Details of PCET for the Initial step of the Co-C bond Cleavage ..................................................................................173

7.3.2 Reductive Co-C5’ Bond Cleavage and Subsequent H-atom abstraction from the Substrate in one Electron Reduced Form ..................................................................................................176

7.3.3 Co-C5’ Bond Cleavage and Subsequent H-atom abstraction from the Substrate in Neutral AdoCbl ............................................................................184

7.3.4 Free energy surface of the Reductive Co-C5’ Bond Cleavage and Subsequent hydrogen abstraction step in one Electron Reduced AdoCbl ..................................................................................190

7.4 Conclusion ..................................................................................................................193

8 CONCLUSIONS .................................................................................................................196

REFERENCES ....................................................................................................................203

APPENDIX .........................................................................................................................221

LIST OF ABBREVIATIONS .................................................................................................231

CURRICULUM VITAE .........................................................................................................234
### LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Key Structural Parameters of the Cofactor Binding Domain of MetH in the Me-Cob(III)alamin Resting State (MeCbl:MetH) and the Cob(I)alamin Intermediate (Co(I):MetH)</td>
</tr>
<tr>
<td>2.2</td>
<td>Main Structural Parameters of the Full and Truncated Cob(I)alamin Gas Phase Models</td>
</tr>
<tr>
<td>2.3</td>
<td>Composition of the CASSCF Wave Function of Cob(I)alamin for each Co–NIm Distance, in Terms of the Mulliken Occupation Numbers of the Three Natural Orbitals Involved in the Open-Shell Singlet Description and the Weights of the Major Configurations State Functions (CSF1, CSF2, CSF3, and CSF4)</td>
</tr>
<tr>
<td>3.1</td>
<td>Key Structural Parameters of the Optimized Systems of MeCbl-MetH Binding Domain and protonated MetH-cobalamin:CH$_3$-H$_4$Folate Reaction Complex (RCN5H) in Comparison with the Experimental Data</td>
</tr>
<tr>
<td>3.2</td>
<td>Relative Energies (in kcal/mol), Mulliken Charges and Spin Densities of Cobalamin:CH$_3$-H$_4$Folate Complexes Calculated with ONIOM-based QM/MM calculations</td>
</tr>
<tr>
<td>4.1</td>
<td>Weight Percentage of Three Major Configurations (CSFs) in the $\Psi_{\text{dirad}}^L$, $\Psi_{\text{dirad}}^{\text{II}}$, and $\Psi_{\text{ionic}}$ Wave Functions for SM</td>
</tr>
<tr>
<td>4.2</td>
<td>Weight Percentage of Three Major Configurations (CSFs) in the $\Psi_{\text{dirad}}^L$, $\Psi_{\text{dirad}}^{\text{II}}$, and $\Psi_{\text{ionic}}$ Wave Functions for LM</td>
</tr>
</tbody>
</table>
5.1 Relevant Co–C BDEs (kcal/mol) and Axial Bond Lengths (Å) of the Neutral, 1e-Ox, and 1e-Red MeCbl Models Computed at the BP86/631G* and B3LYP/6-31G* Levels of Theory..................................................123

5.2 Relevant Geometrical Parameters of the 1e-Ox B-[Co^{III}(Por)]-Me (B = Im or Pyr) Optimized at the BP86/6-31G* and B3LYP/6-31G* Levels of Theory.................................................................124

5.3 Composition of SOMO Molecular Orbital (in %) of the 1e-Ox MeCbl Determined at the BP86/6-31G* and B3LYP/6-31G* Levels of Theory..................................................................................127

5.4 Composition of CASSCF/MC-XQDPT2 Wave Function of 1e-Ox MeCbl in terms of Major Configurations State Functions (CSFs) Computed for First Three Excited States...........................................139

A1. Relative Energy and Composition of the CASSCF/MC-XQDPT2 Wave Function of the Ground State Singlet and three-low Lying Excited States for each Co-N\textsubscript{Im} Distances. The Wave Function is given in Terms of the corresponding Three Major Configurations State Functions (CSFa, CSFb and CSFc) of each State.................................................................216

A2. Weight of CASSCF/MS-CASPT2 calculations in $\Psi_{\text{dirad}}^{\|}$, $\Psi_{\text{dirad}}^{-}$ Wave Functions for Small Model Computed at Different C-S Distance........217

A3. Relevant energetical and geometrical parameters of the neutral and 1e-Ox MeCbl optimized at the BP86/6-31G* level of theory..................218

A4. Relevant energetical and geometrical parameters of the neutral and 1e-Ox MeCbl optimized at the B3LYP/6-31G* level of theory...............218
**LIST OF FIGURES**

<table>
<thead>
<tr>
<th>FIGURES</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Molecular structure of vitamin B₁₂ (CNCbl = cyanocobalamin, R = -CN), methylcobalamin (R = -Me) and adenosylcobalamin (R = Ado).</td>
</tr>
<tr>
<td>1.2</td>
<td>(A) The reaction scheme catalyzed by methionine synthase (MetH), (B) Structure of methylcobalamin and cob(I)alamin cofactor, and (C) Structure of different substrates used in MetH.</td>
</tr>
<tr>
<td>2.1</td>
<td>(a) Molecular structure of free methylcobalamin, where R = CH₃, R₁ = CH₂CONH₂, R₂ = CH₂CH₂CONH₂, and R₃ = (CH₂)₂CONHCH₂CH(CH₃)OPO₃. (b) Gas phase model of the Im⁻[Cob(I)alamin] enzymatic intermediate in the present chapter.</td>
</tr>
<tr>
<td>2.2</td>
<td>Close view of active site of MeCbl in the enzyme showing the interaction of lower axial ligand His759 with the other residues of the triad (Asp759, Ser810) as well as the nearby Leu806 based on the 1BMT crystal structure.</td>
</tr>
<tr>
<td>2.3</td>
<td>QM(DFT)/MM optimized structures of MeCbl binding domain of the MetH (a) hexa-coordinated Me-Cob(III)alamin resting state and (b) Cob(I)alamin intermediate.</td>
</tr>
<tr>
<td>2.4</td>
<td>Spin density of (a) the full cob(I)alamin model and (b) the truncated cob(I)alamin model with C₂ symmetry, computed at the B3LYP/6-31G(d) 5d level of theory. Left, spin populations; right, R and β spin density distributions colored as green and magenta, respectively.</td>
</tr>
</tbody>
</table>
2.5 Evolution of the expectation value of the total spin, $<S^2>$, along with the spin density distributions during the elongation of the Co–NIm distance in the Im$^-$[Cob(I)alamin] model system.

2.6 Variation of HOMO and HOMO–1 based orbital energies with the Co–NIm distance calculated at the UB3LYP level of theory.

2.7 CASSCF active space orbitals used in the calculations of the Im$^-$[Cob(I)alamin] model system at a Co–NIm distance = 2.2 Å.

2.8 Weight (%) of the major configurations (CSFs) contributing to the ground state CASSCF wave function as a function of the Co–NIm distance for the Im$^-$[Cob(I)alamin] model system. The description of the CSFs is given in Table 2.3.

2.9 Pure fragment localized orbitals, describing the electron transfer between the cobalt and the corrin. a) Co–NIm distance < 2.5 Å; the electron is transferred from Co(d$^2$) to corrin($\pi^*$). b) Co–NIm distance > 2.5Å; the electron is shifted from Co(d$_{yz}$) to corrin($\pi^*$).

2.10 Change in the ground state energy with the Co–NIm distance, computed with different DFT functionals and ab initio CASSCF/MCXQDPT2 methods.

3.1 The catalytic cycle and reactivation cycle of methionine synthase (MetH): [I] the first-half catalytic reaction, [II] the second-half catalytic reaction, and [III] the reactivation reaction.

3.2 The active site of CH$_3$-H$_4$Folate substrate (PDB ID: 1Q8J at 1.9 Å resolution) showing the polar environment near the pterin ring of the CH$_3$-H$_4$Folate. Here, the pterin ring is positioned by hydrogen bonds with four important protein residues: Asn508, Asp473, Asp390, and Asn411 and out of these, the N and O of the Asn508 are directly interacting via H-bond with the N5 of pterin ring.

3.3 The QM/MM-ME(css) optimized geometries of the enzyme-substrate complex in the N5H protonation state: (a) reactant complex (RC$_{N5H}$) in
His-off conformation, (b) transition state ($TS_{N5H}$), and (c) product complex ($PC_{N5H}$) in closed shell singlet (css) state calculations....................70

3.4 The QM/MM-ME optimized geometry of the N5H protonated cob(I)alamin:CH$_3$-H$_4$Folate reactant complex (RC$_{N5H}$) in a) triplet state, b) open shell singlet state calculations which shows His-on conformation of Co, consistent with Co$^{II}$(d$^7$)-pterin radical ($\pi^*$) diradical configuration.................................................................73

3.5 Potential energy profile (kcal/mol) for the Co-C$_{Me}$ bond formation in the N5H protonated state, based on the QM/MM-ME calculations and the QM/MM-EE energies are given in the bracket. The corresponding optimized geometries are shown in Figures 3.3 and 3.4.........................75

3.6 Potential energy profiles with energy barrier for the formation of Co-C$_{Me}$ bond in N5H protonated reaction complex of MetH-cob(I)alamin:CH$_3$-H$_4$Folate shown in Figure 3.4, calculated for the triplet (T) as well as open shell singlet (oss) QM/MM calculations. The energies are given relative to the RC$_{N5H}$ where the Co-C$_{Me}$ is 4.16 Å shown in Figure 3.4. The second embedded plot is the evolution of spin density distribution along the Co-C$_{Me}$ distance extracted from open shell singlet calculations.............85

4.1 Close up of the active sites of MetH: (a) 3BOF-based crystal structure of Hcy substrate binding domain and (b) 1BMT-based crystal structure of MeCbl domain.................................................................98

4.2 (a) Molecular structure of methylcobalamin (MeCbl) and its structural simplifications referred to as (b) large (LM) and (c) small (SM) models respectively in this chapter.........................................................102

4.3 Singly occupied sulfur and corrin natural orbitals representing two different symmetry solutions based on RASSCF calculations for small model (SM).................................................................107

4.4 Orbitals in the active spaces for the $\Psi_{\perp \text{dirad}}$, $\Psi_{\| \text{dirad}}$, and $\Psi_{\text{ionic wave}}$ functions used for SM plotted at a C···S distance of 4.0 Å.............113
4.5 Energy curves for the $\Psi_{\text{dirad}}$, $\Psi_{\text{dirad}}^\perp$, and $\Psi_{\text{ionic}}$ wave functions for SM computed at CASSCF/CASPT2 level of theory……………………………..113

4.6 Orbitals in the active space used for CASSCF/MC-XQDPT2 calculations for LM at a C···S distance of 4.0 Å…………………………………….114

4.7 Energy curves for the $\Psi_{\text{dirad}}$, $\Psi_{\text{dirad}}^\perp$, and $\Psi_{\text{ionic}}$ wave functions for LM computed at CASSCF/MC-XQDPT2 level of theory…………………..116

5.1 Molecular structure of B$_{12}$ cofactors (left panel) where R = Me for MeCbl and R = Ado for AdoCbl ($R_1 = \text{CH}_2\text{CONH}_2$, $R_2 = \text{CH}_2\text{CH}_2\text{CONH}_2$, $R_3 = (\text{CH}_2)_2\text{CONHCH}_2\text{CH(\text{CH}_3)OPO}_3^-)$). The right panel shows the structural model of MeCbl employed in this chapter and denoted as Im-[Co$^{\text{III}}$(corrin)]-Me$^+$.……………………………………………………...122

5.2 Orbital energy level diagram for 1e-Ox MeCbl: (a) BP86/6-31G* level of theory and (b) B3LYP/6-31G* level of theory. SOMO is indicated in green; the $\sigma^*$ energy level is indicated in blue…………………………131

5.3 Spin density for the 1e-Ox MeCbl: (a) BP86/6-31G* level of theory and (b) B3LYP/6-31G* level of theory……………………………………..134

5.4 Spin density for the 1e-Ox Im-[Co$^{\text{III}}$(Por)]$^-$-Me$^+$: (a) BP86/6-31G* level of theory and (b) B3LYP/6-31G* level of theory…………………………135

5.5 Spin density for the 1e-Ox Pyr-[Co$^{\text{III}}$(Por)]$^-$-Me$^+$: (a) BP86/6-31G* level of theory and (b) B3LYP/6-31G* level of theory…………………………135

5.6 Comparison of dissociation curves for neutral and 1e-Ox forms of MeCbl calculated using the BP86/6-31G* level of theory…………………………137

5.7 BP86/6-31G* calculated: (a) the ground state and low-lying electronic excited states for 1e-Ox MeCbl as a function of Co–C bond length and (b) spin density changes along the Co–C bond length…………………………138

5.8 B3LYP/6-31G* calculated: (a) the ground state and low-lying electronic excited states for 1e-Ox MeCbl as a function of Co–C bond length and (b) spin density changes along the Co–C bond length…………………………139
5.9 Active space orbitals of 1e-Ox MeCbl used in CASSCF (11,12) calculations

6.1 Active site close-up of the X-ray crystal structure of substrate-bound AdoCbl (PDB-code: 4REQ @ 2.2 Å resolution). To enhance the visualization picture of the active site, the deoxyadenosyl (Ado) moiety has not been shown.

6.2 Dual descriptor contour surface of the minimal structural model of MCM enzyme

6.3 (a) Active site view of the QM/MM-optimized structure of substrate-bound MCM enzyme (the Ado moiety has not been shown here). (b) Corresponding dual descriptor contour surface

7.1 Molecular structure of coenzyme B$_{12}$ where $R = \text{Ado} (R_1 = \text{CH}_2\text{CONH}_2, \ R_2 = \text{CH}_2\text{CH}_2\text{CONH}_2, \ R_3 = (\text{CH}_2)_2\text{CONHC}_6\text{H}_4(\text{CH}_3)\text{OPO}_3^-)$ (right). Structural models of coenzyme B$_{12}$ where the side chains are replaced by hydrogen atoms (left)

7.2 The QM/MM optimized active site of MCM enzyme comprising 1e-reduced AdoCbl cofactor with C2'-endo conformation of the ribose ring: a) reactant complex (RC), b) transition state (TS) between Co-C bond cleavage and subsequent H abstraction from the substrate (Co-C5'-H), C) substrate radical (SR)

7.3 The QM/MM optimized active site of MCM enzyme comprising 1e-reduced AdoCbl cofactor with C3'-endo conformation of the ribose ring: a) reactant complex (RC), b) transition state (TS) between Co-C bond cleavage and subsequent H abstraction from the substrate (Co--C5'--H), c) substrate radical (SR)

7.4 Potential energy surface of 1e-reduced cofactor of the MCM enzyme calculated using QM/MM method where corresponding active sites of RC, TS$_{\text{Co-C5'-H}}$, and SR are shown in Figure 7.2 and 7.3
7.5 The QM/MM optimized active site of MCM enzyme comprising neutral AdoCbl cofactor with C2'-endo conformation of the ribose ring: a) reactant complex (RC), b) transition state (TS) of the H abstraction from the substrate (C5′--H), C) substrate radical (SR)……………………………………..185

7.6 The QM/MM optimized active site of MCM enzyme comprising neutral AdoCbl cofactor with C3'-endo conformation of the ribose ring: a) reactant complex (RC), b) transition state (TS) of the H abstraction from the substrate (C5′--H), C) substrate radical (SR)……………………………………..186

7.7 Potential energy surface of the neutral AdoCbl cofactor of MCM enzyme calculated using hybrid QM/MM method where corresponding active sites of the RC, TS_{C5'-H}, and SR are shown in Figure 7.5, Figure 7.6 and Figure 7.8……………………………………………………………………….188

7.8 The QM/MM optimized active site of MCM enzyme comprising neutral AdoCbl cofactor with C3'-exo conformation of the ribose ring of the reactant complex (RC)……………………………………………………………..188

7.9 2-D Free energy surface (KJ/mol) for reductive cleavage of the Co-c bond and subsequent hydrogen abstraction step (epsilon) computed with the metadynamics………………………………………………………………………..191

A1. Evolution of the spin density of the truncated Im″[Cob(I)alamin] model system with the Co-N_{Im} distance (Å). Left: spin populations; Right: alpha and beta spin density distributions, coloured in green and magenta, respectively………………………………………………………………………..224

A2. CASSCF Active Space Orbitals used in the calculations of the free cob(I)alamin model system………………………………………………………………………..225

A3. Five major configurations (CSFs) contributing to the CASSCF ground state singlet of the Im″[Cob(I)alamin] model system before localization and their weight (%) as a function of the Co-N_{Im} distance. (Note that the notation of the CSFs described here are different from the rest of the discussion)………………………………………………………………………..225
A4. The quantum mechanics (QM) region of the MetH-cob(I)alamin:CH$_3$-H$_4$Folate(THF) reaction complex that was used in the ONIOM-based quantum mechanics /molecular mechanics (QM/MM) calculations is shown in the sticks framework and rest of the protein is treated as the part of the MM region…………………………………………………………..226

A5. The ONIOM-EE based QM(BP86/6-31G(d)/MM optimized structure of the methionine synthase (MetH)-bound methylcobalamin (MeCbl) cofactor……………………………………………………………………227

A6. The ONIOM-ME based QM(BP86/6-31G(d)/MM optimized structure of the methionine synthase (MetH)-bound cob(I)alamin cofactor which were generated by removing the methyl group of MeCbl and adjusting the number of electrons consistent with Co$^{+1}$ oxidation state. ……………227

A7. The ONIOM-based QM(BP86/6-31G(d)/MM-ME optimized structure of the unprotonated (N5) MetH-cob(I)alamin:CH$_3$-H$_4$Folate reaction complex of: (a) reactant complex (RC$_{N5}$), (b) transition state (TS$_{N5}$) in closed shell singlet calculations……………………………………………………228

A8. Larger Model of B12-dependent coenzyme containing full AdoCbl,Y89 and truncated substrate that shows dual descriptor contour surface for the reactant complex……………………………………………………………229

A9. Active site view of the ONIOM-based QM/MM optimized structure of substrate-bound MCM enzyme (BP86 Functional). For viewpoint purpose Ado group has not been shown…………………………………………………229
# LIST OF SCHEMES

<table>
<thead>
<tr>
<th>SCHEMES</th>
<th>PAGES</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>The first-half (1) and second-half (2) catalytic reaction of the MetH……………16</td>
</tr>
<tr>
<td>1.2</td>
<td>General catalytic mechanism for AdoCbl-dependent enzymes.........................23</td>
</tr>
<tr>
<td>1.3</td>
<td>Reactions catalyzed by Methylmalonyl CoA mutase (MCM).................................24</td>
</tr>
<tr>
<td>2.1</td>
<td>The Catalytic Cycle (Shown in Blue) and the Reactivation Cycle (Shown in Red) for the Cobalamin-Dependent Methionine Synthase (MetH).................................29</td>
</tr>
<tr>
<td>3.1</td>
<td>(I) ( S_{N2} )-type, and (II) ET-based radical mechanistic pathways for the methyl transfer reaction from a ( \text{CH}_3\text{-H}_4\text{Folate} ) substrate to a cob(I)alamin intermediate catalyzed by methionine synthase (MetH)................................................76</td>
</tr>
<tr>
<td>3.2</td>
<td>The schematic representation of the ET-mediated radical mechanism where the N5-C\text{Me} bond is elongated and Co-C\text{Me} bond is systematically shortened so that the Co-N5 distance (5.68 Å) is fixed. .........................................................84</td>
</tr>
<tr>
<td>4.1</td>
<td>First Half-Catalytic Cycle of Methionine Synthase (MetH) Enzyme Showing Transfer of a Methyl Group from Methylcobalamin (MeCbI) Cofactor to Homocysteine (Hcy).................................................................97</td>
</tr>
<tr>
<td>6.1</td>
<td>Proposed PCET Mechanistic Route for MCM Enzyme........................................153</td>
</tr>
<tr>
<td>7.1</td>
<td>General catalytic mechanism for the AdoCbl-dependent enzymatic reactions...165</td>
</tr>
<tr>
<td>7.2</td>
<td>Reactions catalyzed by MCM. Carboxylic group that interacts with tyrosine residue has been indicated by red box.................................165</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

Vitamin B$_{12}$ is one of the largest and most complex cofactor in biology in terms of molecular weight and number of functional groups. It was discovered from its relationship to pernicious anemia that was a fatal disease until the 1920s,$^{1,2}$ when it was found that liver contain a substance called “anti pernicious anemia factor” that can act as therapeutic solution to pernicious anemia and was later named as a vitamin B$_{12}$. In mammals including human, the deficiency of this factor can also cause severe neurological diseases. The crystal structure of vitamin B$_{12}$ was successfully solved by Dorothy Hodgkin in 1956,$^5$ and for this monumental achievement, she received the Noble Prize for Chemistry in 1964. Soon after, Woodward and Eschenmoser reported a total synthesis of the vitamin B$_{12}$ in a collaborative effort.$^6$ Since then, a vast majority of studies have been reported describing the chemistry, biochemistry, enzymology of B$_{12}$ cofactor including the elucidation of biosynthetic pathway of vitamin B$_{12}$ in aerobic and anaerobic microorganisms.$^6$-$^{11}$

Although vitamin B$_{12}$ or cyanocobalmin (CNCbl) is one of the essential vitamins, but higher organisms cannot synthesize it, thus it must be supplied with the diet. Moreover, CNCbl is not a biologically active species and it does not known to have physiological role itself.$^{12,13}$ However, it is a water-soluble cofactor that dissociates into
two other active forms when orally administered namely, methylcobalamin (MeCbl) and Adenosylcobalamin (AdoCbl) or coenzyme B\textsubscript{12} which are biologically active cofactors in both mammalian and bacterial organisms.\textsuperscript{4,13} The crystal structure of vitamin B\textsubscript{12} was found to be octahedral (Figure 1.1) containing a biochemically rare element cobalt (Co) in its +3 oxidation state.\textsuperscript{5} It is a tetrapyrrolic cofactor, in which the central Co metal is coordinated equatorially to the four nitrogen ligands donated by the pyrroles of the corrin macrocycle. In contrast to the highly symmetrical porphyrins,\textsuperscript{14} the corrin ring of B\textsubscript{12} cofactors is missing one of the bridging carbon atoms that results in a direct bond

---

**Figure 1.1.** Molecular structure of vitamin B\textsubscript{12} (CNCbl = cyanocobalamin, R = -CN), methylcobalamin (R = -Me) and adenosylcobalamin (R = -Ado).
between the two pyrrole rings. In addition, the corrin ring of cobalamin is fairly reduced as compared to its structural analogues chlorins and porphyrins, which makes it more flexible macrocycle. In solution and at physiological pH, the lower or α-axial coordination site is occupied by the 5,6-dimethylbenzimidazole (DBI) of the nucleotide loop that is communicating to the corrin ring by one of its seven amide chains. However, B$_{12}$ cofactors (or cobalamin) are generally distinguished by the nature of an upper (β)-axial ligand (R) bound to the Co, if the R position is occupied by -CN group then it is called cyanocobalmin (CNCbl) or vitamin B$_{12}$. More importantly, when a -CN group is hydrolyzed and it transformed into one of the two biologically active cofactors MeCbl or AdoCbl, which contains alkyl group such as methyl (-Me) and 5'-deoxyadenosyl (-Ado), respectively. These are the only cofactors known to date, containing naturally occurring organometallic metal-carbon (Co-C) bond that involved in enzymatic reactions. There are other non-alkyl groups, which are also found to be coordinated at the β-site of the cofactor, such as water, hydroxy, and glutathionyl that form aquocobalamin (H$_2$OCbl), hydroxycobalamin (HOCbl), and glutathionylcobalamin (GSCbl), respectively.

The X-ray crystal structures of the most important derivatives for the vitamin B$_{12}$, i.e., free AdoCbl$^{16}$ and MeCbl$^{17}$ were successfully characterized, first time in 1961 and 1985, respectively. The most striking feature of MeCbl and AdoCbl cofactors is the first ever known organometallic Co-C bond that provides attractive avenues for the experimental as well as computational research in the field of bioinorganic chemistry. In fact, the bond dissociation energy (BDE) for the Co-C bond is not relatively high rather it is in the range of 32-40 kcal/mol for the MeCbl$^{18-20}$ and 26-34 kcal/mol for the AdoCbl$^{21}$. 
indicating the moderate strength of the Co-C bond. As discussed above, MeCbl and AdoCbl are the cofactors of several enzymes, which catalyze enzymatic reactions in mammals and bacteria.\textsuperscript{4,13} The rapid cleavage of the Co-C bond constitutes the key steps in the catalysis of each enzymatic reaction due to its inherent reactivity. However, both the cofactors are known to bind with the apoenzyme in the base-off form where the lower axial ligand DBI is known to be dissociated and most of the time it is replaced by the histidine residue from the protein scaffold that is called base-off/His-on binding conformation of the corrin.\textsuperscript{8} In nature, B\textsubscript{12}-dependent enzymatic reaction can be divided into three different classes: (I) methyltransferases that catalyze methyl transfer reactions, (II) isomerases that catalyze 1,2-rearrangement reactions, and (III) reductive dehalogenasis which are able to reductively dechlorinate aliphatic and aromatic chlorinated hydrocarbons, one of the least studied classes of B\textsubscript{12}.

MeCbl is the cofactor of a broad class of the methyltransferases.\textsuperscript{25-36} Among them, methionine synthase (MetH) is one of the well-studied corrinoid proteins that catalyze the biosynthesis of methionine, which is one of the essential amino acids both in mammals including human and in bacteria. The structural analogues of the MeCbl such as methylcobamides (i.e., methylcorrinoids) are also the cofactor of enzymes participating in one-carbon metabolism and CO\textsubscript{2} fixation in anaerobic acetogenic bacteria\textsuperscript{37} and methanogenic archaea.\textsuperscript{38} Thus, all the methyltransferases catalyze the transfer of a methyl group from a methyl donors to a methyl acceptor in multi domains or/and multi proteins where methylcorrinoids binding domain serve as an intermediate for the methyl group. The enzymatic mechanism for the methyl transfer requires the formally cleavage and formation of the Co-C bond to initiate the catalytic turnover.\textsuperscript{39}
However, the methyl transfer reactions to and from the cobalamin are extremely challenging and the way in which MeCbl-dependent methyltransferases activates the methyl donors through conformational changes remains ambiguous, the enzymatic mechanisms for the methyl transfer reactions remains the subject of discussion. Particularly, MetH is a single polypeptide (136 kDa and 1227 amino acids) and has a modular architecture with four different binding modules including a CH$_3$-H$_4$Folate-binding donor domain, an homocysteine (Hcy)-binding acceptor domain and a MeCbl-binding domain, which are required for the main catalytic cycle and the fourth one adenosylmethionine (AdoMet) domain is essential for the reactivation cycle. The crystal structure of a full length of MetH enzyme has not been solved yet due to its high degree of conformational flexibility but the crystal structure of each individual domain has been successfully characterized. Moreover, there is no available X-ray structure for the reaction complex comprised of two binding modules, thus the details of the mechanistic pathway for the methyl transfer reactions remain elusive. Therefore, the computational tools are of great help to model the chemical reactions of two binding modules in order to understand the mechanistic pathways for the methyl transfer reactions.

AdoCbl or coenzyme B$_{12}$ is the cofactor of a wide class of enzyme including eliminases and isomerases that catalyze 1,2-rearrangement between an hydrogen atom and an electronegative group on adjacent carbon atoms. Among them, methylmalonyl- CoA-mutase (MCM) is the most studied enzyme that catalyzes the reversible isomerisation of methylmalonyl into the succinyl residue in mammals including human and bacteria. In mammals, MCM is mitochondrial matrix enzyme
that converts methylmalonyl-CoA to succinyl-CoA in catabolic pathways. The coenzyme B₁₂ is also a cofactor of the class II ribonucleotide reductase (RNR), which catalyzes the reduction of ribonucleoside to deoxyribonucleoside in bacteria and is important for DNA replication and repair. In all the mutases, the reversible homolytic cleavage of the Co-C(Ado) bond is required to initiate the catalytic reaction. The most important features of the enzymatic catalysis is the observed rate enhancement, indicating AdoCbl-dependent enzyme catalyzed reaction is trillion-fold faster than that of the uncatalyzed reaction taking place in solution. However, considering the fact that the cleavage of Co-C bond is not the rate determining step, thus its strength must be reduced by ~ 50% in order to achieve such rate enhancement. Moreover, the reaction intermediates involved in the AdoCbl-dependent enzymatic catalysis are very short-lived and most of them are spectroscopically silent, so the actual mechanism for the Co-C bond activation remains elusive. The computational tools based on quantum mechanical methodology are complementary to study the electronic and structural properties of intermediates as well as transition states involved in the reaction mechanism. We applied wide variety of computational tools to understand the mechanism of the initial step of B₁₂ catalysis i.e., the homolytic cleavage of the Co-C bond and subsequent hydrogen atom abstraction from MCM substrate.

This thesis research is divided into two main parts. In first part (chapters 2-5), we employed different computational techniques to investigate the electronic and structural properties of the intermediates as well as reaction complexes involved in the methyl transfer reactions catalyzed by MetH. We explored the possibility of the mechanistic details of the methyl transfer reaction complementing the previous experimental studies.
In chapter 2, we analyzed the electronic structure of enzyme-bound cob(I)alamin intermediate of MetH by means of computational tools using the x-ray structure of the MeCbl-binding domain in order to understand the first half-catalytic cycle of MetH. Because of its high reactivity and supernucleophilic nature, there is no X-ray crystal structure available for an enzyme-bound cob(I)alamin intermediate, therefore the enzyme-bound structure was prepared computationally. Most importantly, the observed displacement of the histidine (His759) residue from the Co(I) center is in agreement with structure of the free cob(I)alamin in solution, indicating the Co(I) is not axially coordinated inside the MetH enzyme. This study further reveals that the transfer of the methyl group to the substrate is associated with the displacement of the (His759) axial base of protein scaffold. In addition, high-level ab initio calculations show that the ground state of the cob(I)alamin is multiconfigurational where the diradical Co(II)-corrin radical configuration (formed by electron transfer from the cobalt to the corrin ring) contributes the electronic structure of the cob(I)alamin intermediate, indicating a non-innocent behavior of the corrin ring. In chapter 3, we applied docking studies based on the electron structure of the enzyme-bound cob(I)alamin intermediate of MetH to investigate the mechanistic details for the methyl transfer reaction from 5-methyltetrahydrofolate to cob(I)alamin. Out computational study reveals the traditionally assumed S_N2 mechanism for the formation CH_3-cob(III)alamin and the activation energy barrier for S_N2 reaction was found to be comparable with respect to the determined experimental rate constant. However, the possibility of ET-based radical mechanism consistent with the close-lying diradical states has also been suggested, where an electron transfer (ET) from His-on cob(I)alamin to pterin ring of the protonated CH_3-H_4Folate
radical transfer. In addition, the mechanistic details of the second-half catalytic cycle of MetH is discussed in chapter 4, where the methyl transfer takes place from MeCbl form of the cofactor to homocysteine. In this chapter, the possibility of both SN2 and ET-based radical mechanisms have analyzed based on the energetic of ionic and diradical states as a function of the C-S distance between the methyl group of MeCbl and the sulfur of the methylthiolate. The most important finding of this work is the energetic modulation of ionic and diradical states with respect to distance between reacting modules indicating SN2-type and ET-based radical mechanism, respectively. In chapter 5, we investigate the electronic and structural properties of MeCbl and its one-electron-oxidized form to understand the oxidative mechanism of the Co-C bond, which remains unclear from experimental point of view. The most noteworthy result of this study, regardless of the type of method used, is the significant lowering of the Co-C bond dissociation energy in one electron oxidized form. The second part of the dissertation is focused on AdoCbl-dependent enzymes, i.e., the investigation of enzymatic mechanism of the initial step in B12 catalysis (MCM) that comprises the Co-C bond cleavage and subsequent H-atom abstraction from the substrate. In chapter 6, the alternative role of the tyrosine residue in the vicinity of AdoCbl cofactor has been studied which serves as an ET source to reduced the AdoCbl in case of MCM enzyme. The suggested proton coupled electron transfer (PCET) mechanism is consistent with the homolytic cleavage of the Co-C bond in the one electron reduced AdoCbl cofactor instead of its neutral analogue. Therefore, the energetic profile of the Co-C bond cleavage followed by a hydrogen transfer from methylmalonyl substrate in the reduced AdoCbl cofactor has been investigated in chapter 7, by considering the entire protein environment. The purpose of this study is to further
elaborate the reductive cleavage of the Co-C bond and subsequent hydrogen atom abstraction step.

1.1 Computational Methodology

Three different theoretical methods were used to study the electronic structure of intermediates and reaction complexes involved in B$_{12}$-dependent enzymatic reactions. At first, we employed hybrid quantum mechanics/molecular mechanics (QM/MM) approach to study the reaction complexes that incorporate the protein environment. This technique combines two different levels of theories, particularly, a quantum mechanical methods with a force field based molecular mechanics methodology. In QM/MM methodology, the active part of the molecular system or the region where the cleavage and the formation of chemical bond takes place is usually treated quantum mechanically whereas the rest of the protein environment is considered at molecular mechanics level. In addition, we employed density functional theory (DFT) based quantum mechanical methods for the further characterization of the electronic structure of the molecular systems. Although DFT methods are very efficient and relatively less computer costly, but there are some limitations which cannot be included at this level of theory. The major shortcoming of the DFT is that it is based on the single reference wave function and it fails to reproduce multiconfigurational nature as well as near-degeneracy of a molecular systems. Therefore, in a third place, we employed complete active space self consistent field (CASSCF) methodology that is based on the multireference wave function. The main advantage of using this multiconfigurational methodology for the bioinorganic molecular system is that it can characterize more than one configuration, which could be
used to understand the mechanistic details of the enzymatic reactions. The details of all computational methodologies employed in this thesis are given below:

1.1.1 Hybrid QM/MM Calculations. To investigate electronic structure and mechanistic details of the reaction complexes of B₁₂-mediated enzymatic reactions, ONIOM-based QM/MM analysis was carried out using experimental crystal structures. The crystal structures of the protein was usually obtained from the Protein Data Bank (PDB ID) which were then used to build a reasonable computational model. At first, the experimental artifacts were removed from the crystal structure and therefore, no counterions were included in the prepared structure model. Subsequently, the H-atoms were incorporated using VMD and GaussView assuming the normal protonation state of all titrable residues except histidine. The protonation state of histidine residue that can be protonated at δ or/and ε was determined based on the local H-bonding network via visual inspection and using PROPOKA suite of program. As a next step, AMBER parameters were assigned for the B₁₂ cofactor as developed by Marques et al. The model structure was then subjected to geometry optimization of all the hydrogens while keeping the heavy atoms fixed using the AMBER force field as implemented in Gaussian09 program. Subsequently, AMBER optimization was performed on the side chains of the protein residues except lower axial His residue. We carried out ONIOM (BP86/6-31G(d):AMBER) Mechanical Embedding (ME) single-point calculations to determine Merz–Kollman electrostatic potential (ESP) atomic charges for the Quantum Mechanical (QM) region. In this ONIOM-ME calculation, the B₁₂ cofactor was used as a part of the QM model system and the radius of the Co was set at 0.79 Å. After the preparatory calculations, the entire B₁₂ cofactor, as well as His was included in the QM
region where His was capped as an H-link atom. The protein residues within 20 Å of the Co center was used for geometry optimization first by the ONIOM-ME method followed by ONIOM-electronic embedding (EE) formalism.

In addition, the hybrid QM/MM calculations were also performed with other implementation using the method developed by Laio et al.,\textsuperscript{70} which combines the first-principles MD method of Car and Parrinello\textsuperscript{71} with a force-field MD methodology (i.e., QM/MM CPMD). These calculations are used in chapters 2 and 6. In short, the geometries of the B\textsubscript{12} based cofactors and intermediates were optimized as a closed-shell singlet using the BP86 functional,\textsuperscript{67,68} a plane wave basis set with a 70-90 Ry kinetic energy cutoff, and Martins–Troullier pseudopotentials\textsuperscript{72} to describe the interaction between the ionic cores and the valence electrons. Two pseudopotentials were tested for the cobalt atom: one with nine valence electrons supplemented with nonlinear core corrections \textsuperscript{73} and another with 17 valence electrons. To explore the possibility of spin polarization between the cobalt and the corrin, a single point calculation within the local spin density approximation (LSD) was also performed.

1.1.2 DFT Calculations. To get further insight into the complex electronic properties of the B\textsubscript{12}-mediated enzymatic reactions suggested by the QM/MM calculations, we utilized DFT calculations for the model systems. In all the chapters, the geometry optimization of all the structural models was carried out employing the Becke–Perdew (BP86)\textsuperscript{67,68} functional and the 6-31G(d) (5d components) basis set, as implemented in Gaussian 03 and 09.\textsuperscript{74} This level of theory constitutes an appropriate platform for describing the structural and electronic properties of alkyl-cobalt-(III) complexes, as documented in the literature.\textsuperscript{75-77} However, a pure generalized gradient approximations
(GGA) functionals such as PBE\textsuperscript{78} or BP86 are known to be unable to converge to a spin polarized solution for Co(I) complexes.\textsuperscript{79,80} This is not the case for hybrid functionals such as B3LYP, despite the fact that this functional underestimates the strength of the Co–C bond of the B\textsubscript{12} cofactors.\textsuperscript{79,81} Therefore, we used B3LYP to test the possibility of a spin-polarized solution, which may be indicative of a more complex electronic wave function.\textsuperscript{82} In particular, B3LYP calculations were initially performed on structural models guided by QM/MM calculations having an even number of electrons, assuming a closed-shell singlet wave function in chapter 2. Subsequently, the unrestricted Kohn–Sham formalism using UB3LYP was applied by mixing HOMO and LUMO orbitals to obtain the corresponding open-shell singlet.

In addition, we also carried out Density Functional Reactivity Theory (DFRT) to characterize the charge separation propensity of donor-acceptor complex in Chapter 6. We employed the electrophilic Fukui function and the dual descriptor formalism from DFRT\textsuperscript{83-87} to demonstrate thermodynamic-like support of the PCET mechanism (Scheme 1) by analyzing the propensity of electron to be transferred to a site other than the proton acceptor site. The dual descriptor concept in DFRT with N total number of electrons in the system is defined\textsuperscript{88-90} as:

\[
f^{(2)}(\mathbf{r}) = \left( \frac{\partial f(\mathbf{r})}{\partial N} \right)_{\nu}
\]

\(f(\mathbf{r})\) is the Fukui function\textsuperscript{91} which indicates the response in the electron density when the number of electrons is changed in the presence of external potential \(\nu\) from the atomic nuclei. Acting as reactivity descriptors, the \(f^+(\mathbf{r})\) or \(f^-(\mathbf{r})\) functions measure the response of electron density change upon the addition or removal of an electron, thus they can be
used to quantify electrophilicity or nucleophilicity of the system. Considering these two approximations, dual descriptor becomes:

\[ f^{(2)}(\mathbf{r}) = f^+(\mathbf{r}) - f^-(\mathbf{r}) \approx \rho_{\text{LUMO}}(\mathbf{r}) - \rho_{\text{HOMO}}(\mathbf{r}) \]

The \( \rho_{\text{LUMO}}(\mathbf{r}) \) and \( \rho_{\text{HOMO}}(\mathbf{r}) \) are the LUMO and HOMO densities, respectively. In other words, dual descriptor \( (f^{(2)}(\mathbf{r})) \) provides a framework to determine electrophilic and nucleophilic regions in an acid-base (or donor-acceptor) complex and once these regions are sufficiently separated, there will be a propensity for charge separation.

1.1.3 CASSCF Calculations. The Kohn–Sham formalism (which is the base for DFT-based computations) is restricted to a single Slater determinant description, it cannot describe the multiconfigurational character of the \( \text{B}_{12} \) based cofactors. Thus, we carried out CASSCF multireference calculations, followed by quasi-degenerate perturbation theory (QDPT2)\textsuperscript{92} calculations with a multiconfigurational self-consistent-field reference function (MC-QDPT2) to include the dynamical correlation, as implemented in the PC GAMESS/Firefly QC package.\textsuperscript{93} The MC-QDPT2 calculations, aimed to correct the CASSCF excitation energies for the leading dynamical correlation effects that were performed using the modified version of the MC-XQDPT2 approach, abbreviated as MC-XQDPT2, as implemented in the PCGAMESS/Firefly QC package.\textsuperscript{94} We recall that the MC-XQDPT2 approach is a genuine multi-root (i.e., perturb-then-diagonalize) multi-reference perturbation theory technique using the CASSCF state-averaged (SA) electron density over the ground and excited states of interest, which allows a true mixing of the CASSCF zero-order states via the construction and the subsequent diagonalization of the effective Hamiltonian. All of the CASSCF calculations were performed on the DFT

13
optimized structures of the B_{12} models using the 6-31G(d) basis set. The details regarding the active space chosen for these multireference calculations are described in chapter 2.

1.2 Mechanistic Details of Methyl Transfer Reactions in Methionine Synthase (MetH)

Cobalamin-dependent methionine synthase (MetH) enzyme from *E. coli* and *H. Spein* is one of the best-studied members of cobalamin- and corrinoid-dependent methyltransferases.\textsuperscript{25-29,31-36,95-97} MetH is a single polypeptide (136 kDa and 1227 amino acids) enzyme and has a modular architecture and the methyl transfer reactions in the catalytic cycle involve interaction among four different binding domains.\textsuperscript{41,98} However, the crystal structure of the whole enzyme has not been yet resolved mainly owing to the very high degree of conformational flexibility, but the X-ray crystal structure of the individual domains,\textsuperscript{41,98-100} including the one that binds MeCbl,\textsuperscript{41} Hcy,\textsuperscript{98} and CH_{3}-H_{4}folate\textsuperscript{100} have been successfully characterized.

MetH enzyme catalyzes the transfer of a methyl group from methylcobalamin (MeCbl) to homocysteine (Hcy) to generate methionine (Met) and cob(I)alamin (Figure 1.2) intermediate. The resulting cob(I)alamin is further remethylated by methyltetrahydrofolate (CH_{3}-H_{4}Folate or CH_{3}-THF) substrate to generate back the MeCbl and the tetrahydrofolate (H_{4}Folate). These two half-reactions involved in the catalytic turnover of the MetH where the cobalamin cofactor plays a crucial role in the catalysis (Figure 1.2A), acting both as methyl donor and acceptor.\textsuperscript{41,51} However, in the reactivation reaction cycle, a cob(I)alamin (Figure 1.2B, lower part) form of the cofactor which serves as methyl acceptor, is occasionally oxidized to inactive cob(II)alamin once in every 2000 catalytic cycles and under aerobic conditions. This inactive form of the
cofactor further reactivated to MeCbl resting state by taking an electron from reduced flavodoxin and methyl group from S-adenosyl-L-methionine (AdoMet). The cobalamin

![Diagram of the catalytic cycle and reactivation cycle of Cobalamin-Dependent Methionine Synthase (MetH).](image)

Figure 1.2. (A) The reaction scheme catalyzed by methionine synthase (MetH), (B) Structure of methylcobalamin and cob(I)alamin cofactor, and (C) Structure of different substrates used in MetH.

cofactor in MetH participates in three different methyl transfer reactions involving three different metabolites: homocysteine and \( \text{CH}_3\text{-H}_4\text{folate} \) in the catalytic reaction cycle while AdoMet exists in the reactivation cycle (Figure 1.2C). The coordination of cobalamin cofactor plays a major role in controlling the access to catalytic and reactivation cycles of the MetH that is subsequently responsible for the positive charge and oxidation state of cobalt metal.\(^{101,102}\) The methylcob(III)alamin is six coordinated in a free state with a nitrogen atom of the lower axial DBI as shown in Figure 1.2B. In
contrast, when MeCbl is bound to MetH, the DBI ligand, which is coordinated to the metal in a free state is replaced by a histidine (His759) residue from rossman subdomain (α-face) of the cobalamin binding domain.

The catalytic cycle of the MetH comprises of two half-reactions that are given in equations 1 and 2 (Scheme 1.1A). The most important and less understood reaction is the transfer of a methyl group from a CH₃-H₄Folate substrate to a cob(I)alamin/cob(I)amide form of the cofactor (eq. 1). This particular reaction is common in a wide range of methyltransferases, namely, methionine synthase (MetH) from E. coli, methyltetrahydrofolate-corrinoid-iron/sulfur protein (CFeSP) AcsE from C. thermoaceticum, and bacterial mercury methylation HgcAB from D. desulfuricans ND132. During the course of first half-catalytic reaction (eq.1), the two domains enclosing cob(I)alamin and CH₃-H₄folate substrate form a reaction complex in which the substrate interacts with the cob(I)alamin from the upper (β) face of the cofactor.

Scheme 1.1. The first-half (1) and second-half (2) catalytic reaction of the MetH.

In the second half-catalytic reaction, the transfer of the methyl group takes place from MeCbl to Hcy substrate (eq. 2). However, there is no structural information available
with regard to the reaction complexes and therefore, the details of the reaction mechanism involving demethylation CH$_3$-H$_4$folate substrate and methylation of Hcy and remains elusive.$^{39,41,99,100}$ Therefore, the challenge for chemists is to determine the enzymatic mechanisms and the source of activation of an unreactive methyl donor such as CH$_3$-H$_4$Folate and MeCbl by modeling the intermodular methyl transfer reactions.

The net retention of stereochemistry of the methyl group, which is being transferred to and from cobalamin, is observed for the overall reaction catalyzed by the MetH.$^{110}$ This means the displacement of the methyl group takes place with the same mechanistic pathway in both half-catalytic MetH reactions in order to exhibit the similar stereochemistry at the carbon. However, several mechanistic routes have been proposed for both the methyl transfer reactions involved in the catalytic cycle of the MetH.$^{48,111}$ Although it is generally believed that the methyl transfer catalyzed by MetH enzyme proceed via S$_N$2-type nucleophilic displacement,$^{48,111}$ but alternative mechanistic pathways including oxidative addition/reductive elimination as well as mechanism involving single electron transfer (SET) have also been suggested by Matthews et al.$^{111}$ The major limitation of the oxidative addition mechanism in the first-half catalytic reaction is a three-centered bond requirement between the C$_{Me}$-N5 bond of the CH$_3$-H$_4$Folate and the Co metal of cob(I)alamin complex, which might not be possible from enzymatic point of view due to steric interactions in the protein backbone of the two reacting modules. Alternatively, according to the reductive elimination proposal for the second half-catalytic reaction,$^{111}$ a three-centre bond is also required among the Co center of the MeCbl, C$_{Me}$ of the methyl and the S$_{Hcy}$ of the substrate moieties of the reacting modules that lead to the formation of triangle-shape intermediate. Taking into account the
requirement for the three-centered intermediate, it is very difficult to achieve the proper alignment of all three entities involved in the bond formation from computational point of view because of the steric nature protein backbone of two reacting modules.

The mechanistic details of the methyl transfer reaction are unclear from experimental point of view, mainly due to the lack of crystal structure of the reaction complexes as well as due to the very short life time of its intermediates; we turned our attention to computational modeling in order to gain mechanistic insights for the catalytic methyl transfer reactions. In order to study first-half catalytic reaction of the MetH that is the transfer of a methyl group from CH$_3$-H$_4$Folate substrate to the cob(I)alamin intermediate, first we need to investigate the electronic structure of the cob(I)alamin intermediate inside the enzyme because there is no available X-ray structure. During the catalytic methyl transfer reaction (Figure 1.2A), the different binding domains interact with each other in order to either cleave or generate the Co-C bond. Simultaneously, the His759 residue is expected to move with respect to the Co metal center to cleave/form the Co-N$_{His759}$ bond by analogy with the solution chemistry of the B$_{12}$ cofactor. However, the crystal structure of the Me-cob(III)alamin resting state (PDB code: 1BMT, 3 Å resolution) shows that His759 interacts with Asp757 and Ser810 through a network of hydrogen bonds, suggesting that His759 would be fixed with respect to the Co center. On the other hand, a structural study of the reactivation complex (Figure 1.2A) formed by the AdoMet and cob(I)alamin binding domains shows that the axial His759 moves away from the cobalt center and makes specific contacts with the AdoMet domain. Thus, it is still ambiguous how His759 interacts with the cob(I)alamin cofactor during the catalytic cycle. Therefore, in chapter 2, the coordination of the axial His759 and the cob(I)alamin
cofactor inside the enzyme was explored by employing the hybrid quantum mechanical/molecular mechanical (QM/MM) approach. Subsequently, the complex electronic structure of the cob(I)alamin intermediate suggested by the QM/MM results was further analyzed using density function theory (DFT) and complete active space self-consistent field (CASSCF) followed by second order perturbation theory (QDPT2) calculations on gas phase models. In this chapter, the implications of this computation studies are discussed in the context of the methyl transfer reaction between the Me-H4Folate and cob(I)alamin.

In chapter 3, we further applied hybrid QM/MM calculations to investigate and understand the E.S reaction complex involved in the first-half catalytic reaction, which was prepared by docking the CH3-H4Folate substrate in the β-face of the enzyme-bound cob(I)alamin. Based on the electronic structure and coordination of the cob(I)alamin intermediate in the cob(I)alamin:CH3-H4Folate reaction complex, two different mechanistic pathways were explored: S\textsubscript{N}2-type nucleophilic displacement and ET-based radical mechanism where the electron is transferred from His-on cob(I)alamin to pterin ring of the CH3-H4Folate followed by methyl radical. Furthermore, the results of this chapter including mechanistic details and the protonation process of the CH3-H4Folate were extensively discussed that can be applied to a broad class of corrinoid-methyltransferases.

The similar mechanistic details of the methyl transfer reaction in second half-catalytic cycle that is from MeCbl to Hcy substrate were discussed in chapter 4. Both S\textsubscript{N}2 as well as reductive cleavage pathways have been investigated computationally for model systems using DFT.\textsuperscript{114,115} In addition, the feasibility of ET-based reductive
cleavage mechanism has been demonstrated using QM/MM calculations.\textsuperscript{116} Although these studies proved basic concepts, they opened several issues that require further analysis. The main difference between these two mechanisms is the C-S distance in which methyl transfer takes place between the Hcy substrate and the MeCbl. An approximate information can be obtained from DFT which is based on a single reference wave function, and thus a more complex wave function is required to describe the reactant complex. Therefore, the aim of the chapter 4 was to analyze the electronic and structural properties of MeCbl:Hcy reaction complex using the CASSCF method followed by the second-order perturbation theory (MR-PT). In addition, the electronic structure of the reaction complex was further investigated as a function of C···S distance between reaction centers and its possible implications with regard to the enzymatic reaction mechanism are discussed.

At last, the implication of the Co-C bond cleavage in MeCbl and its one electron oxidized (1e-Ox) form was investigated by means of DFT and CASSCF calculations in chapter 5. Based on the previous experimental studies,\textsuperscript{117-128} the electronic structure of the 1e-Ox cobalt corrinoids has been the subject of debate, because the exact location of the unpaired electron remains elusive. Thus, we applied a DFT for the 1e-Ox MeCbl and its related metalloporphyrins in order to gain insight into the Co–C bond properties. Subsequently, the time-dependent DFT (TD-DFT) calculations were carried out to investigate the nature of low-lying excited states and validate the oxidative cleavage mechanism. The complex electronic structure of the 1e-Ox MeCbl revealed by the DFT was further analyzed using the CASSCF calculations. A particular emphasis was placed
on the Co–C bond cleavage in the 1e-Ox form of MeCbl and the relevant mechanism responsible for lowering of the dissociation energy as compared to its neutral analogue.

1.3 Catalysis in Adenosylcobalamin-Dependent Mutases: Methylmalonyl Coenzyme A Mutase (MCM)

Coenzyme B$_{12}$ or AdoCbl is the cofactor to a number of enzymes from a range of organisms that catalyze a wide variety of challenging reactions.$^{9,25-27,44-52,95}$ In all the cases, the Co-C bond of the AdoCbl undergoes homolytic cleavage upon substrate binding to initiate the enzymatic catalysis (Scheme 1.2). However, the generated 5′-deoxy-5′-adenosyl (Ado$^\cdot$) radical plays a critical role in catalyzing various difficult reactions by abstracting H-atom from the substrate and subsequently generating substrate radical that undergoes intramolecular 1,2-rearrangement with the functional groups (such as alkyl, hydroxy) to an adjacent carbon in order to form a product radical. At last, the product radical re-abstracts the H-atom from Ado moiety to form the product complex that completes the catalytic turnover. There are many fundamental questions related to the radical mediated catalysis, which need to be addressed. The most important features of the AdoCbl-dependent enzymatic catalysis is the achieved trillion ($10^{12}$)-fold rate enhancement$^{18,19,129-131}$ of the Co-C bond (catalytic rate, $k_{cat}=300$ s$^{-1}$) as compared to the thermal homolysis of the free cofactor in solution (with rates of $10^{-9}$ S$^{-1}$).$^{57}$ Although, it has been known that the homolytic cleavage of the Co-C bond is very fast but it is not a rate-determining step for the AdoCbl-dependent enzymatic reactions indicating the Co-C bond strength must be decreased in the enzymatic reactions as compared to its strength in the solution.$^{57,130-132}$ Moreover, the Co-C bond cleavage in isolated coenzyme B$_{12}$ is energetically unfavorable with the bond dissociation energy (BDE) estimated at ~31.0
kcal/mol (with rates of $10^9$ S$^{-1}$ at 37°C).$^{21-24}$ In contrast to the homolysis of the Co-C bond inside enzyme,$^{57,130-132}$ its strength must be significantly reduced by 50%

Scheme 1.2. General catalytic mechanism for AdoCbl-dependent enzymes.

($\sim$17.0 kcal/mol) during the catalytic turnover.$^{133}$ In addition, an enzyme environment also shifts the equilibrium constant towards the homolysis products giving equilibrium constant close to unity.$^{130-132,134-136}$ It has been also suggested that homolytic cleavage of the Co-C bond is kinetically coupled to the hydrogen atom abstraction from the substrate.$^{130-132,134-139}$ Therefore, understanding the legitimate reason for the Co-C bond activation in AdoCbl-dependent enzymes is the most critical problem in B$_{12}$ chemistry.

Although, the unified mechanism for the enormous rate enhancement of AdoCbl-enzymatic reactions is not known, the weakening of the Co-C bond in one electron reduced (1e-Red) form has attracted the interest of experimental as well as computational scientists.$^{101,140-144}$ Initially, it has been demonstrated by Lexa and Seveant$^{101,140}$ based on the electrochemical studies that the addition of an electron to the B$_{12}$ cofactors reduced
the BDE of the Co-C bond by a significant amount which might be important from enzymatic point of view in order to understand the rate acceleration of Co-C bond inside the enzyme. Despite the attractive features of the reductive cleavage mechanism of the Co-C bond, this mechanism was not realized inside the enzyme until Birk et al., and our research group reported their findings.\textsuperscript{115,116,143-146} Therefore, the Co-C bond cleavage in the 1e-reduced (1e-Red) form of the cobalt corrinoids has attained a considerable interest because the Co-C bond dissociation energy (BDE) is significantly decreased.\textsuperscript{115,141,143,144,146} The 1e-Red B\textsubscript{12} cofactors (Figure 1.1) and related model complexes have been investigated experimentally\textsuperscript{14,140,147-151} as well as computationally to explore the possibility of the reductive cleavage mechanism.\textsuperscript{115,116,146} It has been found using density functional theory (DFT) that the addition of an electron results in the formation of a $\pi$-corrin radical anion.\textsuperscript{115,116,146} Subsequently, it has been shown that when the Co-C bond was stretched the electron shifts from corrin-based $\pi^*$ to bond-centered $\sigma^*_{\text{Co-C}}$ orbital, which results in the cleavage of three-electron $(\sigma)^2(\sigma^*)^1$ bond. Thus, these findings have demonstrated $\sim 45\%$ lowering of the BDE in the case of MeCbl, which might be due to the involvement of two electronic states. However, the possibility of the reductive cleavage mechanism of the Co-C bond has also been demonstrated in chapter 4, i.e., in the second half-catalytic cycle of the MetH where the electrochemical reduction takes place form upronated methylthiolate to MeCbl that induced the Co-C cleavage of the cofactor followed by methyl radical transfer to methylthiolate radical. Although there is significant amount of reduction of the Co-C bond strength of the B\textsubscript{12} cofactors, the reductive cleavage mechanism has not been considered as a possible route for the trillion-fold rate acceleration of AdoCbl-dependent enzymes because of the unfavorable redox
potential. The reduction potential of the isolated AdoCbl in the formamide/propanol solvent mixture has been determined to be -1.31 V versus saturated calomel electrode (SCE) which is significantly negative that of the isolated tyrosine residue (+0.51 V and +0.76 V vs. SCE), an enzyme-bound AdoCbl can be altered by a various interactions inside the protein scaffold. In that context, it has been widely reported in the case of photosystem II, superoxide dismutase that protein residues can tune the redox potential of the cofactor or the active site by a sizeable amount. These studies suggest that the redox potential of B12 cofactor can also be controlled by the protein residues, implying the possibility of the ET-based activation of the AdoCbl-dependent enzymes.

Methylmalonyl CoA mutase (MCM) is one of the well-studied enzymes among all the known mutases that utilize AdoCbl cofactor for catalysis. MCM enzyme catalyzes the isomerisation of methylmalonyl-CoA to succinyl-CoA (Scheme 1.3) and is

![Scheme 1.3. Reactions catalyzed by Methylmalonyl CoA mutase (MCM).](image)

the only example of B12-dependent family, which is present in both mammals and bacteria. The first step is common in all mutases i.e., homolytic cleavage of Co-carbon bond of AdoCbl cofactor to form Ado radical and cob(II)alamin form of the cofactor. The most remarkable aspect of the AdoCbl-dependent enzyme catalyzed reaction is the observed trillion-fold rate enhancement as compared to the uncatalyzed reaction in solution. Although several mechanisms have been proposed for the Co-C bond labilization on substrate binding, including upward distortion of the corrin
ring, trans effect of the axial ligand and steric interference with the adenosyl group, the precise mechanism responsible such enormous rate acceleration remains elusive. On the hand, the X-ray crystallography studies of MCM (4REQ) has revealed that the binding of substrate induce the large conformational changes and the active side residue that undergoes significant changes is Y89 which moves from its positions and forms a hydrogen-bond contacts with the substrate. The hypothesis of the involvement of Y89 residue in the active site of MCM has been further investigated by site directed mutation where Y89 is mutated by phenylalanine (Y89F) and found that the mutated crystal structure (PDB id: 5REQ) has minor effect on structural properties of coenzyme B12. Indeed, the estimated effect of Y89F mutation on the Co-C bond homolysis was observed to be $\sim 10^3 - 10^4$ of the total $10^{12}$ -fold rate enhancement which might be due to contribution of a single Y89 residue.

Due to the low resolution of the X-ray structure of MCM, the precise core residues (in the vicinity of cofactor) involved in the activation of Co-C bond as well as relevant mechanism associated to trigger the enzymatic catalysis is yet to be realized in order to understand the enzymatic reactions. Recently, a family of AdoCbl-based enzymes have shown a common structural logo of Y residue in close proximity to the cofactor, which acts as internal redox center to reduce the cofactor (i.e., [AdoCbl$^-$]). However, the feasibility of Proton-Coupled Electron Transfer (PCET) route has been suggested as an alternative mechanism for the family of AdoCbl-dependent mutases, but the source of proton transfer is yet to be identified. Alternatively, the PCET mechanistic proposal can involve the substrate as an alternative source of proton acceptor to trigger PCET mechanism. In particular, the crystallography analysis of the active site
of MCM enzyme reveals that the pKa of Y89 (~10) is higher than that of carboxylic group (–COO\(^-\), ~4) in physiological environments i.e. at pH ~7,\(^{53}\) thus the displacement of phenoxy proton of Tyr residue towards the –COO\(^-\) group of substrate might be occurring via PCET route as the concerted coupling of PT and ET steps will avoid the generation of charged intermediates.

Taking into account the experimental\(^{53,54,158}\) as well as theoretical studies\(^{160,161}\) of the MCM enzyme, an interesting question arises regarding the role of the substrate binding to displace the Y89 active site and how precisely these structural changes induce the radical-based enzymatic reactions. Thus, in chapter 6, the thermodynamic feasibility of PCET process for MCM enzyme employing the dual descriptor and the electrophilic Fukui functions from density functional reactivity theory (DFRT) is demonstrated. In this chapter, we reported an independent proof of a PCET mechanistic pathway where the coupling of both proton-transfer (PT) and electron-transfer (ET) steps are taken into consideration and analyzed using a newly advocated density functional reactivity theory (DFRT). This theoretical approach advances our understanding to categorize the charge separation propensity of electron–proton transfer reactions as radical-based enzymatic reactions are often difficult to investigate. Considering the fact that one electron reduction of the B\(_{12}\) cofactors reduced the significant amount of Co-C BDE and exerts significant catalytic effect, the initial step of the Co-C bond cleavage and their subsequent hydrogen abstraction step has been explored in chapter 7. In this chapter, an alternate mechanism for the activation of Co-C bond cleavage in MCM enzyme has been explored by applying hybrid QM/MM calculations. The first objective of this chapter is to further elaborate the proposal and provide energetic implications regarding involvement of the
[AdoCbl]- radical anion in the initial step of AdoCbl-dependent enzymatic catalysis. The energetic of the one-electron reduced AdoCbl cofactor containing all intermediates as well as the transition states along the reaction pathway has been characterized and compared with its neutral analogues. The second aim of this work is to characterize how the orientations of the ribose moiety of coenzyme B_{12} influence the ability of Co-C to facilitate the H-atom-abstraction step. In addition, another hybrid QM(DFT)/MM calculations were performed using the method developed by Laio et al.,\textsuperscript{70} which combines the first-principles MD method of Car and Parrinello\textsuperscript{162} with a force-field MD methodology (i.e., QM/MM CPMD) to compute the free energy surface the entire enzymatic reaction that includes the cleavage of the Co-C bond and subsequent hydrogen abstraction from the substrate.
CHAPTER 2

ROLE OF THE AXIAL BASE IN THE MODULATION OF THE COB(I)ALAMIN ELECTRONIC PROPERTIES

2.1 Introduction

The key step in the catalytic cycle of the cobalamin-dependent methionine synthase (MetH) enzyme is the transfer of a methyl group from the methylcobalamin (MeCbl) cofactor to the homocysteine (Hcy) substrate.\textsuperscript{25,27,29,31-36,95-97,163} The resulting cob(I)alamin intermediate is re-methylated by methyl-tetrahydrofolate (Me-H\textsubscript{4}folate) to generate back methyl-cob(III)alamin and tetrahydrofolate (H\textsubscript{4}folate; Scheme 2.1).\textsuperscript{41,42,164-166} In other words, during the catalytic cycle, the cobalt center oscillates between methyl-cob-(III)alamin and cob(I)alamin. The MetH enzyme has a modular architecture, and the catalytic methyl transfer involves the interaction among different domains.\textsuperscript{42} The crystal structure of the whole enzyme has not yet been resolved, mainly owing to the very high degree of conformational flexibility, but the X-ray crystal structure of individual domains, \textsuperscript{113,165,167,168} including the one that binds MeCbl,\textsuperscript{41} has been well characterized. During the course of catalytic reaction (Scheme 2.1), the two domains binding the MeCbl cofactor and the Hcy substrate form a reaction complex in which the substrate interacts with the MeCbl from the upper face of the cofactor.

\textsuperscript{1} The material presented in this Chapter was reprinted (adapted) with permission from Kumar, N.; Alfonso Prieto, M.; Rovira, C.; Lodowski, P.; Jaworska, M.; Kozlowski, P. M. J. Chem. Theory Comput. 2011, 7, 1541. Copyright 2011 American Chemical Society.
Scheme 2.1. The Catalytic Cycle (Shown in Blue) and the Reactivation Cycle (Shown in Red) for the Cobalamin-Dependent Methionine Synthase (MetH)

$^{a}$H$_4$folate, tetrahydrofolate; CH$_3$-H$_4$folate, methyl-tetrahydrofolate; Fld$_{\text{oxd}}$, oxidized form of flavodoxin; Fld$_{\text{red}}$, reduced form of flavodoxin; AdoHcy, adenosyl-homocysteine; AdoMet, adenosyl-methionine.

However, there is no structural information available with regard to the reaction complex, and therefore, the details of the reaction mechanism involving methylation of the Hcy substrate remain a subject of debate. It is generally believed that the enzyme operates via an $S_N$2-type nucleophilic displacement, though an alternative mechanism in which the one electron reduction of the MeCbl cofactor takes place has also been proposed. This pathway does not impose specific geometrical and distance constraints with respect to the substrate and cofactor as does the $S_N$2 mechanism, which may be advantageous from the enzymatic point of view. The transfer of the methyl group from the MeCbl cofactor to the Hcy substrate results in the formation of the cob(I)alamin intermediate. In solution, this complex is tetra-coordinated because the change of
the Co oxidation state-induced detachment of the axial 5,6-dimethylbenzimidazole (DBI) base (Figure 2.1a).\(^{112}\)

Figure 2.1. (a) Molecular structure of free methylcobalamin, where \(R = \text{CH}_3\), \(R_1 = \text{CH}_2\text{CONH}_2\), \(R_2 = \text{CH}_2\text{CH}_2\text{CONH}_2\), and \(R_3 = (\text{CH}_2)_2\text{CONHCH}_2\text{-CH}(\text{CH}_3)\text{OPO}_3\). (b) Gas phase model of the Im\(^-\)[Cob(I)alamin] enzymatic intermediate in the present chapter.

However, the question is how such displacement of the axial His base takes place inside the enzyme. Due to its high reactivity, there are no structural data available for the
enzyme-bound cob(I)alamin. Several theoretical studies have investigated the electronic and structural properties of the tetra-coordinated cob(I)alamin complex (i.e., without the axial base) using density functional theory (DFT) as well as multireference CASSCF calculations. 80,171-173 Jensen and Ryde 172 first noticed that the B3LYP-based wave function of the ground-state has a singlet instability, indicative of a complex wave function involving other configurations than Co I(d 8). Later, Jensen et al., 80 demonstrated that the ground state of cob(I)alamin is multiconfigurational using CASSCF/CASPT2 calculations. The correlated wave function was found to consist of a closed-shell Co I(d 8) configuration (67%) and a diradical Co II(d 7)-radical corrin (π* 1) configuration (23%). The formation of such unusual electronic configuration was explained by the overlap of low-lying metal d orbitals with ligand orbitals at the singlet ground state, allowing the transfer of an electron from the Co to the corrin ligand. In this regard, the electronic structure of cob(I)alamin resembles the heme-based compound I intermediates, as pointed out by Ryde and Jensen. 174 Both cofactors have a non-innocent macrocycle that can exchange electrons with the metal. In compound I, one electron is transferred from the porphyrin to the Fe metal atom, whereas in the case of cob(I)alamin, electron transfer takes place in the opposite direction, i.e., from Co to the corrin ligand. Indeed, this proposed electron transfer is in agreement with a previous TD-DFT(B3LYP) study by Jaworska and Lodowski, 171 who found that the lowest energy bands (at 700 and 554 nm) in the electronic absorption spectrum of cob(I)alamin have d→π* metal-to-ligand charge transfer (MLCT) character. However, a subsequent TD-DFT(PBE) study by Liptak and Brunold 173 questioned the open-shell antiferromagnetic contribution to the ground state. As will be shown later in this chapter, the use of a nonhybrid DFT functional is probably
the cause of this discrepancy. Nevertheless, all of those previous studies were carried out for isolated models of cob(I)alamin without the presence of the axial base or inclusion of the enzymatic environment. Therefore, the actual coordination number and electronic structure of the cob(I)alamin intermediate in MetH remain an open subject. During the catalytic methyl transfer reaction (Scheme 2.1), different enzyme domains interact with each other in order to either cleave or generate the Co-C bond. Simultaneously, the His759 residue is expected to move with respect to the Co metal center to break/form the Co-N_{His759} bond by analogy with the solution chemistry of the B12 cofactor.\textsuperscript{112} However, the crystal structure of the Me-cob(III)alamin resting state\textsuperscript{41} (PDB code: 1BMT, 3 Å resolution) shows that His759 interacts with Asp757 and Ser810 through a network of hydrogen bonds, suggesting that His759 would be fixed with respect to the Co center (Figure 2.2). On the other hand, a structural study of the reactivation complex\textsuperscript{113} (Scheme 2.1) formed by the adenosyl–methionine (AdoMet) and cob(I)alamin binding domains shows that the axial His759 moves away from the cobalt center and makes specific contacts with the AdoMet domain. Thus, it is still ambiguous how His759 interacts with the cob(I)alamin cofactor during the catalytic cycle
Figure 2.2. Close view of active site of MeCbl in the enzyme showing the interaction of lower axial ligand His759 with the other residues of the triad (Asp759, Ser810) as well as the nearby Leu806 based on the 1BMT crystal structure.\textsuperscript{41}

The purpose of this chapter is two-fold: First, the coordination of the axial His759 and the cob(I)alamin cofactor inside the enzyme is explored by employing the QM/MM approach. Second, the complex electronic structure of the cob(I)alamin intermediate suggested by the QM/MM results is analyzed using DFT and CASSCF/QDPT2 calculations on gas phase models. Finally, the implications of this finding are discussed in the context of the methyl transfer reaction between the Me-H\textsubscript{4}folate and cob(I)alamin.

\section*{2.2 Computational Details}

Three different theoretical methods were used to study the cob(I)alamin intermediate. First, the formation of the Co(I) state was investigated inside the cobalamin-dependent methionine synthase (MetH) enzyme employing QM(DFT)/MM. Second, the complex electronic properties of the cob(I)alamin and the influence of the axial base were
investigated using gas phase DFT calculations. Finally, the multiconfigurational character of the Im⁻⁻⁻⁻[Cob(I)alamin] was further analyzed using CASSCF/MC-XQDPT2 calculations.

2.2.1 QM(DFT)/MM Calculations. The crystal structure of the MeCbI binding module of MetH (PDB code: 1BMT, at 3 Å resolution) was used to model the cob(I)alamin intermediate, by removing the methyl ligand and adjusting the number of electrons of the QM system, consistent with the Co(I) oxidation state. Initially, the His759 residue is bound to the Co center, since the starting crystal structure corresponds to the Me-cob-(III)alamin resting state. However, upon QM/MM geometry optimization, the axial base is observed to detach from the Co(I) center. The hybrid QM(DFT)/MM calculations were performed using the method developed by Laio et al., which combines the first-principles MD method of Car and Parrinello with a force-field MD methodology (i.e., QM/MM CPMD). In short, the geometry of the cob(I)alamin intermediate was optimized as a closed-shell singlet using the BP86 functional, a plane wave basis set with a 70-90 Ry kinetic energy cutoff, and Martins–Troullier pseudopotentials to describe the interaction between the ionic cores and the valence electrons. Two pseudopotentials were tested for the cobalt atom: one with nine valence electrons supplemented with nonlinear core corrections and another with 17 valence electrons. To explore the possibility of spin polarization between the cobalt and the corrin, a single point calculation within the local spin density approximation (LSD) was also performed. Finally, since the GGA BP86 functional is suspected to have problems describing a possible open-shell singlet state in cob(I)alamin, the B3LYP functional was also tested.
2.2.2 Gas Phase DFT Calculations. To obtain further insight into the complex electronic properties of the cob(I)alamin intermediate suggested by the QM/MM calculations, we carried out gas phase calculations. First, we studied the His-off form of cob(I)alamin for comparison with previous studies.\textsuperscript{80,171-173} Two different gas phase models were used: a big model containing the full cofactor with all of the side chains and a truncated model with C2 symmetry where the side chains have been replaced by hydrogen atoms. Second, we investigated the influence of the axial His759 in the electronic properties of cob(I)alamin. An imidazole (Im) molecule modeling the axial His759 was placed under the Co center in accordance with the orientation obtained from the QM/MM calculations, and a series of displacements along the Co(I)–N\textsubscript{Im} coordinate was generated. Geometry optimization of all models was carried out employing the Becke–Perdew (BP86)\textsuperscript{67,68} functional and the 6-31G(d) (5d components) basis set, as implemented in Gaussian 03.\textsuperscript{175} This level of theory constitutes an appropriate platform for describing the structural and electronic properties analysis of alkyl-cobalt-(III) complexes, as documented in the literature.\textsuperscript{75-77} However, pure GGA functionals such as PBE\textsuperscript{78} or BP86 are known to be unable to converge to a spin polarized solution for Co(I) complexes.\textsuperscript{79,80} This is not the case for hybrid functionals such as B3LYP, despite the fact that this functional underestimates the strength of the Co–C bond.\textsuperscript{79,81} Therefore, we used B3LYP to test the possibility of a spin-polarized solution, which may be indicative of a more complex electronic wave function.\textsuperscript{82} In particular, B3LYP calculations were initially performed on all models having an even number of electrons, assuming a singlet closed-shell wave function. Then, the unrestricted Kohn–Sham formalism using UB3LYP was applied by mixing HOMO and LUMO orbitals to obtain the corresponding
open-shell singlet. For each broken-symmetry solution, we examined the extent of spin polarization between the cobalt and the corrin by analyzing the spin density distributions.

2.2.3 CASSCF/MC-XQDPT2 Calculations. Since the Kohn–Sham formalism (which is the base for DFT-based computations) is restricted to a single Slater determinant description, it cannot describe the multiconfigurational character of the cob(I)alamin system. Thus, we carried out CASSCF multireference calculations, followed by quasi-degenerate perturbation theory (QDPT2) calculations with a multiconfigurational self-consistent-field reference function (MC-XQDPT2) to include the dynamical correlation, as implemented in the PC GAMESS/Firefly QC package. All of the CASSCF calculations were performed on the DFT optimized structures of the cob(I)alamin models (without and with the Im base) using the 6-31G(d) basis set. The details regarding the active space chosen for these multireference calculations are described in section 3.3.

2.3 Result and Discussion

2.3.1 Structure of Cob(I)alamin Inside MetH. MetH is a modular enzyme composed of four functional domains. The domain considered here is the MeCbl binding module, in which the His759 side chain serves as the lower axial ligand (Figure 2.2), instead of the DBI base of the free MeCbl cofactor (Figure 2.1a). During the course of enzymatic methyl transfer, the cobalt center oscillates between the Co(III) and the Co(I) oxidation states, and the axial His is expected to dissociate from the cobalt center by similarity with the behavior of the B_{12} cofactor in solution. However, this assumption remains to be proven since the crystal structure of the cob(I)alamin intermediate is not available. Here, we have investigated the coordination of the axial His759 and the cob(I)alamin in the MetH enzyme by means of QM/MM calculations.
Figure 2.3. QM(DFT)/MM optimized structures of MeCbl binding domain of the MetH (a) hexa-coordinated Me-Cob(III)alamin resting state and (b) Cob(I)alamin intermediate.

The QM/MM optimized structure of the cob(I)alamin:MetH intermediate is shown in Figure 2.3, and the main structural parameters are listed in Table 2.1. The optimized structure of the Me-cob(III)alamin:MetH resting state\textsuperscript{116} is also included for comparison.
Table 2.1. Key Structural Parameters of the Cofactor Binding Domain of MetH in the Me-
Cob(III)alamin Resting State (MeCbl:MetH) and the Cob(I)alamin Intermediate (Co(I):MetH)\(^a\)

<table>
<thead>
<tr>
<th>parameter</th>
<th>MeCbl:MetH</th>
<th>Co(I):MetH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>QM/MM(^{16})</td>
<td>X-ray(^{41})</td>
</tr>
<tr>
<td>Co-C</td>
<td>1.99</td>
<td>1.96</td>
</tr>
<tr>
<td>Co-N(_{\text{Im}})</td>
<td>2.27</td>
<td>2.24</td>
</tr>
<tr>
<td>Co-N1</td>
<td>1.88</td>
<td>1.91</td>
</tr>
<tr>
<td>Co-N2</td>
<td>1.86</td>
<td>1.93</td>
</tr>
<tr>
<td>Co-N3</td>
<td>1.93</td>
<td>2.02</td>
</tr>
<tr>
<td>Co-N4</td>
<td>1.92</td>
<td>2.02</td>
</tr>
</tbody>
</table>

\(^a\) All distances are given in Å. \(^b\) The study in this chapter.

The fact that the X-ray Co–N\(_{\text{ax}}\) bond length for Me-cob(III)alamin:MetH is very well
reproduced at the QM/MM level strengthens our confidence that the demethylated
cob(I)alamin:MetH form would also be well-characterized from a structural point of view
using the same level of theory (i.e., the BP86 functional). The calculations show that the
Co–N\(_{\text{ax}}\)(His759) distance increases from 2.27 Å in the Co(III) oxidation state to 2.78 Å
for Co(I). In other words, the axial ligand is displaced due to the change of the Co
oxidation state. This is in agreement with the X-ray absorption spectroscopic studies
showing that cob(I)alamin in solution is not axially coordinated.\(^{112}\) Nevertheless, it
should be noted that, differently from the cofactor in solution, the Co atom inside the
enzyme remains weakly coordinated to the His. Most likely the axial His in MetH cannot
move away further from the Co because it is hydrogen bonded to Asp757 (Figure 2.2).
Both His759 and Asp757 residues are in a loop, conferring to them a certain degree of
flexibility that allows the axial base to be displaced during the catalytic cycle without
breaking the hydrogen bond between them. However, this flexibility is not unlimited, because Asp757 is also interacting with a serine of an α helix (Ser810) and a leucine of a β sheet (Leu806), and these secondary structures are not as flexible as the loop. As a consequence, the Co–N(His) cannot increase beyond ~2.8 Å without disrupting this hydrogen bond network, something that probably has a high energy cost. Maintaining a weak Co(I) –His coordination may be advantageous from the enzymatic point of view, since it would allow the axial His to recoordinate easily to the cobalt center when the cofactor is re-methylated. The His–Asp–Ser triad is expected to play a key role in the Co(I) re-methylation reaction by modulating the interaction between the Co and the axial base. Interestingly, the His–Asp–Ser triad is conserved in all corrinoid-based proteins catalyzing methyl transfer reactions (except in an iron–sulfur corrinoid protein), and thus it is tempting to suggest that the triad plays a similar role in all of the members of this protein family. Indeed, Hegemeier et al. drew a similar conclusion for another cobalamin dependent enzyme, i.e., the methanol–cobalamin methyltransferase (MetABC), on the basis of the long Co–N distance (2.51 Å) observed in the crystal structure of the cob(I)amide intermediate and the methylation rate of Co(I) being completely dependent on the presence of the axial base.

The structure of the cob(I)alamin:MetH intermediate was initially optimized considering a closed-shell singlet electronic configuration. However, since previous DFT as well as CASSCF calculations have shown that free cob(I)alamin has a more complex electronic structure, we checked the possibility of having a spin polarized solution inside the enzyme by using the local spin density approximation. Despite several attempts, we were unable to converge the open-shell singlet either with BP86 or B3LYP. The obtained
electronic state was found to be indeed an intermediate state between two electronic configurations: the closed-shell Co\textsuperscript{I}(d\textsuperscript{8}) singlet and the open-shell Co\textsuperscript{II}(d\textsuperscript{7})-corrin radical \((\pi^*\textsuperscript{1})\) singlet. The spin density distribution showed some unpaired electronic density on the cobalt and the corrin ring with opposite signs, i.e., some diradical character. Moreover, the total integrated absolute value of the spin density was 1e\textsuperscript{−} (intermediate between the 0 unpaired electrons expected for a pure closed-shell singlet and the two unpaired electrons for an open-shell singlet). This suggests that the closed-shell and the open-shell singlet states are very close in energy, so unless we force the system to be a closed-shell singlet, the single-determinant DFT calculations give an intermediate state between Co(I) and Co(II)-corrin radical. In summary, our QM/MM calculations confirm the early proposal of Wirt et al.\textsuperscript{112} that the axial His ligand oscillates between the His-on form in the Me-cob(III)alamin:MetH resting state (Co–N(His) distance = 2.27 Å) and the His-off form in the cob(I)alamin:MetH intermediate (2.78 Å). Such movement of His\textsubscript{759} in the enzymatic environment is due to the changes in the electronic structure at the metal center, i.e., from Co(III) to an intermediate Co(I)/Co(II) state. Hereafter, the complex electronic structure of the cob(I)alamin intermediate as well as the changes associated with the displacement of the axial base are further analyzed using DFT and CASSCF/QDPT2.

2.3.2 Structural and Electronic Analysis Based on DFT Calculations.

2.3.2.1 His-off Cob(I)alamin Intermediate. Table 2.2 shows the main structural parameters obtained for the gas phase models of the His-off form of the cob(I)alamin intermediate employing both the BP86 and B3LYP functionals. The geometries do not differ significantly from the one inside the MetH enzyme (Table 2.1), regardless of the
model employed for the cofactor (full model with all of the side chains or truncated model with C2 symmetry) or the functional used. In particular, the calculated Co–N_{eq} distances (1.82–1.85 Å) are in agreement with the average experimental value obtained from EXAFS studies (1.86–1.88 Å). As for the electronic structure, the main difference between the two functionals is that only B3LYP gives a spin polarized solution.

Table 2.2. Main Structural Parameters of the Full and Truncated Cob(I)alamin Gas Phase Models

<table>
<thead>
<tr>
<th>Model</th>
<th>Full</th>
<th>Truncated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B3LYP\textsuperscript{b}</td>
<td>BP86\textsuperscript{b}</td>
</tr>
<tr>
<td>fold angle\textsuperscript{79}</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Co-N1</td>
<td>1.84</td>
<td>1.82</td>
</tr>
<tr>
<td>Co-N2</td>
<td>1.85</td>
<td>1.83</td>
</tr>
<tr>
<td>Co-N3</td>
<td>1.91</td>
<td>1.90</td>
</tr>
<tr>
<td>Co-N4</td>
<td>1.90</td>
<td>1.89</td>
</tr>
</tbody>
</table>

\textsuperscript{a} All distances are given in Å, and angles are in deg. N1–N4 refers to the corrin nitrogen atoms. \textsuperscript{b} The study in this chapter.

This may reflect that the nonhybrid functionals such as BP86 do not describe correctly the spin polarization,\textsuperscript{180,181} even though they give correct Co–C bond dissociation energies (BDE).\textsuperscript{79} Figure 2.4 shows the (B3LYP) spin density distribution along with the number of unpaired electrons. There is one unpaired electron on the cobalt atom coupled antiferromagnetically with an unpaired electron on the corrin ring, consistent with a CoII(d\textsuperscript{7})-corrin radical (π\textsuperscript{5}) diradical state. The spin density distributions of the full (Figure 2.4a) and the truncated (Figure 2.4b) models are almost identical. The side chains
do not show any spin density, and the density exhibits 2-fold symmetry. Consequently, the truncated model with C2 symmetry was used for further analysis (Figure 2.4b).

Figure 2.4. Spin density of (a) the full cob(I)alamin model and (b) the truncated cob(I)alamin model with C2 symmetry, computed at the B3LYP/6-31G(d) 5d level of theory. Left, spin populations; right, R and β spin density distributions colored as green and magenta, respectively.

From the energetic point of view, the open-shell singlet was found to be \(~4\) kcal/mol (B3LYP) lower in energy than the corresponding closed-shell configuration. Moreover, the ferromagnetic counterpart of the open-shell singlet, a triplet Co\(^{II}(d^7)\)-radical corrin (\(\pi^*\))^1 state, was also found to be 3 kcal/mol lower in energy than the closed-shell singlet, further validating the diradical contribution to the cob(I)-alamin intermediate. Although
these small energy differences are within the error of the B3LYP calculation, they indicate that the closed-shell and the diradical states are close in energy, thus suggesting that the electron transfer from the Co$^I$(d$^8$) to the corrin(π) to originate the Co$^{II}$(d$^7$)-corrin radical(π$^*$)$^1$ may be feasible (see section 3.3.).

2.3.2.2 Influence of the Axial Base His. The electronic properties of cob(I)alamin were also evaluated in the presence of an imidazole (Im) as a model of the His axial base. Since the His759 moves between 2.27 and 2.78 Å with respect to the Co center inside the enzyme, the Co(I)–N$_{Im}$ distance was systematically varied between 2.1 and 2.8 Å in the gas phase DFT calculations. For each Co–N$_{Im}$ distance, the structure was optimized with both the BP86 and the B3LYP functionals. As for the His-off model (section 3.2.1), the structure does not differ significantly between the two functionals, but a spin polarized solution was only obtained in the case of the B3LYP functional.

The spin density profiles (Figure 2.5 and Figure A1, Appendix) show unpaired spin density on the Co and the corrin ring, with opposite signs, consistent with a Co$^{II}$(d$^7$)-corrin radical (π$^*$)$^1$ diradical state. However, it should be noted that the expectation value of the total spin $\langle S^2 \rangle$ is much larger (1.05–0.75) than the value expected for a singlet, $S(S + 1) = 0$, indicating that the open-shell singlet is significantly contaminated by the triplet state. Therefore, the single-determinantal DFT results need to be interpreted with caution. Nevertheless, we believe that the analysis of $\langle S^2 \rangle$ with respect to the Co–N$_{Im}$ distance (Figure 2.5) could help to assess the extent of the diradical character. The initial $\langle S^2 \rangle \sim 1.05$ value remains constant until the distance reaches $\sim$2.5 Å and decreases up to $\langle S^2 \rangle \sim 0.75$ at longer distances. This could indicate a decrease in the mixing between the
open-shell singlet and triplet states due to an increase in the energy gap between them (see below).

Figure 2.5. Evolution of the expectation value of the total spin, $\langle S^2 \rangle$, along with the spin density distributions during the elongation of the Co–N$_{Im}$ distance in the Im$^-$[Cob(I)alamin] model system.

Interestingly, the change in $\langle S^2 \rangle$ is accompanied by a change in the symmetry of spin density distribution on the cobalt atom. This switch in the d orbital of the cobalt atom bearing the unpaired electron can be further explained by analysis of the energies of relevant molecular orbitals near the HOMO/LUMO gap as a function of the Co–N$_{Im}$ distance (Figure 2.6). The antiferromagnetic coupling occurs between the upper occupied orbital of the corrin ($\pi^*_{corr}$) and the singly occupied d orbital of the cobalt (d$_{Co}$). This d$_{Co}$ orbital is the d$_{z^2}$ orbital at short Co–N$_{Im}$ distances, but it is the d$_{yz}$ orbital at long
distances, in agreement with the \( d_{yz} \rightarrow \pi^* \) MLCT band observed in a previous study of the cob(I)alamin intermediate without the axial base.\(^{80}\) Moreover, the energy gap between the \( \pi^*_{\text{corr}} \) and the \( d_{Co} \) orbitals increases with the Co–N\(_{Im} \) distance, further supporting our previous suggestion that the energy gap between the open-shell singlet and the triplet states increases with the Co–N\(_{Im} \) distance. In summary, the DFT calculations suggest that the cob-(I)alamin intermediate has significant diradical character. They also indicate that the axial base modulates the nature of the coupling associated with Co\(^{II}(d^7)\)-corrin radical (\( \pi^* \))\(^1\) configuration by changing the energy of the d orbitals of the cobalt atom. However, the multiconfigurational character of the cob(I)alamin can only be assessed correctly using a multireference method.

Figure 2.6. Variation of HOMO and HOMO–1 based orbital energies with the Co–N\(_{Im} \) distance calculated at the UB3LYP level of theory.


2.3.3 CASSCF/MC-XQDPT2 Analysis. The multiconfigurational nature of the wave function of cob(I)alamin was further investigated using multireference methods (CASSCF and CASSCF/MC-QDPT2). The selection of active orbitals for complex systems such as cob(I)alamin always represents a challenging problem. Using previously reported CASSCF calculations on corrin or corrol based complexes as a guide, we have examined several kinds of active spaces and finally chosen to use a space of 10 active electrons distributed in 11 active orbitals.

Initially, CASSCF(10,11) calculations were carried out for the His-off form of cob(I)alamin (truncated model with C2 symmetry, see Section 3.2.1.) in order to reproduce the weight of the diradical contribution reported by Jensen. The active space, as shown in Figure A2 (Appendix), is very similar to that reported in ref. Second, we probed the effect exerted by the axial ligand on the electronic structure of the Im--[CoI-(corrin)] complex by performing a series of different CASSCF calculations with the same active space (Figures 2.7) but varying the Co–NIm distance from 2.1 Å to 2.8 Å. In particular, the active space for Im--[CoI(corrin)] is comprised of the d\(x^2-y^2\)-N(\(\sigma^*\)), d\(xy\), d\(xz\), d\(yz\), and 3d\(z^2\) + \(\pi^*\) orbitals, and the respective correlating orbitals are \(\pi^* + 3d_{x^2-y^2}\), 4d\(z^2\), d\(x^2-y^2\)+ N(\(\sigma^*\)), 4d\(xy\), 4d\(xz\), and 4d\(yz\) (Figure 2.7). Note that this active space includes not only cobalt orbitals but also the corrin orbitals that may be important for the charge transfer between the corrin ring and the metal. The pair of correlating orbitals numbered 109 and 116 describe a \(\sigma^*\) donation from the lone electron pairs of the equatorial nitrogen atoms to the d\(x^2-y^2\) orbital of cobalt. In addition, we have included the lowest unoccupied \(\pi^*\) orbital of corrin, because it has been shown to be necessary to describe the electron configuration of the His-off form of cob(I)alamin. This orbital mixes with the 3d\(z^2\)
cobalt orbital at a short Co–N\textsubscript{Im} distance (CASSCF orbitals numbered 113 and 114, see Figure 2.7), but with the 3d\textsubscript{yz} orbital at long distances (CASSCF orbitals numbered 112 and 114). Finally, an extra orbital has been added to the correlating orbitals, in order to include the double shell effect\textsuperscript{185} to all d orbitals.

Figure 2.7. CASSCF active space orbitals used in the calculations of the Im\textsuperscript{−}[Cob(I)alamin] model system at a Co–N\textsubscript{Im} distance = 2.2 Å.

The calculations reveal that the overall wave function of cob(I)alamin has multireference character, comprising five major configuration state functions (CSFs) that include single and double excitations (Figure A3, Appendix). By means of a unitary transformation to
localized orbitals, this complex wave function can be simplified to four dominant configuration state functions (CSFs), as plotted in Figure 2.8.

![Figure 2.8](image)

Figure 2.8. Weight (%) of the major configurations (CSFs) contributing to the ground state CASSCF wave function as a function of the Co–N$_{Im}$ distance for the Im$^+\text{[Cob(I)alamin]}$ model system. The description of the CSFs is given in Table 2.3.

CSF1 is the closed-shell singlet, corresponding to the Co$^I$(d$^8$)-corrin ($\pi^*$)$^0$ configuration. The open-shell singlet, representing the Co$^{II}$$(d^7)$-corrin radical ($\pi^*$)$^1$ configuration, consists of two major configurations (CSF2 and CSF3), both corresponding to charge transfer states where one electron from a d orbital of Co (d$_{z^2}$ or d$_{yz}$, respectively) has been shifted to the corrin $\pi^*$ orbital. In addition, there is a small contribution of the double metal-to-ligand excitation, i.e., Co$^{III}$$(d^6)$-corrin anion($\pi^*$)$^2$ (CSF4). Interestingly, the weight of these different CSFs varies with the Co–N$_{Im}$ distance. The weight of the closed shell Co$^I$(d$^8$) configuration increases with the Co–N$_{Im}$ distance, as the overall diradical character decreases. However, the behavior of the two-diradical
Table 2.3. Composition of the CASSCF Wave Function of Cob(I)alamin for Each Co–N₃m Distance, in Terms of the Mulliken Occupation Numbers of the Three Natural Orbitals Involved in the Open-Shell Singlet Description and the Weights of the Major Configurations State Functions (CSF1, CSF2, CSF3, and CSF4)

<table>
<thead>
<tr>
<th>Co–N₃m (Å)</th>
<th>orbital occupation⁴</th>
<th>CSF1</th>
<th>CSF2</th>
<th>CSF3</th>
<th>CSF4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(dₓz)²(π*)⁰</td>
<td>(dₓz)¹(π*)¹</td>
<td>(dₓz)¹(π*)¹</td>
<td>(dₓz)²(π*)⁰</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(weight %)</td>
<td>(weight %)</td>
<td>(weight %)</td>
<td>(weight %)</td>
</tr>
<tr>
<td>2.1</td>
<td>.... (dₓz)¹.⁹⁷(dₓz)¹.¹⁰(π*)⁰⁸⁴ ....</td>
<td>6</td>
<td>80</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>2.2</td>
<td>.... (dₓz)¹.⁹⁷(dₓz)¹.¹²(π*)⁰⁷₉ ....</td>
<td>10</td>
<td>76</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td>2.3</td>
<td>.... (dₓz)¹.⁹⁷(dₓz)¹.¹²(π*)⁰₇₆ ....</td>
<td>11</td>
<td>72</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td>2.4</td>
<td>.... (dₓz)¹.₉₆(dₓz)¹.₃₄(π*)⁰₆₆ ....</td>
<td>19</td>
<td>60</td>
<td>8</td>
<td>-</td>
</tr>
<tr>
<td>2.5</td>
<td>.... (dₓz)¹.₉₀(dₓz)¹.₅₀(π*)⁰₅₇ ....</td>
<td>35</td>
<td>36</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>2.6</td>
<td>.... (dₓz)¹.₅₀(dₓz)¹.₉₄(π*)⁰₅₁ ....</td>
<td>51</td>
<td>13</td>
<td>21</td>
<td>4</td>
</tr>
<tr>
<td>2.7</td>
<td>.... (dₓz)¹.₅₉(dₓz)¹.₉₅(π*)⁰₄₂ ....</td>
<td>57</td>
<td>4</td>
<td>24</td>
<td>5</td>
</tr>
<tr>
<td>2.8</td>
<td>.... (dₓz)¹.₅₉(dₓz)¹.₉₅(π*)⁰₄₂ ....</td>
<td>59</td>
<td>1</td>
<td>24</td>
<td>5</td>
</tr>
</tbody>
</table>

⁴ The Mulliken population analysis is shown for three natural orbitals involved in the open-shell singlet state. The rest of the active space orbitals are represented by dots.

Contributions is different. The weight of CSF2 decreases abruptly for Co–N₃m distances > 2.5 Å, whereas the weight of CSF3 increases. The crossing between the two diradical configurations takes place at ~2.6 Å and can be explored by analyzing the Mulliken populations (Table 2.3) and the pure fragment orbitals obtained after localization (Figure 2.9).
Figure 2.9. Pure fragment localized orbitals, describing the electron transfer between the cobalt and the corrin. a) Co–N Im distance < 2.5 Å; the electron is transferred from Co(d^2_z) to corrin(\pi^*). b) Co–N Im distance > 2.5 Å; the electron is shifted from Co(d_{yz}) to corrin(\pi^*).

When the Co–N Im distance is shorter than 2.5 Å, the electron is transferring from the d^2\_z cobalt orbital to the corrin \pi^* orbital, whereas when it is longer than 2.5 Å the electron is shifted from d_{yz} to \pi^*. Indeed, the localized orbitals are consistent with the orbitals resulting from adding and subtracting the original HOMO (or HOMO–1 for longer Co–N Im distances) and LUMO orbitals for each Co–N Im distance:

\[ \text{Co–N Im} < 2.5 \text{ Å} \quad \psi_{113} + \psi_{114} \rightarrow d^2_z \quad \text{and} \quad \psi_{113} - \psi_{114} \rightarrow \pi^* \]

\[ \text{Co–N Im} > 2.5 \text{ Å} \quad \psi_{112} + \psi_{114} \rightarrow d_{yz} \quad \text{and} \quad \psi_{112} - \psi_{114} \rightarrow \pi^* \]

This switch in the cobalt d orbitals with the change in the axial bond length is consistent with the spin-polarized results obtained from DFT(UB3LYP) calculations. The availability of different cobalt d orbitals in the active space chosen for the calculations
allows the system to change the d<sub>co</sub> orbital that participates in the diradical state during the elongation of the Co–N<sub>Im</sub> distance. It should also be noted that the overlap between the π* corrin orbital and the singly occupied cobalt d orbital is larger for d<sub>yz</sub> than for d<sub>x^2</sub>, resulting in a higher weight for the CSF3 configuration than for CSF2, although the sum of both contributions reduces significantly with increasing Co–N<sub>Im</sub> distance. This is in line with the increasing closed-shell character as well as the appearance of the small contribution from the double excitation (CSF4).

In addition, state average CASSCF/MC-XQDPT2 calculations also reveal that each excited state has a significant diradical character, showing metal-to-ligand charge transfer (Table A1, Appendix). Thus, at Co–N<sub>Im</sub> = 2.1 Å, the correlated wave function of the ground state is mainly composed of the open-shell singlet, Co<sup>II</sup>(d<sup>7</sup>)-corrin radical (π*)<sup>1</sup> configuration. This diradical weight decreases as the axial bond is elongated, such that at Co–N<sub>Im</sub> = 2.8 Å the diradical weight is only 24%, similar to the result obtained by Jensen31 for the cob(I)alamin model without an axial base. In other words, the dominant contribution to the ground state wave function is the diradical configuration when the His is bound to the cobalt atom, but the closed-shell singlet, Co<sup>I</sup>(d<sup>8</sup>) configuration when the axial His is weakly coordinated.

**2.3.4 Energy Changes as Function of Co–N<sub>Im</sub> Distance.** Finally, we have calculated the energy of the Im<sup>−</sup>[CoI(corrin)] complex as a function of the distance between Co and N<sub>Im</sub> atoms in order to evaluate the energy cost of displacing the axial ligand (Figure 2.10).
Figure 2.10. Change in the ground state energy with the Co–N\textsubscript{Im} distance, computed with different DFT functionals and ab initio CASSCF/MCXQDPT2 methods.

We have employed single-determinant (DFT with either the BP86 or the B3LYP functional) and multireference (CASSCF and CASSCF/MC-XQDPT2) methods and different basis sets (6-31G(d) and 6-311G(d,p)). Regardless of the computational method used, the curve is clearly repulsive. However, it should be noted that the (CASSCF/MC-XQDPT2) energy at a Co–N\textsubscript{Im} distance ~ 2.3 Å (i.e., in the Me-Cob-(III)alamin resting state of the MetH enzyme) is only 4 kcal/mol higher than at ~2.8 Å (i.e., the calculated value for the cob-(I)alamin intermediate, see section 3.1). Therefore, although the change of distance causes noticeable changes in terms of electronic properties of the Co center (from Co\textsuperscript{II}-corrin radical to Co\textsuperscript{I}), it is energetically not costly.

2.3.5 Implications for the Re-methylation of the Cob-(I)alamin Intermediate in MetH. The methyl transfer reaction from Me-H\textsubscript{4}folate to cob(I)alamin to generate the
Me-cob-(III)alamin resting state (Scheme 2.1) is generally assumed to occur through a $S_{N2}$-type displacement. However, in view of the present study, it can be suggested that (i) a radical mechanism, consisting of an electron transfer (ET) followed by a methyl radical transfer, is also possible and that (ii) the axial base would have a decisive role regarding the enhancement of the ET-based mechanism by modulating the cob(I)alamin electronic properties. At shorter Co–N$_{Im}$ distances (< 2.5 Å), the dominant electronic configuration is diradical, with a Co(II) and an electron located on the corrin ring. This would favor the radical mechanism, in which an electron is initially transferred from the corrin to the Me-H$_4$folate in order to generate a Co(II) and [Me-H$_4$folate]$^\cdot$ state. Subsequently, a methyl radical would be transferred from [Me-H$_4$folate]$^\cdot$ to Co(II), forming Me-Cob-(III)alamin. This methyl radical transfer would be in line with the reductive cleavage mechanism proposed for the methyl transfer from MeCbl to the Hcy substrate (Scheme 2.1), where the electron is transferred from the Hcy to the MeCbl, followed by a methyl radical transfer from the one electron reduced form of MeCbl to the Hcy substrate. In other words, the reaction mechanism of MetH would not only involve the Co metal but also the corrin ligand, which can activate both substrates (Me-H$_4$folate and Hcy) of the MetH enzyme by ET. On the other hand, at longer Co–N$_{Im}$ distances (>2.5 Å), the dominant configuration is closed-shell, with an electron pair located on the metal center consistent with a Co(I) oxidation state. This would favor an $S_{N2}$-type mechanism in which the Co(I) nucleophile directly abstracts the methyl group of Me-H$_4$folate.

### 2.4. Conclusion

In this chapter, the electronic and structural properties of the cob(I)alamin intermediate have been analyzed using QM(DFT)/MM, gas phase DFT, and CASSCF/QDPT2
calculations. Because previous studies\textsuperscript{80,171-173} did not take into account the influence of the protein environment, we initially performed QM/MM calculations to study the formation of the Co(I) state inside the MeCbl domain of the methionine synthase (MetH) enzyme. The observed displacement of the axial His759 from the cobalt center is in agreement with the model proposed by Wirt et al.\textsuperscript{112} based on the tetracoordinated state of free cob(I)alamin. However, His759 is weakly coordinated inside the enzyme (QM/MM optimized Co–N(His) distance = 2.78 Å), where the presence of the catalytic triad (His759–Asp757–Ser810) and the hydrogen bonds with other residues reduce its conformational freedom. This implies that the re-methylation of the cob(I)alamin cofactor by Me-H\textsubscript{4}folate (Scheme 2.1) occurs in the presence of the axial ligand, and thus His759 may have an influence in this methyl transfer reaction. Consequently, we have analyzed the influence of the axial ligand in the electronic structure of cob(I)alamin cofactor using gas phase DFT followed by CASSCF/QDPT2 calculations. Irrespective of the theoretical method used, our results show that the ground state of cob(I)alamin is multiconfigurational, in agreement with a previous study of the cob(I)alamin cofactor without the axial ligand.\textsuperscript{80} In addition to the closed-shell Co(I), a diradical Co(II)-corrin radical configuration (formed by electron transfer from the cobalt to the corrin ring) contributes to the electronic structure of the cob(I)alamin intermediate, revealing the non-innocent behavior of the corrin ring.\textsuperscript{174} The weight of these two configurations depends on the distance of the axial base His from the Co center. The main contribution to the ground state wave function at short Co–N\textsubscript{Im} distances is the diradical configuration, whereas at long distances, it is the closed-shell. Therefore, our results suggest that (i) the standard description of the Co(I) nucleophile is not appropriate
for cob(I)alamin, due to the non-innocent character of the corrin ring, and (ii) the distance between cob(I)alamin and the axial His plays an important role in modulating the nucleophilicity of Co(I).

In view of our results, we have proposed that the re-methylation reaction in MetH (Scheme 2.1) could involve not only the metal (i.e., the closed-shell Co(I) configuration) but also the corrin ring (i.e., the diradical Co(II)-corrin radical configuration). In other words, in addition to the traditionally assumed $S_N2$ mechanism, the findings of this chapter suggest the possibility of an alternative radical mechanism, in which an electron is transferred from the corrin to the Me-H$_4$folate in order to generate Co(II) and [Me-H$_4$folate]$^-$. It should be noted that this ET does not require the presence of any strong reducing agent near the CH$_3$–H$_4$folate substrate but rather the cofactor-induced formation of anion-radical-like species within the cob-(I)alamin:[Me-H$_4$folate] reactant complex. Indeed, earlier studies by Marcus$^{186}$, Shaik et al.$^{187}$ and Zipse$^{188}$ already discussed that such ET bond-breakage mechanisms significantly enhance the reaction rates in comparison to $S_N2$ mechanisms. The energetic and dichotomy of the $S_N2$ and radical mechanisms in the methyl transfer reaction from Me-H$_4$folate to cob(I)alamin are currently being investigated in our group.
CHAPTER 3

MECHANISTIC INSIGHTS FOR THE METHYLATION OF THE COB(I)ALAMIN FROM CH$_3$-H$_4$FOLATE CATALYZED BY METHIONINE SYNTHASE

3.1 Introduction

B$_{12}$ (cobalamin)-dependent methyltransferases play a critical role in one-carbon metabolism and CO$_2$ fixation in anaerobic microbes, as well as in amino acid metabolism in several organisms including mammals and bacteria.$^{48,189-191}$ The corrinoid-dependent methyltransferases catalyze the transfer of a methyl group from a methyl donor to a methyl acceptor in multi domains or/and multi proteins. However, the most important and less understood reaction is the transfer of a methyl group from a methyltetrahydrofolate (CH$_3$-H$_4$Folate or CH$_3$-THF) substrate to a cob(I)alamin/cob(I)amide form of the cofactor. This particular reaction is common in a wide range of methyltransferases, namely, methionine synthase (MetH) from *E. coli*,$^{25-28,31,95}$ methyltetrahydrofolate-corrinoid-iron/sulfur protein (CFeSP) AcsE from *C. thermoaceticum*,$^{103-109}$ and bacterial mercury methylation HgcAB from *D. desulfuricans ND132*. $^{192}$ Among all the known corrinoid proteins, the cobalamin dependent MetH (Figure 3.1) is one of the best studied enzyme that catalyzes the transfer of a methyl group from CH$_3$-H$_4$Folate to homocysteine(Hcy) to form a methionine, in which cobalamin cofactor serves as an
intermediate carrier oscillating between methylcobalamin (MeCbl) and cob(I)alamin. \(^{170, 193, 194}\) MetH is a single polypeptide (136 kDa and 1227 amino acids) and has a modular architecture with four different binding modules including a CH\(_3\)-H\(_4\)Folate-binding donor domain, an Hcy-binding acceptor domain and a MeCbl-binding domain, which are required for the main catalytic cycle and the fourth one adenosylmethionine (AdoMet) domain is essential for the reactivation cycle (Figure 3.1).\(^{40-43}\) During the course of catalytic turnover, each substrate-binding domain must interact with the \(\beta\)-face of the enzyme-bound cobalamin prosthetic group, which requires various conformational changes to position the donor and acceptor domains in alternate manner.\(^{41}\) However, the crystal structure of a full length of MetH enzyme has not been solved yet due to the high degree of conformational flexibility but the crystal structure of each individual domain has been successfully characterized.\(^{39, 41, 99, 100}\)

The methyl transfer reactions to and from the cobalamin are extremely challenging and the way in which enzyme activates the methyl donors such as MeCbl and the CH\(_3\)-H\(_4\)Folate through conformational changes remains ambiguous. In the first-half catalytic reaction (Figure 3.1), the cob(I)alamin is methylated by CH\(_3\)-H\(_4\)Folate, thus forming MeCbl and tetrahydrofolate (H\(_4\)Folate or THF). During the second-half catalytic reaction, the resulting MeCbl form of the cofactor is demethylated by Hcy substrate forming cob(I)alamin and methionine (Figure 3.1). The key step in the MetH catalytic reaction is the formation of organometallic carbon-cobalt (CH\(_3\)-Co) bond, i.e., Me-Cob(III)alamin cofactor where the CH\(_3\)-H\(_4\)Folate donates a methyl group to exogenous cob(I)alamin, however exactly how it is formed in the enzymatic reaction remains an open question. Thus, the investigation of the Co-C bond formation in the reaction
complex (cob(I)alamin:CH₃-H₄Folate) is important in order to understand the role of the MetH in catalysis. Moreover, it has been estimated that the methyl transfer from enzyme-bound CH₃-H₄Folate to cob(I)alamin is accelerated by 35-million-fold as compare to the model reaction. Therefore, the challenge for chemists is to determine the enzymatic mechanism(s) and the source of activation of an unreactive methyl donor CH₃-H₄Folate by modeling the intermodular methyl transfer reactions.

Figure 3.1. The catalytic cycle and reactivation cycle of methionine synthase (MetH): [I] the first-half catalytic reaction, [II] the second-half catalytic reaction, and [III] the reactivation reaction.

On the other hand, the displacement of a methyl group (as cation/radical) attached to a tertiary amine (i.e., CH₃-H₄Folate) is not a facile reaction in bioorganic chemistry even in the presence of strong nucleophile such as cob(I)alamin because the H₄Folate
anion is a very poor leaving group. It has been successfully demonstrated by Pratt and his coworkers that 5,5-dimethyltetrahydropterin (structural analogue of protonated CH$_3$-H$_4$Folate) can serve as a methyl donor to cob(I)alamin. Therefore, the electrophilic activation of the CH$_3$-H$_4$Folate is required for the methyl transfer reaction, most probably by protonation at N5 position that will form an aminium complex. In addition, Smith and Matthews have revealed that the CH$_3$-H$_4$Folate protonate the N5 position but not the conjugated carbon next to it using $^{13}$C distortionless enhancement by polarization transfer (DEPT) NMR spectroscopy. Thus, it is of great interest to understand the source of a protonation at N5 position of the CH$_3$-H$_4$Folate in order to investigate the methyl transfer reaction.

The X-ray structure of CH$_3$-H$_4$Folate-binding domain of the MetH (PDB ID: 1Q8J at 1.9 Å resolution, Figure 3.2) show that the pterin ring of the CH$_3$-H$_4$Folate is interacting with the hydrophilic protein residues such as Asn508, Asn473 and Asp390 via H-bond network, indicating the polar environment might be involved to increase the pK$_a$ at N5 position. Most importantly, Asn508 interacts from 2.8 Å distance with the N5 of the CH$_3$-H$_4$Folate substrate and this H-bond interaction could be the key for the stabilization of protonation state. Interestingly, the active site residues near the pterin ring of the CH$_3$-H$_4$Folate are conserved both in MetH and corrinoid/Fe-S protein (CFeSP) including the Asparagine (Asn508 in PDB id: 1Q8J and Asn199 in PDB id: 2E7F, respectively) which is H-bonded to the N5 position. However, it has been shown that H$_4$Folate which is the product of the methyl transfer reaction contains a proton at N5 position. Thus, the most critical and somehow unanswered question, that arise here is the timing of the proton uptake with respect to methyl group transfer; nonetheless, the
crystal structure of the CH$_3$-H$_4$Folate binding domain of the MetH shows no general acid (proton donor, BH) other than a H-bond network with the Asn508 (Figure 3.2), indicating that an unprotonated CH$_3$-H$_4$Folate bound by the protein.\textsuperscript{193} Moreover, it has been further postulated by Matthews et al. that N5 atom of the CH$_3$-H$_4$Folate cannot be protonated in the binary CH$_3$-H$_4$Folate:MetH complex. Instead, the protonation event occurs in a later stage, i.e., in the ternary MetH:Cob(I)alamin.CH$_3$-H$_4$Folate complex.\textsuperscript{111} In the case of corrinoid-Fe/S protein AcsE, the kinetic studies have shown that the CH$_3$-H$_4$Folate bind to the MeTr enzyme in the unprotonated form and then undergoes rapid protonation, thus protonation occur in the binary complex.\textsuperscript{104,196}

\textbf{Figure 3.2.} The active site of CH$_3$-H$_4$Folate substrate (PDB ID: 1Q8J at 1.9 Å resolution) showing the polar environment near the pterin ring of the CH$_3$-H$_4$Folate. Here, the pterin ring is positioned by hydrogen bonds with four important protein residues: Asn508, Asp473, Asp390, and Asn411 and out of these, the N and O of the Asn508 are directly interacting via H-bond with the N5 of pterin ring.
Although several mechanistic pathways have been suggested for the methyl transfer reactions catalyzed by MetH, the exact route responsible for these reactions remains the subject of ongoing discussion. It is generally believed that the methyl transfer reaction from the CH₃-H₄Folate to the cob(I)alamin proceed via an S_N₂-type nucleophilic displacement, nevertheless, alternative pathway including single electron transfer (SET) as well as oxidative addition (reductive elimination in a second-half reaction) have also been proposed by Matthews et al. The major limitation of the oxidative addition mechanism is a three-centered bond requirement between the CMe-N5 bond of the CH₃-H₄Folate and the Co metal of cob(I)alamin complex, which might not be possible from enzymatic point of view due to steric interactions in the protein backbone of the two reacting modules. In addition, the S_N₂-type mechanism was observed in the model complexes using DFT calculations and the activation energy of the transition state was found to be 13 kcal/mol. However, this model study lacks the crucial role of the enzyme, especially the cap subdomain that packs over the β-face of the cofactor, which could be involved in stabilizing the CH₃-H₄Folate substrate and subsequently would help in facilitating methyl transfer reaction. Despite the extensive studies of the methyltransferases, the crucial role of enzyme-bound cob(I)alamin (containing cap and Rossman subdomain) for the methyl transfer reaction from CH₃-H₄Folate remains unclear.

In addition, the feasibility of an ET-mediated reductive Co-C bond cleavage has also been explored in the second-half catalytic reaction of the MetH using QM/MM and multireference CASSCF/MC-XQDPT2 computations. According to the reductive cleavage mechanism, the electrochemical reduction from an unprotonated Hcy
(methylthiolate) to the MeCbl cofactor induces the homolytic cleavage of the Co-C bond, followed by the methyl radical transfer. However, there is no structural information available with regard to the reaction complex in both half-catalytic reactions of MetH, therefore, the details of the reaction mechanism including the methylation of the cob(I)alamin intermediate by CH$_3$-H$_4$Folate substrate is not clearly understood.

In the computational study performed in this chapter, we applied ONIOM-based quantum mechanics/molecular mechanics (QM/MM) calculations to investigate the E.S reaction complex, which was prepared by docking a CH$_3$-H$_4$Folate substrate in the β-face of the enzyme-bound cob(I)alamin. Based on the electronic structure and coordination of the cob(I)alamin intermediate in the cob(I)alamin:CH$_3$-H$_4$Folate reaction complex, two different mechanistic pathways are explored: $S_N$2-type nucleophillic displacement and ET-based radical mechanism where the electron is transferred from His-on cob(I)alamin to pterin ring of the CH$_3$-H$_4$Folate followed by methyl radical. Furthermore, the results of the study in this chapter including mechanistic details and the protonation process are discussed that can be applied to a broad class of corrinoid-methyltransferases.

3.2 Computational Methods

3.2.1 System Preparation and ONIOM-based QM/MM Calculations. To investigate the electronic structure and mechanistic details of the cob(I)alamin:CH$_3$-H$_4$Folate reaction complex, ONIOM-based QM/MM analysis was first carried out for the enzyme-bound cob(I)alamin intermediate. Since there is no X-ray crystal structure available for enzyme-bound cob(I)alamin intermediate because of its high reactivity and
supernucleophilic character, we used the crystal structure of the enzyme-bound MeCbl. The initial structure of the MeCbl-binding domain of the MetH was obtained from the Protein Data Bank (PDB ID: 3BMT at 3.0 Å resolution)\(^{41}\) which were then used to build a computational model. At first, the experimental artifacts were removed from the crystal structure and therefore, no counterions were included in the structure model. Subsequently, the H-atoms were incorporated using VMD\(^{58}\) and GaussView\(^{198}\) assuming the normal protonation state of all titrable residues except histidine. Thus, the protonation state of histidine residue that can be protonated at $\delta$ or/and $\varepsilon$ was determined based on the local H-bonding network via visual inspection and using PROPOKA suite of program.\(^{60,61}\) The lower axial His759 residue was treated as a HID in which $N_\delta$ atom was protonated. Consequently, AMBER parameters were assigned for the $B_{12}$ cofactor as developed by Marques et al.\(^{62}\) The model structure was then subjected to geometry optimization of all the hydrogens while keeping the heavy atoms fixed using the AMBER force field as implemented in Gaussian09 program.\(^{63}\) Subsequently, AMBER\(^{64}\) optimization was performed on the side chains of the protein residues except lower axial HIS759 residue. Furthermore, we carried out ONIOM (BP86/6-31G(d):AMBER)\(^{64-68}\) Mechanical Embedding (ME) single-point calculations to determine Merz–Kollman\(^{69}\) electrostatic potential (ESP) atomic charges for the Quantum Mechanical (QM) region (using the $+3$ oxidation state of the cobalt). In this ONIOM-ME calculation, the cofactor was used as a part of the QM model system and the radius of the Co was set at 0.79 Å. After the preparatory calculations, the entire $B_{12}$ cofactor, as well as His759 was included in the QM region (contains 194 atoms) and His759 was capped as a H-link atom. The protein residues within 20 Å of the Co center was used for geometry optimization first by
the ONIOM-ME method followed by ONIOM-electronic embedding (EE) formalism (Figure S2). The overall geometrical parameters obtained from the QM(BP86)/MM-EE geometry optimization of the MeCbl-binding domain of MetH agree well with the experiments X-ray structure (Table 3.1, column 2 and 3).

Furthermore, we removed the methyl group of MeCbl and adjusted the number of electron of the QM system consistent with the Co(I) oxidation state to prepare the structure of enzyme-bound cob(I)alamin intermediate and then optimized the resulting structure with QM/MM calculations. Upon geometry optimization, the α-axial His759 ligand was found to be detached from the Co(I) center (See Figure S3) as expected. The calculated distance between the Co and N of the His759 was 3.02 Å, which clearly indicated the Co(I) is four coordinated inside the enzyme.  

Table 3.1. Key Structural Parameters of the Optimized Systems of MeCbl-MetH Binding Domain and protonated MetH-cobalamin:CH$_3$-H$_4$Folate Reaction Complex (RC$_{N5H}$) in Comparison with the Experimental Data $^a$

<table>
<thead>
<tr>
<th></th>
<th>MeCbl-MetH</th>
<th>Co(I)-MetH</th>
<th>Co(I):[CH$_3$-H$<em>4$Folate] (RC$</em>{N5H}$(css))</th>
<th>Co(II):[CH$_3$-H$<em>4$Folate]$^*$ (RC$</em>{N5H}$(T/oss))</th>
<th>Me-Co(III) :H$<em>4$Folate (PC$</em>{N5H}$(css))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>calcd</td>
<td>X-ray$^{41}$</td>
<td>calcd</td>
<td>calcd</td>
<td>Calcd</td>
</tr>
<tr>
<td>Co-C</td>
<td>1.97</td>
<td>1.96</td>
<td>3.17</td>
<td>3.02</td>
<td>2.07</td>
</tr>
<tr>
<td>Co-N$_{ax}$(His)</td>
<td>2.14</td>
<td>2.24</td>
<td>1.82-1.90</td>
<td>1.82-1.91</td>
<td>1.87-1.93</td>
</tr>
<tr>
<td>Co-Nc</td>
<td>1.87-1.95</td>
<td>1.91-2.02</td>
<td>1.37-1.51</td>
<td>1.33-1.50</td>
<td>1.32-1.51</td>
</tr>
<tr>
<td>Nc-C</td>
<td>1.32-1.51</td>
<td>1.23-1.44</td>
<td>1.56</td>
<td>1.55</td>
<td>1.56</td>
</tr>
<tr>
<td>C1-C2</td>
<td>1.56</td>
<td>1.46</td>
<td>73.9-103.6</td>
<td>75.1-100.13</td>
<td>86.9-98.4</td>
</tr>
<tr>
<td>&lt;C-Co-N$_{ax}$</td>
<td>84.5-95.0</td>
<td>83.4-99.5</td>
<td>88.6-91.3</td>
<td>82.1-90.8</td>
<td>88.3-92.7</td>
</tr>
</tbody>
</table>
Distances and angles are given in angstroms and degrees, respectively. T and oss stands for triplet and open shell singlet state, respectively.

3.2.2 Enzyme-Substrate Reaction Complex. In order to analyze the structure of the cob(I)alamin:CH$_3$-H$_4$Folate reaction complex we docked the CH$_3$-H$_4$Folate substrate on the β-face of the cofactor in the cob(I)alamin-binding module using Maestero graphical user interface from Schrödinger suite of program. Herein we used the truncated CH$_3$-H$_4$Folate where the long side chain of the CH$_3$-H$_4$Folate was replaced with a methyl group to decrease the size of the QM model. We also placed a proton at the N5 position of the CH$_3$-H$_4$Folate substrate which is referred to as a N5H protonation state in this article. Subsequently, Merz-kollman$^{69}$ electrostatic potential atomic charges were determined for the CH$_3$-H$_4$Folate substrate (first stationary point of each model and then same charges were used in all optimizations). In the enzyme-substrate reaction complex with N5H protonation state, the total number of QM atoms was 218 including the entire cofactor, lower axial His759 ligand and protonated CH$_3$-H$_4$Folate (Figure A4, Appendix).

The steric interaction between protonated CH$_3$-H$_4$Folate substrate and protein residues of the Cap subdomain (i.e., β-face of the cofactor) of corrinoid protein would be minimized during the QM/MM-ME geometry optimization. To attain more reasonable structure, we used a number of different starting conformations of the CH$_3$-H$_4$Folate relative to the cob(I)alamin intermediate. We found that, after the QM(BP86/6-31G(d)/MM(AMBER)-ME optimization, all the conformations converge to a point where the orientation of methyl group is pointing towards the cob(I)alamin (Figure S1) and energy of all these conformations were almost comparable. In all the conformations, the distance between the Co and the C$_{Me}$ of CH$_3$-H$_4$Folate was never less than 3.6 Å. In addition, the propanamide side chain of the cofactor interacts with N3H of the pterin ring and this
intermolecular interactions act as an anchor, which would stabilize the pterin ring in the reaction complex. Therefore, we selected the final reaction model complex which has Co-C distance 3.59 Å that is shown in Figure 3.3a (hydrophobic surface is shown in Figure S5). All the resulting structures and transition states were optimized by QM/MM-ME in which the QM and MM parts were treated by the BP86/6-31G(d) method and AMBER96 force field, respectively. The QM/MM-ME vibrational frequency calculation at the same level of theory was further performed to characterize and verify the key transition states and intermediates. Since the ME scheme does not include polarization of MM region into a QM Hamiltonian, the additional single point energy calculations using QM/MM-EE scheme were further carried out based on the QM/MM-ME optimized geometry.

3.3 Results and Discussion

3.3.1 Chemical Models and Enzyme-Substrate Reaction Complex. The MetH enzyme is a modular enzyme that is composed of four different functional domains. During the course of catalytic methyl transfer, both substrate-binding domains (Hcy and CH$_3$-H$_4$Folate) alternatively come closer to cobalamin-binding domain through conformational changes and form a reaction complex to the upper(β)-face of the cofactor in which cobalt center cycles between the Co(III) and the Co(I) oxidation states. Due to high conformational motion of enzyme, there is no X-ray crystal structure available for the enzyme-substrate reaction complex enclosing the cob(I)alamin form of the cofactor and the CH$_3$-H$_4$Folate substrate. Therefore, we applied our chemical intuition to computationally study the cob(I)alamin:CH$_3$-H$_4$Folate reaction complex by means of its electronic structure and mechanistic detail for the methyl transfer. Thus, a
reasonable chemical model was prepared based on the available X-ray structure of MeCbl-binding domain of the MetH (PDB code: 1BMT, at 3.0 Å resolution)\textsuperscript{41} by docking the isolated CH$_3$-H$_4$Folate on the β-face.

As a first step, we optimized the geometry of Me-cob(III)alamin-MetH resting state by ONIOM-based QM/MM-ME computations that is shown in the Appendix (Figure A5, Appendix) and their main structural parameters are listed in Table 3.1 (column 2 and 3) for the direct comparison with the experimental structure. The fact that the key structural parameters of the MeCbl are correctly reproduced using hybrid QM/MM-ME calculations strengthened our confidence that a cob(I)alamin intermediate could be well characterized by the same level of theory. Therefore, the structure of the cob(I)alamin intermediate was prepared from enzyme-bound MeCbl by removing the methyl group and keeping the number of electrons of the QM (entire cofactor + His759) system consistent with the Co(I) oxidation state. However, upon QM/MM-ME geometry optimization of the cob(I)alamin-MetH, the α-axial His759 residue found to be displaced from the Co(I) center (Figure A6). The QM/MM-ME simulations show that the Co-N(His759) distance elongated to 3.02 Å in Co(I) from 2.14 Å in Co(III) oxidation state. The axial His759 ligand displaced due to the change in oxidation state of the Co metal which is in agreement with the X-ray absorption spectroscopic studies,\textsuperscript{112} indicating cob(I)alamin form of the cofactor is not axially coordinated in solution. Moreover, the displacement of the axial His759 residue from Co(I) center is in agreement with the previous Car Parinello Molecular Dynamics (CPMD) QM/MM computational studies.\textsuperscript{145} It is important to further mention that a short Co(I)-N(His759) axial distance may be advantageous from the enzymatic point of view, since it would allow the axial His759 to
re-coordinate easily to the cobalt center when the cofactor would be re-methylated by CH₃-H₄Folate substrate.

To further investigate the electronic structure and mechanistic details of the MetH-cob(I)alamin:CH₃-H₄Folate reaction complex, we subsequently docked CH₃-H₄Folate substrate on the β-face of the QM/MM optimized enzyme-bound cob(I)alamin intermediate. It has been suggested that an activation of CH₃-H₄Folate takes place by protonation (from general acid, BH) at N5 position; therefore, we utilized the protonated CH₃-H₄Folate referred as a N5H protonation state. We first optimized N5H protonated E.S reactant complex for the methyl transfer (Figure 3.3a) without applying any structural and geometrical constraints either on the cofactor or on the CH₃-H₄Folate in order to attain a more reasonable structure of the model complex. It should be noted here that we tested a number of different starting conformations of the CH₃-H₄Folate on the upper of face the cofactor, however, it appears that all the conformations converge to a point (after the optimization) where methyl group of N5 is perpendicular to Co(I)corrin. In addition, an O-end of propanamide side chain of the Co(I)corrin cofactor is interacting with the -HN3 of the pterin ring and such an intermolecular interaction serve as an anchor between E.S complex. The communication between the CH₃-H₄Folate substrate and the cob(I)alamin through H-bond interactions could be the key in stabilizing the orientation of the pterin ring for a methyl transfer.

The optimized geometry of N5H protonated cob(I)alamin:CH₃-H₄Folate reaction complex ($RC_{N5H}$) is shown in Figure 3.3a, and the corresponding key structural parameters are given in the Table 3.1 (column 5, Co(I):[CH₃-H₄Folate]). Interestingly, the overall parameters of CH₃-H₄Folate substrate-bound cob(I)alamin (column 5 in Table
3.1) and substrate-free cob(I)alamin (column 4 in Table 3.1) are comparable. The distance between the Co and the carbon of the transferred methyl ($C_{Me}$) group found to be 3.59 Å. On the other hand, the Co-$N_{ax}$(His759) distance increased to 3.17 Å which is referred as a His-off conformation of the cofactor because His759 is not bound to the Co center. It should be noted that axial His759 residue in $R_{CN5H}$ cannot move further because it is H-bonded with Asp757 and Ser810 protein residues that act as catalytic triad (His-Asp-Ser) for a methyl transfer. Both His759 and Asp757 residues are tethered in a loop, conferring them a certain degree of flexibility that allows the axial base to be displaced during the catalytic cycle without breaking the hydrogen bond network between them. However, this flexibility is limited because Asp757 is also interacting with a Ser810 of a $\alpha$-helix and this secondary structure is not as flexible as the loop. As a result, the Co-$N$(His759) cannot increase beyond the 3.17 Å in $R_{CN5H}$ without disrupting the hydrogen bond network, something that probably might have a high energy cost.
Figure 3.3. The QM/MM-ME(css) optimized geometries of the enzyme-substrate complex in the N5H protonation state: (a) reactant complex (RC_{N5H}) in His-off conformation, (b) transition state (TS_{N5H}), and (c) product complex (PC_{N5H}) in closed shell singlet (css) state calculations.
The electronic structure of the $\text{RC}_{\text{N5H}}$ was initially optimized considering a closed-shell singlet electronic configuration which is consistent with the Co(I) oxidation state (Figure 3.3a). However, it has been noticed using hybrid B3LYP functional\textsuperscript{78} that the open-shell singlet (oss) state of the free cob(I)alamin intermediate is energetically lower (1-2 kcal/mol) than the closed-shell singlet state (css) indicating a singlet instability of the ground state wave function.\textsuperscript{80,145,172} The spin polarized solution of Co(I)corrin obtained from open-shell singlet calculations shows an unpaired electron on the cobalt metal coupled antiferromagnetically with an unpaired electron on the corrin ring, which was consistent with a $\text{Co}^{\text{II}}(d^7)$-corrin radical ($\pi^*$)\textsuperscript{1} diradical state. Moreover, the ferromagnetic counterpart of the open-shell singlet state, a triplet $\text{Co}^{\text{II}}(d^7)$-radical corrin ($\pi^*$)\textsuperscript{1} state, was also found to be lower in energy (3kcal/mol) than that of the closed-shell singlet, further validating the diradical contribution to the Co(I)corrin.\textsuperscript{80,145,172} Although, the small energy differences between the states might be within the error of the hybrid B3LYP functional, nevertheless, they signify that the closed-shell and the diradical states are close in energy, suggesting that the electron transfer from the Co(I) to the corrin($\pi$) to originate the $\text{Co}^{\text{II}}(d^7)$-corrin radical ($\pi^*$)\textsuperscript{1} may be possible. The multiconfigurational character of the cob(I)alamin intermediate has been further confirmed by the high level CASSCF calculations, and correlated ground state wave function was found to be consistent with that of a closed-shell $\text{Co}^{\text{I}}(d^8)$ configuration (67%) and a diradical $\text{Co}^{\text{II}}(d^7)$-corrin radical ($\pi^*$)\textsuperscript{1} configuration (23%).\textsuperscript{80,145} However, the formation of such unusual oxidation state at ground state was due to the overlap of low-lying metal d-orbitals with the corrin ligand orbitals that allow the electron transfer from the Co(I) to the corrin ring. Therefore, a singlet instability of the wave function can be tested either with triplet state
(ferromagnetic coupling) or with an open-shell singlet state (antiferromagnetic coupling) calculations of the RC_{N5H}.

Subsequently, we checked the possibility of having a diradical state i.e., electron transfer at ground state for the protonated cob(I)alamin:CH$_3$H$_4$Folate complex (RC$_{N5H}$) using triplet state as well as open-shell singlet state calculations. It should be noted that in the this contribution we employed GGA based BP86 functional instead of hybrid B3LYP for the QM/MM calculations because it is well known that the BP86 functional produced correct Co-C bond dissociation energy while B3LYP underestimate the BDE.$^{75,76,115,200}$ Consequently, we first optimized the triplet state of the RC$_{N5H}$ using QM(BP86)/MM-ME calculations and found that from energetic point of view, the triplet state was 2.32 kcal/mol higher in energy than the corresponding closed-shell singlet state (Figure 3.5). The QM/MM-ME optimized geometry of the triplet state is shown in Figure 3.4a and the key structural parameters are listed in Table 3.1 (column 6). Most interestingly, we noticed that the $\alpha$-axial His759 ligand displaced and coordinated to the Co center during the triplet state optimization (Figure 3.4a). The QM/MM calculations of the RC$_{N5H}$ reveal that the Co-N(His759) distance decreased from 3.17 Å in the closed-shell singlet state to 2.07 Å in the triplet state. In other words, the axial ligand displaced from four coordinated Co$^I$(d$^8$) closed-shell singlet state to five coordinated Co$^{II}$(d$^7$)-pterin radical $(\pi^*)^1$ diradical state. Since the cofactor is in a His-on conformation, so it is referred as RC$_{N5H}$ (His-on). However, the distance between the Co and C$_{Me}$ of the methyl group of CH$_3$H$_4$Folate to be transferred is elongated to 4.16 Å as compared to the closed-shell singlet state (3.59 Å). The spin density distribution shows that there is one unpaired electron on the Co atom coupled ferromagnetically with an unpaired electron on the
pterin ring of the CH₃-H₄Folate, consistent with the triplet Co⁺(d⁷)-pterin radical (π*)¹ state.

Figure 3.4. The QM/MM-ME optimized geometry of the N5H protonated cob(I)alamin:CH₃-H₄Folate reactant complex (RCₙ₅₇) in a) triplet state, b) open shell singlet state calculations which shows His-on conformation of Co, consistent with Co⁺(d⁷)-pterin radical (π*) diradical configuration.

Furthermore, we optimized the open-shell singlet state to investigate the possibility of a spin-polarized solution for the cob(I)alamin:CH₃-H₄folate reaction complex (RCₙ₅₇ (oss)). The optimized structure (Figure 3.4b) certainly shows the spin
polarization as expected, with an unpaired electron on Co coupled antiferromagnetically with an unpaired electron delocalized on the pterin ring as well as small spin density on the corrin ring, consistent with a Co$^{II}$($d^7$)-pterin/corrin radical ($\pi^*$)$^1$ diradical state. The energy of the open-shell singlet state was found to be 3.18 kcal/mol higher than the closed-shell singlet state (Figure 3.5), further validating the diradical contributions in the RC$_{N5H}$. It is important to mention that the optimized geometry of the open-shell singlet state is similar to the triplet state (as two spins are separated) but there is very small energy difference (~ 0.8 kcal/mol) between these two states. This has been shown previously for the bioinorganic system such as cob(I)alamin$^{80,145}$ and heme-based compound 1 intermediates$^{201,202}$ that the energies and geometries of the open-shell singlet and triplet states are generally similar. Both the intermediates have a singlet instability and a non-innocent macrocycle (corrin and porphyrin) that can exchange electrons with the metal. As a consequence, the QM/MM-ME calculations of the RC$_{N5H}$ indicate that the closed-shell and diradical states (both open-shell singlet and triplet) are close in energy with His-off and His-on conformation, respectively, suggesting that electron transfer from Co$^I$(d$^8$) to the pterin ($\pi$) to originate the Co$^{II}$($d^7$)-pterin radical ($\pi^*$)$^1$ may be feasible.
Figure 3.5. Potential energy profile (kcal/mol) for the Co-CMe bond formation in the N5H protonated state, based on the QM/MM-ME calculations and the QM/MM-EE energies are given in the bracket. The corresponding optimized geometries are shown in Figures 3.3 and 3.4.

The methyl transfer from the CH$_3$-H$_4$Folate substrate to the cob(I)alamin to generate a Me-cob(III)alamin resting state (Scheme 3.1) is generally assumed to proceed via S$_{N}$$^{2}$-type displacement. In this chapter, S$_{N}$$^{2}$-type mechanism is consistent with the closed-shell singlet state (His-off) transferring methyl cation in a rate-limiting step. However, in view of the energetically close-lying diradical states obtained from triplet and open-shell singlet state calculations (His-on), it can be suggested that a radical mechanism, consisting of an electron transfer (ET) from cob(I)alamin to CH$_3$-H$_4$Folate followed by a methyl radical transfer, is also possible. For the ET-based radical mechanism, the axial His759 residue plays a crucial role to enhance the electron transfer step. Based on the electronic structure of the RC$_{N5H}$ as well as the coordination between the Co center and His759 residue, two possible mechanistic pathways for the methyl
transfer reaction catalyzed by MetH can be envisioned (Scheme 3.1): I) $S_N2$-type pathway i.e., a methyl cation transfer, II) ET-based methyl radical transfer.

\[
\text{Co(I)Cbl} + \text{CH}_3\text{-H}_4\text{Folate} + H^+ \rightarrow [\text{CH}_3\text{-Co(III)Cbl}]^+ + \text{H}_4\text{Folate} \quad (I)
\]

\[
\text{Co(I)Cbl} + \text{CH}_3\text{-H}_4\text{Folate} + H^+ \rightarrow [\text{Co(II)Cbl}]^+ + [\text{CH}_3 - \text{H}_4\text{Folate}]^{•-} \rightarrow [\text{CH}_3\text{-Co(III)Cbl}]^+ + \text{H}_4\text{Folate} \quad (II)
\]

In the next two sections, we would focus on these two mechanisms for the methyl transfer to enzyme-bound cob(I)alamin intermediate, complementing the previously reported experimental study.

Scheme 3.1. (I) $S_N2$-type, and (II) ET-based radical mechanistic pathways for the methyl transfer reaction from a CH$_3$-H$_4$Folate substrate to a cob(I)alamin intermediate catalyzed by methionine synthase (MetH).
3.3.2 S\textsubscript{N}2-type Nucleophilic Displacement (His-off). As discussed above, the S\textsubscript{N}2 pathway proceed in a His-off conformation of the cofactor consistent with closed shell singlet (css) state optimized geometry of the RC\textsubscript{N5H} (Scheme 3.1). As a next step, a transition state (TS\textsubscript{N5H}) for the methyl transfer from the protonated CH\textsubscript{3}-H\textsubscript{4}Folate to the Co(I) center has been optimized with QM/MM-ME simulations. The formation of the TS\textsubscript{N5H} (css) has been characterized by the ONIOM-based vibrational frequency calculations at the same level of theory that displays only one imaginary frequency (329i cm\textsuperscript{-1}). The vibration associated with an imaginary frequency corresponds to the axially directed S\textsubscript{N}2 reaction coordinate containing the N5 of the CH\textsubscript{3}-H\textsubscript{4}Folate, the carbon (C\textsubscript{Me}) of the methyl group, and the Co metal. The transition state (TS\textsubscript{N5H}) located along the S\textsubscript{N}2 pathway, however, can be described as early with respect to N5-C\textsubscript{Me} bond and their key structural parameters are shown in Figure 3.3b and Table 3.2. In the TS\textsubscript{N5H} (css) calculations, the key interacting entities (N5, C\textsubscript{Me} and Co atom) were confined strictly in a linear geometry (\(\angle\text{N5-C-}\text{Co}=174^\circ\)) and both axial (N5-C\textsubscript{Me}= 1.75 Å and C\textsubscript{Me}-Co=2.81 Å) bonds were stretched out along the reaction coordinates. It is important to note that we did not put any constrain on the lower axial His759 residue in the TS\textsubscript{N5H} geometry optimization. The distance between the Co center and the lower axial His759 ligand thus decreased from 3.17 Å in the RC\textsubscript{N5H} to 2.58 Å in the TS\textsubscript{N5H}. The shortening of Co-N(His759) distance in the TS\textsubscript{N5H} could be advantageous from enzymatic point of view as it would facilitate the lower axial protein residue to re-coordinate with the Co center. This important finding also confirm that the proposed S\textsubscript{N}2 pathway for the first-half MetH reaction will lead to the experimentally observed octahedral (d\textsuperscript{6}) Me-Cob(III)alamin species. Alternatively, this movement also validates the crucial role of lower catalytic
triad (His-Asp-Ser) which is known to play a key role for the coordination chemistry of Co metal.

The activation energy barrier for the $S_{N2}$-type methyl cation transfer was found to be 9.01 kcal/mol with the QM/MM-ME calculation (Table 3.2 and Figure 3.5). In addition, we applied single point energy calculations with electronic embedding scheme (EE) known as a QM/MM-EE using the QM/MM-ME optimized geometry, which show the energy barrier of 7.62 kcal/mol. This difference from the two types of schemes (ME and EE) in ONIOM-based calculations is obvious as reported previously, and due to the fact that the electrostatic interactions between two level i.e., QM and MM part are treated differently. The main advantage of the QM/MM-EE calculations is that it includes the electrostatic interaction between the QM and MM part correctly by incorporating the MM charges into a QM Hamiltonian. Most importantly, an approximate ~ 8-9 kcal/mol energy barrier for the methyl transfer from protonated CH$_3$-H$_4$Folate to the cob(I)alamin intermediate, are well in line with the experimental rate constant (7.0 $\times$ 10$^4$ mol$^{-1}$s$^{-1}$ at 37°C), which indicate that the ONIOM-based QM/MM calculations reproduced the enzymatic effect of reaction acceleration by the methionine synthase (MetH). The Mullikan charges computed from the QM/MM-EE simulations (Table 3.2) show that the charge on the C$_{Me}$ (-0.41) in TS$_{5NH}$ increased as compared to the RC$_{N5H}$, indicating methyl group transferred as a cation. Note that the predicted energy barrier in the present QM/MM calculations is lower than that of the reported for small model complex (13 kcal/mol), signifying the crucial role of the MetH.

Upon methyl transfer to the $\beta$-face of the Co center and subsequently optimizing Co(III) oxidation state displaced the lower axial His759 ligand close to the Co center and
thus formed hexacoordinated His-on Me-Cob(III)alamin and protonated H₄-Folate in product complex (PC₅₅₅₅). It is important to mention that we did not apply any constrains to the His759 ligand, therefore present QM/MM calculations validate the consensus that the ligation of the His759 residue with a Co center is certainly controlled by the oxidation state of the Co (Co(I) to Co(III)) as has been concluded by Matthews and coworkers. Moreover, the key structural parameters of the MeCbl in the PC₅₅₅₅ (css) are comparable with the available crystal structure of MeCbl (Table 3.1, column 3 and 7). In other words, our QM/MM calculations proved that the Rossman subdomain (α/β-fold) containing catalytic triad (His-Asp-Ser) communicating with cobalamin through a Co-N(His759) bond during the methyl transfer reaction in the first-half catalytic cycle of MetH. Thus, the role of the MetH is to stabilize the product complex using the Rossman subdomain that certainly plays a critical role in catalysis. An interesting observation here is that, in going from RC₅₅₅₅ to PC₅₅₅₅ via TS₅₅₅₅, the distance between the N3H of the CH₃-H₄Folate and O-end of the propinamide side chain of the corrin ring is elongated from 1.98 Å to 2.68 Å (2.01 Å in TS₅₅₅₅). This indicates that the side chain of corrin ring plays a crucial role in stabilizing the RC₅₅₅₅ and TS₅₅₅₅ for the methyl transfer reactions which were missing in the previous model studies. In other words, the communication between the cofactor and the CH₃-H₄Folate substrate and the displacement of the His759 with respect to Co center are the key components in order to understand the catalytic methyl transfer reactions.

The QM/MM-ME optimizations from the TS₅₅₅₅ to the product complex (PC₅₅₅₅) lead to the formation of six-coordinated MeCbl which is found to be -36.74 kcal/mol (-37.94 kcal/mol with QM/MM-EE) lower in energy than the RC₅₅₅₅ (Figure 3.4 and Table
3.2. The energies of the product complex relative to the RC$_{N_{5}H}$ indicate that the reaction is strongly exothermic as would be expected from charge neutralization reaction (protonation at N5 in the product complex). However, such a highly exothermic nature of the reaction (35.95 kcal/mol in a proteinlike environment ($\varepsilon=4$)) has been also noticed for the model complex that mimic the second-half MetH reaction in which the S$_{N}$2 methyl transfer takes place from MeCbl to unprotonated methylthiolate (CH$_3$S$^-$).$^{114}$ Alternatively, the methyl transfer in the backward direction (first-half MetH reaction) i.e., from MeCbl to protonated H$_4$Folate requires a high-energy barrier for the reaction to proceed as an S$_{N}$2-type, which is energetically unreachable. This is due to the fact that the protonated H$_4$Folate product includes a proton at N5 and the methyl transfer to a protonated N5 species makes this reaction highly uphill. As discussed in the introduction, the active site polar residues (Figure 3.2) such as Asn503 near the pterin ring of the CH$_3$-H$_4$Folate could be important for the stabilization of protonation state in reverse methyl transfer.$^{111}$

Table 3.2. Relative Energies (in kcal/mol), Mulliken Charges and Spin Densities of Cobalamin:CH$_3$-H$_4$Folate Complexes Calculated with ONIOM-based QM/MM calculations.

<table>
<thead>
<tr>
<th></th>
<th>Energy (kcal/mol)</th>
<th>Mulliken charge (Spin density)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>QM/MM-ME</td>
<td>QM/MM-EE</td>
</tr>
<tr>
<td>S$_{N}$2</td>
<td>RC$<em>{N</em>{5}H}$ (css)</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>TS$<em>{N</em>{5}H}$ (css)</td>
<td>9.01</td>
</tr>
<tr>
<td></td>
<td>PC$<em>{N</em>{5}H}$ (css)</td>
<td>-36.74</td>
</tr>
<tr>
<td>ET (T)</td>
<td>RC$<em>{N</em>{5}H}$ (T)</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Barrier (T)</td>
<td>8.05</td>
</tr>
<tr>
<td>ET (oss)</td>
<td>RC$<em>{N</em>{5}H}$ (oss)</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Barrier (oss)</td>
<td>7.72</td>
</tr>
<tr>
<td></td>
<td>PC$<em>{N</em>{5}H}$ (oss)</td>
<td>-38.77</td>
</tr>
</tbody>
</table>
The most important and interesting question that arises here (to understand the reverse methyl transfer) is the timing of proton uptake with respect to the methyl transfer, and unfortunately, this information is unclear from experimental point of view as discussed in the introduction section. Although we have explored the feasibility of $S_N2$ methyl transfer in N5H protonation state, it is worth investigating the activation energy and structural parameters of the reaction complex in unprotonated N5 species of the CH$_3$-H$_4$Folate i.e., N5 state, and that would further provide understanding of the protonation process. As a next step, we removed the proton of the N5H protonation state and reoptimized the unprotonated reactant complex (RC$_{N5}$). The QM/MM-ME optimized geometry is shown in the Appendix (Figure A7a,) which revealed that the Co-C$_{Me}$ and Co-N(His759) distances elongated to 4.92 Å and 4.02 Å, respectively as compared to protonated reactant complex (RC$_{N5H}$). This is due to the fact that the QM part of the QM/MM-ME simulation is relatively neutral because there is no proton at N5, that means neutral CH$_3$-H$_4$Folate is not activated for the methyl transfer reaction and keeps the enzyme-substrate complex apart. Thus, the optimization of the TS$_{N5}$ (Figure A7b, Appendix) for the methyl transfer reaction shows that the activation energy barrier is 39.07 kcal/mol (with negative eigen value of 417i cm$^{-1}$), which is relatively very high. The predicted energy barrier for the RC$_{N5}$ is almost four times more than that of the protonated reaction complex (RC$_{N5H}$). As a result, our QM/MM calculations discard the possibility of methyl transfer reaction from unprotonated CH$_3$-H$_4$Folate to Co(I). On the other hand, the interesting observation is that the optimized TS$_{N5}$ is rather late as compared to the TS found in the protonated complex (TS$_{N5H}$) with respect to the Co-C$_{Me}$ distance of 2.32 Å (Figure 4Sb). It is important to mention that we were unable to optimized the product complex in an
unprotonated N5 state due to the formation of the H₄folate anion which is an unfavorable leaving group, though several attempts were made to optimize it but without success. This can be accounted to the fact that H₄Folate anion is the conjugate base of an acid (formed by methyl transfer from CH₃-H₄Folate) that is expected to have a very high pKa because the pKa of ammonia dissociating to form the ammonia anion is 38.²⁰⁴

Our computational results indicate that without the protonation at the N5 of the CH₃-H₄Folate substrate, the S_N2 methyl transfer to Co(I) intermediate is unreachable because the barrier for the rate-limiting step is very high. Moreover, the H₄folate anion is very poor leaving group. It is important to mention that the methyl transfer from aromatic tertiary amine is not a facile reaction even in the presence of strong nucleophile such as cob(I)alamin and it must requires electrophilic activation of the methyl group probably by protonation to facilitate the methyl transfer. This has been pointed out by Pratt et al. based on their model studies where they demonstrated that a quaternary ammonium salts can function as a methyl donors to cob(I)alamin but not corresponding tertiary amines.¹⁹⁵

As a result, ONIOM based QM/MM calculations of the E.S complex (both protonated and unprotonated) suggest that protonation at N5 position of the CH₃-H₄Folate must take place, either simultaneously with the methyl transfer reaction, or prior to the methyl transfer.

3.3.3 ET-based Methyl Radical Transfer Mechanism (His-on). As discussed above in section 3.1, the electron transfer (ET)-mediated radical mechanism (See equation II) has been suggested for the RC₅₅H to form a CH₃-cob(III)alamin based on the energetically close-lying diradical states calculated from the triplet and open-shell singlet state calculations (Scheme 3.1 and Figure 3.4). The ET-based radical pathway proceed in a
five coordinated His-on conformation (Scheme 3.1), where an electron transfer first takes place from cob(I)alamin to pterin ring of the CH$_3$-H$_4$Folate, consistent with a Co$^{II}$(d$^7$)-pterin radical ($\pi^*$)$^1$ diradical state (unpaired electrons coupled both ferromagnetically and antiferromagnetically). The electrochemical reduction of the CH$_3$-H$_4$Folate substrate, thus induces the homolytic cleavage of N5-C$_{Me}$ bond followed by methyl radical to the Co(II) to generate back Me-Cob(III)alamin intermediate and this pathway is consistent with the ET-mediated reductive cleavage of the N5-C$_{Me}$. However, an ET-based reductive cleavage mechanism has also been suggested for the second-half MetH cycle,$^{115,116,183}$ where the electron transfer first takes places from unprotonated methylthiolate to antibonding $\pi$ orbital of the corrin ring followed by the methyl radical transfer. It should be noted here that in the QM/MM-ME optimized structure of RC$_{N5H}$ (T/oss), the electron transfer take place to an antibonding $\pi$ orbital of the pterin ring with spin density on the terminal and bridging carbon atoms but not directly on the N5 or C$_{Me}$ atoms.

In order to investigate the complete reaction of methyl transfer from the CH$_3$-H$_4$Folate substrate to His-on conformation of the Co center, N5-C$_{Me}$ bond was systematically elongated (with an increment of 0.2Å) towards the Co metal and the Co-C$_{Me}$ was shortened in such a way so that Co-N5 equilibrium distance of 5.68 Å is fixed (See Scheme 3.2). Since there is no structural information available from the existing X-ray data regarding how the two modules containing the cofactor and CH$_3$-H$_4$Folate substrate interact during the methyl transfer. Therefore, the distance between substrate and the cofactor (Co-N5) was kept fixed for all the Co-C$_{Me}$ distances to avoid steric clashes between the substrate and the cofactor (Scheme 3.2). The QM/MM-ME potential energy profile associated with the cleavage of the N5-C$_{Me}$ bond along the Co-C distances
are shown in Figure 3.6. It should be noted that all degrees of freedom were fully optimized in ONIOM-based QM/MM-ME calculations for the every constrained Co-C\textsubscript{Me} distances both in open shell singlet (oss) as well as in triplet state optimizations. In addition, we performed single point QM/MM-EE calculations based on the previously optimized QM/MM-ME geometry for all distances to incorporate the electrostatic interaction between the QM and MM parts more accurately. The single point QM/MM-EE calculations provide a potential energy profile similar to that of the ME scheme (See Table 3.2 and Figure 3.6).

Scheme 3.2. The schematic representation of the ET-mediated radical mechanism where the N5-C\textsubscript{Me} bond is elongated and Co-C\textsubscript{Me} bond is systematically shortened so that the Co-N5 distance (5.68 Å) is fixed.
Figure 3.6. Potential energy profiles with energy barrier for the formation of Co-C\textsubscript{Me} bond in N5H protonated reaction complex of MetH-cob(I)alamin:CH\textsubscript{3}-H\textsubscript{4}Folate shown in Figure 3.4, calculated for the triplet (T) as well as open shell singlet (oss) QM/MM calculations. The energies are given relative to the RC\textsubscript{N5H} where the Co-C\textsubscript{Me} is 4.16 Å shown in Figure 3.4. The second embedded plot is the evolution of spin density distribution along the Co-C\textsubscript{Me} distance extracted from open shell singlet (QM/MM-EE) calculations.

The energy barrier for the methyl transfer reaction in the open shell singlet (oss) calculations appears to be 8.05 kcal/mol and 8.68 kcal/mol with QM/MM-ME and QM/MM-EE scheme, respectively (Table 3.2 and Figure 3.6). The evolution of spin density distribution for the formation of the Co-C\textsubscript{Me} bond is also plotted that is embedded in the lower part of Figure 3.6, and these spin density numbers are extracted from the
open-shell singlet state QM/MM-EE (oss) calculations. In the equilibrium geometry of the $RC_{N5H}$ with the N5-C$_{Me}$ bond length 1.52 Å, the $\alpha$ spin density is centered on a Co metal, while the $\beta$ spin density is delocalized on $\pi^*$ pterin orbital with small contribution of $\pi^*$ corrin orbital (Figure 3.6). Therefore, one electron transfer from the HOMO of the Co(I) metal to the LUMO of the pterin/corrin takes place, leading to a Co$^{II}(d^7)$-pterin/corrin radical ($\pi^*$)$_1$ diradical configuration. Note that the spin density distribution on the pterin ring of the CH$_3$-H$_4$Folate substrate calculated from open-shell singlet QM/MM calculations is the same that can be observed if we reduced the isolated CH$_3$-H$_4$Folate substrate by one electron. As the N5-C$_{Me}$ bond is stretched or the Co-C bond is shortened, the $\beta$ spin density shift from the pterin ring of the CH$_3$-H$_4$Folate to the N-C$_{Me}$ bond. At a Co-C distance of 3.8 Å (which display energy barrier for the methyl transfer reaction), the crossing between the $\beta$ spin densities on the CH$_3$-H$_4$Folate and C$_{Me}$ takes place with only residual spin density on the corrin ring (Figure 3.6). However, the C$_{Me}$ atom attains maximum $\beta$-spin density at a Co-C$_{Me}$ distance of 3.2 Å, consistent with the Co(II)--C$^\bullet_{Me}$--N5-(H$_4$Folate). In other words, the N5-C$_{Me}$ bond cleaved homolytically leading to the formation of a methyl radical. Further, elongation of the N5-C$_{Me}$ bond leads to the formation of a closed-shell configuration with no spin density distribution around Co-C$_{Me}$ distance of 2.8 Å. This is due to the fact that the Co(III)-C$_{Me}$ bond formation take place and moreover, it is well known that the ground state of Me-Cob(III)alamin is closed shell singlet. Therefore, at a Co-C of distance of 2.0 Å, the methyl group is completely transferred to the cofactor, giving the product of the reaction that is six coordinated Me-Cob(III)alamin and protonated H$_4$folate. The potential energy profile (Figure 3.6) further indicates the methyl transfer reaction is highly exothermic by -38.77 kcal/mol calculated
with QM/MM-ME (-37.95 kcal/mol with QM/MM-EE) which is in agreement with the
$S_N2$-type mechanism ($\sim 37$ kcal/mol) from energetic point of view.

The energy barrier for the methyl transfer reaction in the triplet state QM/MM(T)
calculations turns out to be 8.05 kcal/mol and 8.68 kcal/mol with QM/MM-ME and
QM/MM-EE calculations, respectively (Table 3.2 and Figure 3.6) which is almost
comparable to the one calculated with the open-shell singlet state calculations. All these
energy values shown in Figure 3.6 should be considered as an upper bound because zero
point energy corrections were not included in the ONIOM-based QM/MM calculations. It
is important to mention that we were unable to optimize the triplet state QM/MM-ME(T)
calculations after the Co-C$_{Me}$ distance 2.6 Å owing to the fact that the lowest state of the
MeCbl is the closed shell singlet. In Figure 3.6, we also plotted the QM energy
contribution to the overall QM/MM energy in order to understand the role of protein
(MM part), which were extracted from the QM/MM-ME calculation. Although, the actual
trend of QM energy profile for the Co-C$_{Me}$ bond formation is similar to that of the
QM/MM energy in both open-shell singlet and triplet calculations, but there is small
contributions from the MM part. In particular, the protein effect (MM energy components
of the combined QM/MM-ME calculations) on the energy barrier (at Co-C$_{Me}$ distance 3.8
Å) for the methyl transfer reaction is 0.45 kcal/mol and 0.67 kcal/mol in triplet and open
shell singlet state calculations, respectively.

As discussed before in section 3.1, QM/MM optimized structures of the triplet
and open-shell singlet (oss) states are energetically $\sim 2$-3 kcal/mol higher than that of the
closed-shell singlet (css) state (Figure 3.5). Therefore, it can be concluded that the
calculated energy barrier (i.e., in total $\sim$10-11 kcal/mol) for the methyl radical transfer
(with His-on conformation) both in triplet and open-shell singlet state calculations is quantitatively comparable to that of the previously calculated activation energy (~ 8-9 kcal/mol) for the S_N2 pathway. However, the main advantage of ET-based radical mechanism from enzymatic point of view is that a methyl radical can be transferred at a longer distance (Such as 4.16Å in RC_{N5H}), which does not require the close proximity of the two binding modules of MetH, as would be required in an S_N2-type. It should be noted, however, that the proposal based on the ET-based radical mechanism is similar to the one suggested by the Matthews et al. described as a single electron transfer (SET).\textsuperscript{111}

The difference lies in the initial electron transfer from cob(I)alamin to the CH_3-H_4Folate substrate, our theoretical study in this chapter shows that the electron is being transferred to π* orbital of pterin ring of N5H protonated CH_3-H_4Folate, whereas they suggested that the electron is transferred to unprotonated CH_3-H_4Folate followed by the proton transfer to initiate the homolytic cleavage of the N5-C_Me bond.

3.3.4 Understanding the protonation process of the CH_3-H_4Folate Substrate. As it has been shown based on the QM/MM calculations, if the methyl transfer from the CH_3-H_4Folate to cob(I)alamin intermediate proceed via an S_N2 or an ET-based mechanism then one might expect the activation of the CH_3-H_4Folate to transfer the methyl group, presumably by a protonation at the N5 position. Therefore, it is important to understand and investigate the protonation event at N5 position when the CH_3-H_4Folate fragment is bound to its domain in MetH.

The pK_a of the free CH_3-H_4Folate substrate in aqueous solution is determined to be a 5.05 which is associated with the protonation at N5 as detected by the measurement of fluorescence changes when the pH is lowered.\textsuperscript{205} This pK_a value of the CH_3-H_4Folate
is likely to be increased (~ 2 unit) by the hydrophilic protein residues such as aspartatic acid (Asp) and asparagine (Asn) which are usually found to be near the pterin ring (Figure 3.2, PDB ID: 1Q8J at 1.9 Å resolution). Among these residues, O-end side chain of the Asn508 residue is H-bonded with the N5 group of pterin ring and such an H-bonding interaction has been suggested to be involved in the stabilization of the protonation state and high energy intermediates (such as TS), but this interaction is not chemically suitable to act as a proton donor. Moreover, the crystal structure of CH$_3$-H$_4$Folate-MeH shows there is no general acid catalyst (BH) near N5 position and the CH$_3$-H$_4$Folate substrate has been suggested to be bound with the MetH in an unprotonated form. On the other hand, it has been shown that product H$_4$Folate contains proton at N5 when the methyl group is transferred to cob(I)alamin/cob(I)amide. In consequence, the timing of proton uptake relative to methyl group transfer is very crucial but remains unclear from experimental point of view. Therefore, the presence of cob(I)alamin cofactor may be important for the substrate activation, and the proton uptake could take place in a ternary cob(I)alamin.CH$_3$-H$_4$Folate.MetH complex, either prior to or during the methyl transfer reaction. On the other hand, it has been shown experimentally in the case CoFeSP methyltransferases (MeTr) that the binding of CH$_3$-H$_4$Folate to MeTr is associated with a pKa increase and a proton uptake in a binary E.S complex.

In addition, it has been suggested by Matthews et al. using pH dependent and rate constant studies of a methyl transfer reaction from CH$_3$-H$_4$Folate to exogenous cob(I)alamin that a general acid/base analysis could be involved in the ternary cob(I)alamin.CH$_3$-H$_4$Folate.MetH complex. Therefore, based on the present QM/MM calculations of the MetH-cob(I)alamin:CH$_3$-H$_4$Folate complex and the analysis
of the X-ray structure of CH$_3$-H$_4$Folate binding module (Figure 3.2) it can be inferred that the proton transfer and methyl transfer steps are somehow coupled, probably with the general acid catalyst (BH) that is initially hydrogen bonded to the N5 of the CH$_3$-H$_4$Folate. As the N-CMe bond is weakened in the ternary complex (via either an S$_N$2 or ET-based mechanism), the N5 of the H$_4$folate anion leaving group increasingly becomes more basic, and the proton transfer takes place when the pKa of N5 is more than that of the general acid catalyst. Alternatively, the presence of general acid catalyst would decrease the QM/MM energy barrier for the reverse methyl transfer i.e., from MeCbl to protonated H$_4$folate which has been predicted to be very high in the present QM/MM calculations (~37-39 kcal/mol) of the N5 protonated E.S complex (Figure 3.5 and Table 3.2). Therefore, a reverse methyl transfer from MeCbl to protonated H$_4$folate will simultaneously facilitate the proton transfer back to the general base (B') to stabilize the neutral CH$_3$-H$_4$folate. A full active site CH$_3$-H$_4$folate substrate need to be considered which is thus beyond the present study due to the fact that the β-site (cap subdomain) of the corrinoid-protein is not wide enough to accommodate all the active site residues of the CH$_3$-H$_4$folate binding domain. Moreover, there is no available X-ray structure of the reaction complex containing cob(I)alamin- and CH$_3$-H$_4$folate- binding modules owing to their high conformational flexibility.

3.3.5 Implications for the Re-methylation of the cob(I)alamin/cob(I)amide intermediate to other Methyltransferases. In the present theoretical account, we suggested that the electrophilic activation of very difficult methyl donors such as CH$_3$-H$_4$folate take place by the protonation at methyl-bound N5, either simultaneously with the methyl transfer or prior to methyl transfer but not after. Therefore, the results of the
study performed in this chapter including mechanistic details of methyl transfer reaction and protonation process can be applied to broad class of corrinoid-methyltransferases, such as AcsE (CH$_3$H$_4$Folate-CFeSP)$^{208}$ and HgcAB bacterial mercury methylation.$^{192}$ In addition, other MtrH,$^{209}$ an enzyme that catalyzes transfer of a methyl group from N5-methyltetrahydrodromethanopterin (structural analogue of the CH$_3$H$_4$Folate methyl donor) to a corrinoid protein. All these enzymes are able to catalyze methyl transfer to exogenous cob(I)alamin/cob(I)amide intermediate. However, both MetH and MtrH bind the corrinoid cofactor with a histidine ligand from the protein serving as the lower axial ligand while corrinoid proteins of the HgcAB recently found to be bound with a cysteine residue. On the other hand, a distinguished feature of the CFeSP that serves, as a methyl acceptor from AcsE is its base-off and His-off conformation, which means no protein-derived His759 residue bound to the cobalt center. Therefore, the methyl transfer reaction to such corrinoids proteins is likely to proceed via an $S_N$2 pathway, rather than an ET-based radical mechanism in which His759 ligand plays a crucial role in promoting the electron from Co(I) to the CH$_3$-H$_4$Folate. In addition, a methyl donors to cob(I)alamin or to structurally related corrinoid derivatives (cob(I)amide) include not only CH$_3$-H$_4$Folate but also methanol,$^{208}$ chloromethane,$^{210}$ mono-, di- and tri-methylamines.$^{210,211}$ However, the activation of these challenging methyl donors is less understood and thus, a little is known in terms of the chemical mechanisms. Although, one possibility is protonation of the methyl donors while other is the involvement of the Zn metal for the electrophilic activation. In the case of methanol coenzyme M methyltransferase (MtaABC),$^{212,213}$ a Zn(II) is found near the methanol binding module which is believed to be involved for the activation. Most interestingly, the crystal structure of the Methanol: cob(I)amide
(MtaBC) reaction complex reveals the Co-N(His136) distance 2.51 Å, which clearly indicating the possibility of ET-mechanism in addition to $S_N2$ pathway. As a result, the attractive feature of the computational study presented in this chapter can be applied to a wide range of methyltransferases to understand the activation process of the methyl donors and their subsequent enzymatic mechanism.

3.4 Conclusion

In this chapter, the electronic structure and mechanistic details of the first-half reaction of the cobalamin-dependent methionine synthase (MetH) where a methyl transfer takes place from the \( \text{CH}_3-\text{H}_4\text{Folate} \) substrate to cob(I)alamin intermediate has been investigated. The analysis was performed using ONIOM-based QM/MM computations on the enzyme-substrate reaction complex. Since there is no crystal structure available for the cob(I)alamin:CH\(_3\)-H\(_4\)Folate reaction complex, the structural model was prepared using the X-ray crystal structure of cobalamin binding domain (PDB Id: 1BMT, resolution 3.0 Å) by docking the CH\(_3\)-H\(_4\)Folate on the \( \beta \)-face of the cofactor.

The hybrid QM/MM calculations reveal the traditionally assumed $S_N2$ mechanism for the formation CH\(_3\)-Cob(III)alamin resting state that is a methyl transfer from the protonated CH\(_3\)-H\(_4\)Folate to enzyme-bound cob(I)alamin intermediate. The $S_N2$-type mechanism is consistent with the closed-shell singlet ground state configuration (His-off conformation) transferring methyl cation in a rate-limiting step followed by the lower axial His759 ligand. The activation energy barrier in the protonated CH\(_3\)-H\(_4\)Folate:cob(I)alamin complex found to be $\sim 8\text{-}9$ kcal/mol which is reasonable with respect to the experimental rate constant. The key structural and geometrical parameters of the MeCbl resting state formed in the product complex are in agreement with the
existing X-ray crystal structure. The protonation at the methyl-bound N5 species of the CH$_3$-H$_4$Folate substrate results in the formation of quaternary ammonium cation which is essential for the activation of the CH$_3$-H$_4$Folate substrate in an enzyme-bound cob(I)alamin intermediate. However, the QM/MM calculations in the unprotonated CH$_3$-H$_4$Folate substrate show very high-energy barrier ($\sim$ 39 kcal/mol) for the methyl transfer that is almost four times more than that of found in the protonated analogue of the substrate. The enormous energy barrier is attributed to the fact that the product of this reaction i.e., H$_4$Folate anion is a very poor leaving group. These findings indicate that the reaction is inaccessible without the protonation step and the activation of the substrate must take place by protonation. As a result, our QM/MM calculations discard the possibility of a methyl transfer reaction from unprotonated CH$_3$-H$_4$Folate.

However, in view of the energetically close-lying diradical states obtained from triplet and open-shell singlet state QM/MM calculations, it can be suggested that a radical mechanism, consisting of an electron transfer (ET) from cob(I)alamin to protonated CH$_3$-H$_4$Folate followed by a methyl radical transfer, is also possible. The ET-based radical pathway proceed in a five coordinated His-on conformation (Scheme 3.1), where an electron transfer first takes place from Co(I) to pterin ring of the CH$_3$-H$_4$Folate, consistent with Co$^{II}(d^7)$-pterin radical ($\pi^*$) diradical state. The electrochemical reduction of the CH$_3$-H$_4$Folate, thus induce the homolytic cleavage of the N5-C$_{Me}$ bond followed by the transfer of methyl radical to Co(II) to form a six coordinated CH$_3$-cob(III)alamin resting state and a neutral H$_4$Folate. Moreover, the predicted energy barrier ($\sim$ 8 kcal/mol) for the ET-mediated radical reaction is comparable with that of the $S_N$2 reaction pathway. It should be noted that the ET-based radical mechanism is similar to
the single electron transfer (SET) mechanism suggested by Mathews et al.\textsuperscript{111} but there is difference in the initial step of electron transfer to the substrate.

There are two key differences between the \(S_N2\)-type and ET-based radical mechanism. The primary difference between these two mechanisms (ET and \(S_N2\)-type) is the initial step of the reaction complex (\(RC_{N5H}\)) that involves one-electron reduction of the \(CH_3-H_4\)Folate substrate. The ET takes place between the cob(I)alamin intermediate and protonated \(CH_3-H_4\)Folate substrate thus forming a \(\pi^*\)-pterin radical that induce the cleavage of the \(N5-C_{Me}\) bond to form a \(Co-C_{Me}\) bond. The second difference is the coordination of lower His759 residue with the Co center, the ET-mediated radical mechanism takes place in His-on conformation where the His759 plays a crucial role in promoting the electron from \(Co^1(d^8)\) metal to the \(CH_3-H_4\)Folate substrate. On the other hand, \(S_N2\)-type mechanism occurs in His-off conformation where the formation of the \(Co-C_{Me}\) bond takes place first followed by the displacement of His759 ligand. From enzymatic point of view, the main advantage of ET-based radical mechanism is that a methyl radical can be transferred at a longer distance, which does not require the close proximity of the two binding modules of MetH, as would be required for an \(S_N2\)-type.

However, the crystal structure of \(CH_3-H_4\)Folate binding module shows no obvious acid catalyst (BH) near the N5 position, but the product \(H_4\)Folate contains a proton at N5 when the methyl group is transferred to cob(I)alamin. Thus, the timing of proton uptake relative to methyl group transfer is very crucial but remains unclear from experimental point of view. Based on the activation energy of the TS both in the protonated and unprotonated N5 species of the \(CH_3-H_4\)Folate, it can be suggested that the protonation event must takes place, either prior to or during the methyl transfer reaction in a ternary
complex but not after the methyl transfer. The results of this chapter including chemical mechanisms and implications of the protonation process can be applied to a broad class of corrinoid-methyltransferases, which employ CH$_3$-H$_4$Folate substrate or its related structure as a methyl donor.
CHAPTER 4

ELECTRONIC STRUCTURE OF COFACTOR-SUBSTRATE REACTANT COMPLEX OF METHIONINE SYNTHASE²

4.1 Introduction

The cobalamin-dependent methionine synthase enzyme (MetH) catalyzes the transfer of the methyl group from methyltetrahydrofolate (CH₃–H₄folate) to homocysteine (Hcy) thus forming methionine (Met).¹¹,²⁵,³¹,³²,⁵¹,⁹⁵,²¹⁴,²¹⁵ The key step in the catalytic cycle involves displacement of the methyl group from the methylcobalamin cofactor (MeCbl) to Hcy in which the cobalamin cofactor serves as an intermediate cycling between hexa-coordinated methyl-cob(III)alamin and tetra-coordinated cob(I)alamin complexes.¹⁰⁰,²¹⁶,²¹⁸ Owing to the presence of a zinc ion (Zn²⁺) near the reaction center, the Hcy substrate enters the reaction as a deprotonated species (Scheme 4.1).¹⁶⁴,²¹⁹ It has been suggested that the activation of thiol is promoted due to the lewis acidity of catalytic zinc ion.²²⁰

The MetH enzyme has a modular architecture comprising of four different binding domains.²¹⁷ The modules containing Hcy, CH₃–H₄folate, and MeCbl are involved in the main catalytic cycle, while the fourth AdoMet module is essential for the

² The material presented in this Chapter was reprinted (adapted) with permission from Kumar, N.; Jaworska, M.; Lodowski, P.; Kumar, M.; Kozlowski, P. M. J. Phys. Chem. B 2011, 115, 6722. Copyright 2011 American Chemical Society
Scheme 4.1. First Half-Catalytic Cycle of Methionine Synthase (MetH) Enzyme Showing Transfer of a Methyl Group from Methylcobalamin (MeCbl) Cofactor to Homocysteine (Hcy).

...
Several mechanistic routes have been proposed for the methyl transfer mediated by MeCbl. Although it is generally believed that the methyl transfer reactions catalyzed by MetH enzyme proceed via S_N2-type nucleophilic displacement, alternative mechanistic pathways, including reductive elimination/oxidative addition as well as mechanisms involving electron transfer (ET), have also been suggested. The main inadequacy of the S_N2-type mechanism in the context of cobalamin-dependent MetH is
the required proximity between two reacting domains, which may not be effective from the enzymatic point of view. Indeed, the crystallographic studies of Co- and Fe-containing corrinoid iron–sulfur protein (CoFeSP) involved in acetyl-CoA synthesis,\textsuperscript{223} in which methyl transfer reactions takes place between two individual proteins (i.e., cobamide cofactor and Ni–Ni–[4Fe–4S] cluster A) revealed such limitation. Both methyltransferases (i.e., MetH as well as CoFeSP) share the same difficult task involving specific and reversible interaction between methyl donor/acceptor and cobalamin. In the case of CoFeSP, it was expected that a \(\sim3\)–4 Å distance between methyl donor and acceptor would allow an \(S_N2\)-like displacement. However, attempts to manually fit resulting modules failed to bring both metal sites closer than 8 Å because of clashes between the protein backbones. It was also suggested that a conformational switch at both reaction centers might take place in order to induce the \(S_N2\)-like methyl transfer.\textsuperscript{223}

To overcome the difficulties associated with distance constrains, a related mechanistic pathway (referred to as reductive cleavage) has been proposed\textsuperscript{115,116} that does not require close proximity of two reacting centers, and the methyl group can be effectively transferred at longer distances. The key difference between the \(S_N2\) and reductive cleavage mechanisms is the initial step that involves one-electron reduction of the cofactor. The ET takes place between the deprotonated Hcy and the MeCbl cofactor thus forming a \(\pi^*\)-corrin radical followed by the reductive Co–C bond cleavage. Electrochemical\textsuperscript{140,141,143,144,169,224} as well as computational\textsuperscript{115,116} studies revealed that the strength of Co–C bond in one-electron reduced form of MeCbl cofactor is significantly lower than its neutral analogue. It should be noted, however, the proposal based
on reductive cleavage of Co–C bond is different than what has been described as a single electron transfer (SET).\textsuperscript{111} In the SET mechanism, the initial step involves homolytic cleavage of the Co–C bond followed by ET from the methyl radical to cob(II)alamin without involving the formation of one-electron reduced form of the MeCbl cofactor.

Both S\textsubscript{N}2 as well as reductive cleavage pathways have been investigated computationally for model systems using density functional theory (DFT).\textsuperscript{114,115} In addition, the feasibility of ET-based reductive cleavage mechanism has been demonstrated inside the enzyme based on quantum mechanical/molecular mechanical (QM/MM) analysis.\textsuperscript{116} Although these studies proved basic concepts, they opened several issues that require further analysis. The main difference between these two mechanisms is the distance at which methyl transfer takes place between the Hcy substrate and the MeCbl. The S\textsubscript{N}2-type mechanism was studied with the C···S distance equal to 2.8 Å, whereas in ET-based mechanism, the distance was assumed to be 4.0 Å. The later was based on the analogous studies in corrinoid iron-sulfur proteins\textsuperscript{223} where the methyl transfer reaction was suggested to take place at a similar distance. Thus, it appears that the distance between two interacting moieties (MeCbl and Hcy) plays a critical role in the methyl transfer reaction whether it occurs via S\textsubscript{N}2-type or ET-based mechanisms. Moreover, the closed-shell wave function was required for the former, while the open-shell singlet for the latter to describe the formation of the diradical species. Therefore, only an approximate description can be obtained using DFT, which is based on a single reference wave function, and a more complex wave function is required to describe the reactant complex.

The aim of the theoretical study performed in this chapter is to analyze the electronic and structural properties of MeCbl:Hcy complex using the complete active space self-
consistent field (CASSCF) method followed by the second-order perturbation theory (MR-PT). In addition, the electronic structure of the reaction complex is further investigated as a function of C···S distance between reaction centers, and its possible implications with regard to the enzymatic reaction mechanism are discussed.

4.2 Computational Methods and Model System

4.2.1 Structural Models. To explore the electronic properties of MeCbl:Hcy reactant complex two structural simplifications with regard to corrin macrocycle have been employed: the model with a truncated ring representing only the inner conjugated macrocycle (Figure 4.2c) and the other model containing a full corrin macrocycle (Figure 4.2b). The full corrin, denoted as [Co$^{III}$(corrin)]$^-$Me$^+$, refers to the structure with side chains replaced by hydrogen atoms. In both the models, methylthiolate (−S–CH$_3$) was used as a structural mimic of deprotonated Hcy substrate. The axial base was not included in the calculations because the most energetically efficient reductive Co–C cleavage takes place in the base-off form. This has been confirmed by the previous electrochemical$^{144}$ and theoretical studies.$^{115,116}$ Employing different complexities of the corrin ring, two structural models of MeCbl:Hcy have been prepared. The first model, referred to as small model (SM), was based on a corrin ring with inner conjugated macrocycle only (Figure 4.2c). The second based on a full corrin, i.e., [Co$^{III}$,(corrin)]$^-$Me$^+$···S–CH$_3$, and referred to as large model (LM) was the same as used in previous calculations$^{115,116}$ where electronic properties of the complex have been investigated at DFT. The symmetry plane perpendicular to the corrin ring could only be imposed in the case of SM. The use of $C_s$ symmetry was advantageous with regard to a proper orientation of $p_y$ or $p_z$ orbitals of S in order to describe the degeneracy of these
orbitals separately. Note that SM was primarily used to establish electronic configuration required to describe the reactant complex and can be viewed as a truncated version of LM.

Figure 4.2. (a) Molecular structure of methylcobalamin (MeCbl) and its structural simplifications referred to as (b) large (LM) and (c) small (SM) models respectively in this chapter.
The geometries of both reactant complex models were optimized at DFT level of theory employing the BP86 functional. The coordinates of LM complex were initially obtained from previous DFT/Car–Parrinello molecular dynamics (CPMD) calculations, and the C···S distance was systematically varied starting from 2.6 up to 4.6 Å. At each point, the geometry of reactant complex was fully optimized keeping the carbon–sulfur distance fixed. In the case of SM, the plane of symmetry was retained. Such prepared structural models (SM and LM) at varied C···S distances were used for CASSCF followed by the second order perturbation theory calculations.

4.2.2 CASSCF/MR-PT Calculations. Although the Kohn–Sham formalism (i.e., DFT) is widely used nowadays, certain electronic properties of reactant complex, such as near-degeneracy properties of the electronic wave function when SN2 pathway turns to ET (or vice versa), cannot be adequately described. The proper description requires multireference wave function, such as the one based on CASSCF formalism. Due to lack of dynamical correlation energy, CASSCF calculations alone are not sufficiently accurate and need to be corrected by second order perturbation theory. Two different implementations of multireference perturbation theory (MR-PT) have been used to study the reaction complex. For SM, the calculations were performed using MOLCAS-7.4 program. The ANO-L basis set was employed in the calculations: 6s5p3d1f contraction for cobalt, 5s4p2d for sulfur, 3s2p1d for carbon and nitrogen, and 2s for hydrogen. To understand which orbitals are important in CAS calculations, first we carried out the restricted active space CAS (RASSCF) calculations on the optimized geometry using C···S distance equal to 4.0 Å. Second, based on RASSCF results, CASSCF calculations have been performed by varying the C···S distance. The dynamical correlation was taken
into account by performing multistate second-order perturbation theory calculations (MS-CASPT2). Note that they have been carried out with the 0.3 level shifts parameter based on the CASSCF wave function. For LM, the PCGAMESS/Firefly QC package was applied to perform CASSCF calculations followed by the quasidegenerate perturbation theory (QDPT2) with multiconfigurational self-consistent field reference function and referred to as a MC-XQDPT2. The 6-31G(d) basis set was applied in all CASSCF/MC-XQDPT2 calculations. In order to carry out the MC-XQDPT2 analysis, initially state average (SA) CASSCF calculations were performed employing 12 electrons and 11 orbitals in the active space.

4.3 Results and Discussion

4.3.1 Electronic Properties of Reactant Complex Based on DFT Calculations.

Initially, the S_N2 reaction pathway has been computationally investigated assuming closed-shell\(^{114}\) (i.e., RHF-type) wave function for \(\text{Im} - \text{[Co}^{III}\text{(corrin)}]^{+} \cdot \text{Me}^{+} \cdots \text{S} \cdot \text{CH}_3\) complex using the B3LYP functional. The transition state (TS) predicted for the methyl transfer reaction, in which computed C⋯S distance was 2.88 Å, has been described as very early. The low imaginary frequency of \(94.3i\) cm\(^{-1}\), associated with the TS, corresponds to an axially directed S_N2 displacement involving the motion along the S, CH\(_3\), Co, and N\(_\text{His}\) coordinates. From the structural point of view, the geometry of the MeCbl cofactor was only slightly different in TS as compared to that of the MeCbl:Hcy reactant complex. More specifically, Co−C bond in TS was stretched from 2.01 to 2.08 Å along the reaction coordinates, while the Co−N\(_\text{His}\) distance was elongated from 2.57 to 2.70 Å. The B3LYP-based structure of TS along the S_N2 pathway is very similar to one predicted with gradient corrected BP86.
functional,\textsuperscript{227} despite the fact that the B3LYP functional significantly underestimates the strength of the Co–C bond.\textsuperscript{76,79} Such structural similarities of predicted TS using both types of functionals can be attributed to the formal heterolytic cleavage of the Co–C bond, which takes place during the methyl group transfer.

When the model of MeCbl:Hcy reactant complex was investigated using unrestricted (i.e., UHF-type) formalism, the lowest electronic state was found to be consistent with an open-shell singlet, i.e. [Co\textsuperscript{III}(corrin')]–Me···S–CH\textsubscript{3} configuration at a C···S distance of 4.0 Å.\textsuperscript{115} Moreover, the spin-polarized singlet solution has also been observed inside the enzyme.\textsuperscript{116} The open-shell singlet solution (which describes diradical nature of the reactant complex) implies that the ET was induced from the substrate to the cofactor. In summary, at longer C···S distances, open-shell configuration was the lowest energy solution, while at short distances, the closed-shell configuration was the lowest one.

4.3.2 Selection of the Active Space. One of the main objectives of study performed in this chapter was to construct multireference wave function capable of describing both the S\textsubscript{N}2-type as well as ET-based mechanisms as a function of C···S distance. Based on previous DFT calculations,\textsuperscript{114-116} at least two electronic configurations corresponding to closed-shell as well as open-shell singlet need to be taken into consideration. Taking into account that the S\textsubscript{N}2 and ET mechanisms depend on the distance between reacting centers, switching of these is expected to take place. For that reason, the state-of-the-art ab initio CASSCF wave function was constructed to describe near-degeneracy problems of the electronic states.

The selection of the active space plays a critical role in CASSCF calculations. We primarily used the outcome from DFT analysis and selected Kohn–Sham orbitals based
on restricted as well as unrestricted calculations that can be associated with closed- and open-shell electronic configurations, respectively. Taking into account that the selection of active space orbitals is further complicated when applied to molecules containing different fragments as well as the presence of transition metal, we also performed preliminary restricted active space SCF (RASSCF) calculations in order to screen the resulting configurations. Thus, the selection of active orbital requires the inclusion of metal’s d orbitals, corrin π orbitals, and more exclusively sulfur orbitals that are participating in the reaction. In addition, the previous reported CASSCF calculations for tetra-pyrrolic systems, including cobalamins, corrole, and porphyrin systems assist the active space selection.

At first, the preliminary calculations for SM, in which the C···S distance was assumed 4.0 Å, were carried out employing RASSCF wave function with 20 active orbitals and 20 active electrons allowing a maximum of 4-fold excitations. The RASSCF results provide an important insight for the electronic structure of the reactant complex: the lowest solution of A’ symmetry was predicted to be |...S(p_y)^(1)(π^*_{corrin})^(1)...|, while that with A” symmetry was found to be |...S(p_z)^(1)(π^*_{corrin})^(1)...|. The singly occupied natural orbitals representing two different symmetry based solutions have been shown in Figure 4.3.
Figure 4.3. Singly occupied sulfur and corrin natural orbitals representing two different symmetry solutions based on RASSCF calculations for small model (SM).

To confirm the diradical character of the lowest energy configuration, similar calculations were also performed assuming a triplet state. The singlet–triplet energy gap was found to be very small, i.e., \( \sim 1.0 \text{ kcal/mol} \). The RASSCF calculations were interpreted as follows: isolated MeCbl cofactor has closed-shell character and dominant electronic configuration of the form \(|...\left(\pi_{\text{corrin}} + d_z^2\right)^2\left(\pi_{\text{corrin}} + d_{yz}\right)^2\left(\pi^*_{\text{corrin}}\right)^0...|\), in which the lowest unoccupied molecular orbital (LUMO) has an antibonding character \(\pi^*_{\text{corrin}}\) entirely localized on the corrin ring.\(^{229}\) The model of Hcy substrate was described by the configuration \(|...\left(S(p_y)^2S(p_z)^2(\sigma^*_{\text{C-S}})^0\right)...|\), where \(S(p_y)\) and \(S(p_z)\) denoted degenerate orbitals of the sulfur-
containing lone pairs of electrons. More precisely, these orbitals are labeled as \( S(p_{||}) \) and \( S(p_{\perp}) \) where the first one is parallel to the plane of the corrin ring, and the second one is perpendicular. Thus, according to RASSCF results, electronic structure of MeCbl:Hcy reactant complex (Scheme 4.1) can be described as follows:

\[
\Psi_{\text{dirad} \parallel} \approx \ldots (S_{||})^1 (\pi^*_{\text{corrin}})^1 \ldots \quad (A' \text{ symmetry}) \quad (1)
\]

\[
\Psi_{\text{dirad} \perp} \approx \ldots (S_{\perp})^1 (\pi^*_{\text{corrin}})^1 \ldots \quad (A'' \text{ symmetry}) \quad (2)
\]

\[
\Psi_{\text{ionic}} \approx \ldots (S_{||})^2 (S_{\perp})^2 (\pi^*_{\text{corrin}})^0 \ldots \quad (A' \text{ symmetry}) \quad (3)
\]

At the same time, \( S(p_{||}) \) and \( S(p_{\perp}) \) are perpendicular and parallel to the Co–C bond, respectively. Both \( \Psi_{\text{dirad} \perp} \) and \( \Psi_{\text{dirad} \parallel} \) wave functions have a dominant open-shell singlet character and were viewed as quasidegenerate, where singly occupied \( S(p_{||}) \) orbital might be associated with \( \pi \)-type interaction, while \( S(p_{\perp}) \) orbital are associated with \( \sigma \)-type interactions, respectively. On the other hand, the \( \Psi_{\text{ionic}} \) wave function has a dominant closed-shell character, and at 4.0 Å distance, the energy of ionic state predicted at the RASSCF level of theory is higher than that of the corresponding diradical state.

To provide a more accurate description of the energy of the ionic vs diradical states\(^{230}\) and how the energy difference varies with C⋯S distance (Scheme 4.1), dynamical correlation\(^{231}\) energy has been taken into consideration while performing CAS-type quantum calculations. To accomplish this task the electronic structure of the SM was analyzed employing the CASSCF framework followed by the CASPT2 perturbation theory. Based on the RASSCF calculations and the previous CASSCF calculations, different active spaces were tested and analyzed for SM, which resulted in the selection of 12 active electrons distributed in 13 active orbitals as the final active
space (Figure 4.4). The active space comprised of doubly occupied orbitals 3dₓz, 3dᵧz, 3dₓ²−ᵧ², and σₓCo−C, and the combination of four nitrogens lone pair orbitals. The respective correlating orbitals are 4dₓz, 4dᵧz, 4dₓ²−ᵧ², and σ⁎Co−C as well as the unoccupied 3dₓ in antibonding combination with nitrogen lone pairs that formed a σ-type donor–acceptor bond between corrin and cobalt. In addition, the active space comprised of the lowest unoccupied π⁎-corrin, one p orbital of sulfur (doubly occupied for the free thiolate), and the 4dₓ² orbital. The last one was a correlating orbital for the 3dₓ² orbital, after Co–C bond cleavage. The use of 4d orbitals as correlating counterparts for doubly occupied 3d orbitals (so-called the 3d double-shell effect) has been found necessary for the correct description of excited states and molecular properties in the CASPT2 calculations. The active orbitals selected for SM are similar to those which were used to describe the electronic properties of cob(I)alamin apart from σₓCo−C and sulfur orbitals. In Ψ┴dirad as well as in Ψionic, the sulfur orbital, S(pₓ), is perpendicular to corrin, while in Ψ‖dirad, the S(pᵧ) orbital is parallel to the ring. This leads to the inclusion of 8a’ and 5a” orbitals in the CAS space for the first two wave functions, while 7a’ and 6a” are in the third wave function. These active orbitals are summarized in Figure 4.4.
Figure 4.4. Orbitals in the active spaces for the $\Psi_{\perp \text{dirad}}$, $\Psi_{\parallel \text{dirad}}$, and $\Psi_{\text{ionic}}$ wave functions used for SM plotted at a C···S distance of 4.0 Å.

In the case of LM, the selection of active space was based on the CASSCF calculations of SM, as we already set a tool for the selection of orbitals based on RASSCF analysis. Due to all the restrictions regarding the choice of active space, several possible compositions have been tested, before leading to the conclusion that the most appropriate choice was
12 active electrons distributed in 11 active orbitals Figures 4.6. In particular, the active space for LM was comprised of $3d_{xz}$, $3d_{x^2-y^2}$, $3d_{yz}$, $\sigma_{Co-C}$, $S(p_z)$, and $S(p_y)$ orbitals of sulfur (lone pair for the free thiolate). The respective correlating orbitals were the lowest unoccupied $\pi^*$-corrin and $\sigma^*_{Co-C}$ orbitals and $4d_{yz}$, $4d_{x^2-y^2}$, $4d_{xz}$ to consider the double shell effect. It is important to note that the p orbital of sulfur (lone pair) of $-S-CH_3$ does not have a correlating counterpart, so the $\pi^*$-corrin orbital was selected. The use of an active space involves orbitals of sulfur as well as corrin, our main objective was to investigate whether the formation of the diradical state is energetically feasible by taking into account the close proximity between the MeCbl cofactor and the $-S-CH_3$ moiety. If it is feasible, then what would be the appropriate C···S distance to form the dominant diradical state over the ionic state? As in the case of SM, the dynamical correlation was also included for LM via second-order multireference perturbation theory based on the CASSCF wave function where the efficient implementation of the MC-XQDPT2 approach has been used.

4.3.3 Analysis of Ionic and Diradical States. In the case of SM, the SA-CASSCF calculations were performed with seven states for the ionic wave function and four states for the diradical followed by the multistate CASPT2 calculations. As CASSCF does not account for dynamic electronic correlation, the CASSCF states are often different from the correct order predicted by higher-level calculations, such as CASPT2. The diradical state was the lowest electronic state in accordance with CASSCF and CASPT2 calculations. However, the ionic state was found to be the lowest one in CASPT2 computations, but it was placed among higher excited states in CASSCF calculations. For this very reason, a large number of excited states were computed. At short C···S
distances (2.6 Å) for SM, the formation of a weak thiolate-carbon bond has been noticed, especially in case of Ψ_{dirad}^{\perp} state. The active orbitals of Ψ_{dirad}^{\parallel} are quite similar for different C⋯S distances. In Table 4.1 the major configurations of all three electronic states are presented in terms of weight percentage computed as a function of C⋯S distances. In the case of diradical states, the leading configuration has singly occupied sulfur and π* corrin orbitals, whereas in the case of ionic state, the sulfur orbitals are doubly occupied. The configurations with minor weights are produced mainly due to single and double excitations between σ_{Co−C} and σ_{Co−C}^{*} orbitals.

Table 4.1. Weight Percentage of Three Major Configurations (CSFs) in the Ψ_{dirad}^{\perp}, Ψ_{dirad}^{\parallel}, and Ψ_{ionic} Wave Functions for SM

<table>
<thead>
<tr>
<th>State</th>
<th>CSF1 (weight)</th>
<th>CSF2 (weight)</th>
<th>CSF3 (weight)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R(C−S) = 4.2 Å</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ψ_{dirad}^{\parallel}</td>
<td>2220000 u22d00 (80%)</td>
<td>2202000 u22d00 (6%)</td>
<td>22ud000 u22d00 (4%)</td>
</tr>
<tr>
<td>Ψ_{dirad}^{\perp}</td>
<td>222u0000 22d00 (81%)</td>
<td>220u2000 22d00 (6%)</td>
<td>22uud000 22d00 (3%)</td>
</tr>
<tr>
<td>Ψ_{ionic}</td>
<td>22220000 22000 (55%)</td>
<td>222ud000 22000 (11%)</td>
<td>222ud000 u2d00 (17%)</td>
</tr>
<tr>
<td>R(C−S) = 3.4 Å</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ψ_{dirad}^{\parallel}</td>
<td>2220000 u22d00 (76%)</td>
<td>2202000 u22d00 (6%)</td>
<td>22ud000 u22d00 (4%)</td>
</tr>
<tr>
<td>Ψ_{dirad}^{\perp}</td>
<td>222u0000 22d00 (79%)</td>
<td>220u2000 22d00 (6%)</td>
<td>222u0000 22d00 (3%)</td>
</tr>
<tr>
<td>Ψ_{ionic}</td>
<td>22220000 22000 (76%)</td>
<td>22022000 22000 (5%)</td>
<td>222ud000 22000 (4%)</td>
</tr>
<tr>
<td>R(C−S) = 2.6 Å</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ψ_{dirad}^{\parallel}</td>
<td>2220000 u22d00 (65%)</td>
<td>2220000 022200 (7%)</td>
<td>2202000 u22d00 (5%)</td>
</tr>
<tr>
<td>Ψ_{dirad}^{\perp}</td>
<td>222u0000 22d00 (74%)</td>
<td>2220u0000 22d00 (8%)</td>
<td>220u2000 22d00 (5%)</td>
</tr>
<tr>
<td>Ψ_{ionic}</td>
<td>22220000 22000 (74%)</td>
<td>22220000 u2d00 (7%)</td>
<td>222ud000 u2d00 (4%)</td>
</tr>
</tbody>
</table>

*Notation used to describe the CSFs follows the description of orbitals in the active spaces shown in Figure 4.4, where 2 means doubly occupied, u means spin up, d means spin down, and 0 unoccupied.*
The important feature of the electronic structure of SM is the energetic modulation of ionic and diradical states as a function of C···S distance (Scheme 4.1), shown in Figure 4.5 and Table A2. The energy of both states was computed assuming a fixed Co−C distance of 1.95 Å at which Co−C bond is not activated. The energetic variation of the electronic states needs to be interpreted with caution considering that the switching between ionic and diradical states occur (based on DFT). However, the potential energy surface of the ionic state shows greater sensitivity with respect to the C···S distance (Figure 4.5) than the diradical one. For a distance shorter than ~4.0 Å, the ionic state is energetically the lowest electronic state, while the diradical state becomes the lowest state at longer distances. The ionic state can be associated with S_N2-type displacement, where the cleavage of Co−C bond would be heterolytic in nature, while the diradical state with ET-type mechanistic route where it cleaves would be homolytic.

![Figure 4.5](image)

**Figure 4.5.** Energy curves for the \( \Psi^\perp_{\text{dirad}} \), \( \Psi^\parallel_{\text{dirad}} \), and \( \Psi_{\text{ionic}} \) wave functions for SM computed at CASSCF/CASPT2 level of theory.
In order to describe the electronic structure of the LM, a series of SA-CASSCF calculations testing different numbers of electronic states has been performed. It was concluded that only three states are sufficient to describe the complex electronic structure of LM. These three states include an ionic state ($\Psi_{\text{ionic}}$), which is closed-shell, and two diradical states ($\Psi_{\text{dirad}}$ and $\Psi_{\text{dirad}}^{\|}$), transferring of an electron either from S(p$_z$) or S(p$_y$)
respective orbital to \( \pi^* \) corrin. Thus, the electronic structure of LM in terms of electronic states could be described as follows:

\[
\Psi_{\text{ionic}} \approx \ldots (S_{\perp})^2 (S_{\parallel})^2 \left( \pi^*_{\text{corrin}} \right)^0 \ldots \]  
(1)

\[
\Psi_{\perp \text{dirad}} \approx \ldots (S_{\perp})^1 (S_{\parallel})^2 \left( \pi^*_{\text{corrin}} \right)^1 \ldots \]  
(2)

\[
\Psi_{\parallel \text{dirad}} \approx \ldots (S_{\perp})^2 (S_{\parallel})^1 \left( \pi^*_{\text{corrin}} \right)^1 \ldots \]  
(3)

In order to determine correlation between ionic and diradical states as in SM, the energy modulation was analyzed as a function of C···S distance (Figure 4.7) in which the structure of the cofactor remains the same for all distances. All three states show a monotonic change in energy: \( \Psi_{\text{ionic}} \) is the lowest energy state followed by \( \Psi_{\perp \text{dirad}} \) and then \( \Psi_{\parallel \text{dirad}} \) at a short C···S distance (2.6 Å). The main feature of electronic structure of LM is the switch between ionic and diradical states as has been observed for SM, where the ionic state has been found to be the lowest state for C···S distances shorter than 4.0 Å, while the diradical state is the lowest one for C···S distances equal or greater than 4.0 Å. The presence of solvent (water) environment does not affect the energetic behavior of ionic and diradical states, which has been confirmed via CASSCF calculation in PCM. This finding is in agreement with the previous QM(DFT)/MM calculations,\(^{116}\) wherein the solvent was not found to impact the energetic of the electronic states.
Figure 4.7. Energy curves for the $\Psi_{\text{dirad}}^\perp$, $\Psi_{\text{dirad}}^\parallel$, and $\Psi_{\text{ionic}}$ wave functions for LM computed at CASSCF/MC-XQDPT2 level of theory.

The major configuration state functions (CSFs) for each electronic state of LM have been shown in Table 4.2. The energetic behavior of LM and SM for ionic as well as diradical states is almost similar. However, there is a small change in weight of CSFs between two structural models, due to the lack of symmetry in LM. The main difference in the weight of CSFs is noticed in the ionic state (more specifically in CSF1) of two models calculated at $C\cdots S = 4.2 \, \text{Å}$. The ionic state (of CSF1) of SM has a lower weight (55%) as compared to LM; nonetheless, CSF1 is the most dominating weight of ionic state of each model. In case of SM, this distance (4.2 Å) is close to the crossing of ionic and diradical states, and strong mixing between states can take place at a crossing point, leading to the modification of CSFs weight of ionic state. Whereas, in case of LM, crossing of two
states occurs at short distances (≈3.9 Å), which does not affect the weight of CSFs at 4.2 Å. In addition, the participation of CSF2 with higher weight (≈10%) in LM (for diradical states) at 3.4 and 4.2 Å is a result of inclusion of second p orbital of sulfur, which can mix efficiently due to a lack of symmetry constraints in LM.

Table 4.2. Weight Percentage of Three Major Configurations (CSFs) in the $\Psi_{\text{dirad}}^{\perp}$, $\Psi_{\text{dirad}}^{\parallel}$, and $\Psi_{\text{ionic}}$ Wave Functions for LM$^a$

<table>
<thead>
<tr>
<th>State</th>
<th>CSF1 (weight %)</th>
<th>CSF2 (weight %)</th>
<th>CSF3 (weight %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Psi_{\text{dirad}}^{\parallel}$</td>
<td>22222100000 (79%)</td>
<td>22221200000 (10%)</td>
<td>22202110000 (6%)</td>
</tr>
<tr>
<td>$\Psi_{\text{dirad}}^{\perp}$</td>
<td>22222100000 (79%)</td>
<td>22221200000 (10%)</td>
<td>22202110000 (6%)</td>
</tr>
<tr>
<td>$\Psi_{\text{ionic}}$</td>
<td>22222200000 (88%)</td>
<td>22202200000 (7%)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>State</th>
<th>CSF1 (weight %)</th>
<th>CSF2 (weight %)</th>
<th>CSF3 (weight %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Psi_{\text{dirad}}^{\parallel}$</td>
<td>22222100000 (76%)</td>
<td>22221200000 (9%)</td>
<td>22202110000 (6%)</td>
</tr>
<tr>
<td>$\Psi_{\text{dirad}}^{\perp}$</td>
<td>22222100000 (74%)</td>
<td>22221200000 (10%)</td>
<td>22202110000 (6%)</td>
</tr>
<tr>
<td>$\Psi_{\text{ionic}}$</td>
<td>22222200000 (85%)</td>
<td>22202200000 (7%)</td>
<td>22221010000 (2%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>State</th>
<th>CSF1 (weight %)</th>
<th>CSF2 (weight %)</th>
<th>CSF3 (weight %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Psi_{\text{dirad}}^{\parallel}$</td>
<td>22222100000 (82%)</td>
<td>22221200000 (4%)</td>
<td>22202110000 (2%)</td>
</tr>
<tr>
<td>$\Psi_{\text{dirad}}^{\perp}$</td>
<td>22222100000 (82%)</td>
<td>22221200000 (4%)</td>
<td>22202110000 (2%)</td>
</tr>
<tr>
<td>$\Psi_{\text{ionic}}$</td>
<td>22222200000 (88%)</td>
<td>22221200000 (3%)</td>
<td>22202200020 (2%)</td>
</tr>
</tbody>
</table>

$^a$ Notation used to describe CSFs follows description of orbitals in the active spaces shown in Figure 4.6 where 2 means doubly occupied, 1 means singly occupied, and 0 means unoccupied.

Indeed, the lowest electronic state of LM reaction complex is the diradical state at C···S distance of 4.0 Å, which confirms the DFT results showing antiferromagnetic coupling between the p orbital of S and the $\pi^*$ orbital of corrin. Therefore, the results based on ab initio theory for both models are entirely in agreement with previous DFT study.115
4.3.4 ET versus SN2-Type Mechanism. The calculations carried out in this chapter for SM and LM demonstrate that the computed C···S distance can tune energetic behavior of the ionic and diradical states. The main difference in the two types of computations is that all three electronic states can be observed in an active space of LM, which includes all essential orbitals forming ionic and diradical states using the CASSCF/QDPT2 approach. Whereas in case of the CASSCF/CASPT2 method used for SM comprised of symmetry plane, depending upon electronic state (ionic or diradical) of interest, the corresponding wave function can be used for active space used for each state). Two possible mechanistic pathways for methyl transfer can be envisioned depending on the type of wave function used to describe the electronic configuration of the system. The ionic-type closed-shell wave function is used to describe SN2-type methyl transfer consistent with [CoIII(corrin)]−Me+···−SMe electronic configuration, where the cleavage of the Co−C bond occurs in a heterolytic fashion. However, the promotion of one electron from −S−CH3 to corrin gives rise to [CoIII(corrin•)]−Me···•SMe configuration, whose description requires an open-shell singlet wave function. The appearance of this diradical state seems to be promoted by the presence of high-lying molecular orbitals of −S−Me, which are energetically very close to π*corrin. In this pathway, which is referred to as ET, the Co−C bond is cleaved homolytically.

The observed diradical state at longer C···S distance is indicative of the fact that ET-type mechanistic pathway can account for efficient methyl transfer. The ET-based reductive cleavage pathway does not impose specific geometrical and distance constrains with respect to substrate and cofactor, as does SN2, which may be important from an enzymatic point of view.
4.4 Conclusion and Summary

In this chapter, the electronic structure of the reactant complex has been extensively explored at CASSCF/QDPT2 (for large model) and CASSCF/CASPT2 (for small model) levels of theory, respectively. In spite of the structural differences, both the employed models representing reactant complex mimic of cofactor:substrate provide the same conclusion regarding the modulation of two electronic states (ionic and diradical states), i.e., at longer C···S distances the diradical state of the complex is the lowest electronic state. This implies that ET-based reductive cleavage is more in line at 4.0 Å C···S distance than that of $S_N$2-type. In addition, the major contribution of the electronic state represented their weight percentage which also strengthened our hypothesis, as one state is always dominant over the other for a particular C···S distance.

The computational results of this chapter could also help in understanding some of the critical aspects related to the enzymatic methyl transfer reactions catalyzed by MeCbl cofactor. The key question concerned with the methyl transfer reaction is at what distance the Co–C bond is cleaved and at what distances the C···S bond is formed, i.e., how closely the two reacting modules interact with each other to form the reactant complex (Scheme 4.1). Unfortunately, such structural information is not available from existing X-ray crystallographic data. DFT calculations predicted that in the $S_N$2-type transition state, the C···S distance should be ~2.8 Å to facilitate the methyl transfer from cofactor to substrate. Such a short distance may not be effective from the enzymatic point of view or may not be even feasible due to possible strong steric interactions among the enzyme backbones, thus making the actual pathway more consistent with ET-based mechanistic proposal. Moreover, it has been proposed experimentally in acetyl-CoA synthase\textsuperscript{223} that
the traditionally discussed S$_{N}$2-type mechanism is not appropriate at longer distances. On the other hand, if the ET pathway is considered, then methyl transfer can effectively take place even at longer distances. Even in previous DFT and QM/MM studies,\textsuperscript{115,116} it has been demonstrated that the cleavage of the Co–C bond takes place via [Co$^{\text{III}}$(corrin$^-$)]–Me···S–Me diaradical configuration. Although the S$_{N}$2/ET dichotomy was discussed in the context of many chemical reactions,\textsuperscript{186-188} such a connection has never been made in the case of B$_{12}$-mediated methyl transfer reactions. Consequently, the work in this chapter opens up a new horizon to further investigate the proposed mechanisms using experimental techniques.
CHAPTER 5

ELECTRONIC STRUCTURE OF ONE-ELECTRON OXIDIZED FORM OF THE METHYLCOBALAMIN COFACTOR

5.1 Introduction

Methylcobalamin (MeCbl,) is the cofactor of several methyltransferases such as methionine synthase (MetH) or enzymes that participate in the carbon dioxide fixation pathway. The enzymatic mechanisms for the methyl transfer require the cleavage and the formation of a Co–C bond in the presence of an additional cofactor, either a Zn$^{2+}$ ion or a $\text{[Fe}_4\text{S}_4\text{]}$ cluster. In solution, MeCbl is a very sluggish catalyst in contrast to an enzyme-bound MeCbl where transfer of the methyl group is very efficient. The methyl transfer mechanism has been commonly interpreted as an $\text{S}_\text{N}2$-type nucleophilic displacement; however, other proposals including one-electron reduction or oxidation of cobalt corrinoids were also considered. The Co–C bond cleavage in the one-electron-reduced (1e-Red) form of the cobalt corrinoids has attained a considerable amount of interest because the Co–C bond dissociation energy (BDE) is significantly decreased. The 1e-Red $\text{B}_{12}$ cofactors (Figure 5.1) and related model complexes have been investigated experimentally. For MeCbl

---

3 The material presented in this Chapter was reprinted (adapted) with permission from Kumar, N.; Kuta, J.; Galezowski, W.; Kozlowski, P. M. Inorg. Chem. 2013, 52, 1762. Copyright 2013 American Chemical Society.
cofactor$^{115,116}$ (Figure 5.1), it has been found using density functional theory (DFT) that the Co–C bond was stretched, the electron shifted from the corrin-based $\pi^*$ to bond-centered $\sigma^*_{\text{Co-C}}$ orbital, which results in the cleavage of the three-electron $(\sigma^2(\sigma^*)^1$ bond.$^{115}$ A similar DFT analysis of the Co–C bond cleavage in the neutral as well as in the 1e-Red form of alkyl-cobalt(III) phthalocyanines$^{146}$ confirmed the validity of the reductive cleavage mechanism proposed for the MeCbl.$^{115,116}$

![Figure 5.1. Molecular structure of B$_{12}$ cofactors (left panel) where R = Me for MeCbl and R = Ado for AdoCbl ($R_1 = \text{CH}_2\text{CONH}_2$, $R_2 = \text{CH}_2\text{CH}_2\text{CONH}_2$, $R_3 = (\text{CH}_2)_2\text{CONHCH}_2\text{CH(\text{CH}_3)}\text{OPo}_3^{-}$). The right panel shows the structural model of MeCbl employed in this chapter and denoted as Im-[Co$^{\text{III}}$(corrin)]-Me$^+$.](image)

Despite the fact that the oxidative mechanism was considered a long time ago as one of the possible routes for alkyl transfer from alkylcobalamins,$^{238}$ the Co–C bond cleavage of one-electron-oxidized (1e-Ox) cobalt corrinoids has gained less attention compared to the reductive cleavage. The electronic and structural properties of species...
generated upon oxidation remain unexplored, mainly because these species are unstable and difficult to characterize experimentally. It is noteworthy that a quite rapid Co–C bond cleavage was observed during electrochemical oxidation of MeCbl in aqueous solution with aquoCo(III)cobalamin and methanol as the final products. It has been further suggested that one-electron oxidation leads to depopulation of the Co–C σ orbital, which consequently would reduce the bond strength. The resulting species then would have d⁵ cobalt(IV) character, and the BDE of the Co(IV)–C bond should be significantly lower in comparison to Co(III)–C. As an alternative to the Co(IV) state, the corrin ligand could be oxidized leading to a corrin–based radical in analogy to the heme-containing systems where usually the porphyrin ligand is oxidized.

Initially, the formation of Co(IV)–C and π-cation radicals with Co(III)–C species in 1e-Ox organocobalt complexes had been thoroughly investigated by Halpern et al., where they distinguished these two electronic species using the superhyperfine structure of the EPR spectrum. It was demonstrated for 1e-Ox methylcobaloximes that the odd electron resides mainly on the cobalt d orbital which is not directed toward the donor N atoms, thus giving a dominant contribution to the singly occupied molecular orbital (SOMO), with a small contribution from the cobalt 4pₓ orbital. A similar location of the unpaired electron was also suggested for the 1e-Ox alkylcobaltporphyrin. Although these studies established that the 1e-Ox alkylcobalt complexes are usually formulated as Co(IV) species, it was further put into question if the oxidized organocobalt porphyrins (which are structurally related to corrinoids) should
be described as the ligand centered $\pi$-cation radicals Co(III) rather than the Co(IV) species.\textsuperscript{237,249} On the other hand, the electronic properties of the organocobalt porphyrins\textsuperscript{249} imply that the 1e-Ox cobalt corrinoids may have mixed $\pi$-cation corrin radical and Co(III)/Co(IV) character.

Almost two decades later, Fukuzumi et al. pointed out such dichotomy for organocobalt tetraphenylporphyrins, \([B-(TPP)Co-R]^+\), where they studied the role of different axial ligands.\textsuperscript{248} They observed that the oxidized species for both $R = \text{Me}$ and phenyl have a Co(III) $\pi$-cation radical character in dichloromethane (with no respectable B), but in weakly binding acetonitrile, the \([\text{MeCN}-(TPP)Co-\text{Me}]^+\) species has more Co(IV) character than its \([\text{MeCN}-(TPP)Co-\text{Ph}]^+\) analogue. Additionally, with a stronger axial base (B), pyridine, both complexes have a dominant Co(IV) character. Furthermore, it has been suggested that the homolytic cleavage of the Co–C bond in the Co(IV) complexes is easier than in the $\pi$-cation radicals with Co(III).\textsuperscript{248} The formation of Co(IV) species in oxidized coenzyme B\textsubscript{12} model complexes, i.e., cobaloximes, \([B-Co^{\text{IV}}(DH)_{2}-R]^+\) and subsequent Co(IV)–C bond cleavage have been further observed by Fukuzumi and Ohkubo,\textsuperscript{142} wherein they revealed that regardless of the mechanism, the Co–C bond cleavage is facilitated by the presence of an axial base. Alternatively, Fanchiang et al.\textsuperscript{250} proposed that the 1e-Ox MeCbl species is involved in the methyl transfer between MeCbl and diaquo-Co(III)cobinamide. It was suggested that within a bimolecular complex, the diaquo-Co(III)cobinamide form of the cofactor oxidizes MeCbl and subsequently forms Co(II)cobinamide, which serves as a methyl acceptor.\textsuperscript{250} This particular methyl transfer reaction demonstrates that the oxidative cleavage of the Co–C bond might have relevance to the enzymatic catalysis. In addition, it has been
investigated using kinetic measurements of various reactions between the methylcobalamin and oxidizing agents such as AuCl\textsubscript{4}\textsuperscript{−}, IrCl\textsubscript{6}\textsuperscript{2−}, and Fe\textsuperscript{3+} ions that a single electron transfer from MeCbl to these oxidants is feasible which initiates the cleavage of the Co–C bond.\textsuperscript{123,125,251-253} These experimental findings are more in line with oxidative cleavage mechanism of the Co–C bond.

The coordination state is very important for the chemistry of corrinoids to understand the efficient cleavage of the Co–C bond. The MeCbl cofactor has a dimethylbenzimidazole (DBI) axial base coordinated to the cobalt in solution (except for acidic conditions) as well as in the enzymatic environment where this intramolecular base is replaced by a histidine residue from the protein.\textsuperscript{41} Since the MeCbl cofactor preferably exists as a six-coordinate 18-electron species, its 1e-Ox analogue must show even stronger demand for an axial base. This is the reason why the Co–C bond cleavage in the 1e-Red species involves a five-coordinate base-off form, as it was concluded in the recent DFT study,\textsuperscript{115,116} while the 1e-Ox form requires a base at the lower axial position.

On the basis of the previous studies, one can conclude that the electronic structure of the 1e-Ox cobalt corrinoids is a subject of debate, because the exact location of the unpaired electron remains elusive. Herein, we employed two different types of computational tools to analyze the electronic and structural properties of the 1e-Ox MeCbl. We applied a DFT for the 1e-Ox MeCbl and its related metalloporphyrins in order to gain insight into the Co–C bond properties. Subsequently, the time-dependent DFT (TD-DFT) calculations were carried out to investigate the nature of low-lying excited states and validate the oxidative cleavage mechanism. The complex electronic structure of the 1e-Ox MeCbl revealed by the DFT was further analyzed using the
CASSCF/MC-XQDPT2 calculations. A particular emphasis was placed on the Co–C bond cleavage in the 1e-Ox form of MeCbl and the relevant mechanism responsible for the lowering of the dissociation energy as compared to its neutral analogue.

5.2 Computational Details

5.2.1 Structural Models. A simplified model of MeCbl, denoted as Im-[Co\textsc{iii}(corrin)]-Me\textsuperscript{+} (Figure 5.1, right panel), has been employed where all the amide side chains were replaced with hydrogen atoms, and the lower axial ligand was modeled with the imidazole (Im) moiety. These types of truncated models have been extensively used in previous studies, with success, to predict the electronic and structural properties of B\textsubscript{12} cofactors.\textsuperscript{75-77} Consequently, the truncated model of MeCbl has been used to generate 1e-Ox MeCbl species. In order to gain further insight into the Co–C bond properties of the 1e-Ox MeCbl species, we have also utilized structurally related porphine-based complexes, B-[Co\textsc{iii}(Por)]-Me with pyridine or imidazole as an axial base (B).

5.2.2 DFT Calculations. The DFT method has been used due to its effectiveness in the investigation of the electronic and structural properties of the alkyl-cobalt(III) complexes, including cobalt corrinoids.\textsuperscript{75-77,172,254} The use of a suitable functional is important for the accurate prediction of the Co–C bond dissociation energy. Thus, the geometry optimization of all models as well as electronic structure calculations has been carried out using the BP86\textsuperscript{67,68} functional and the 6-31G(d) (5d components) basis set, as implemented in Gaussian 03.\textsuperscript{255} It has been pointed out for the 1e-Ox nitrosyl cobaltporphyrin that the B3LYP functional gives spin density distribution which is similar to the experimental data as compared to the BP86 functional.\textsuperscript{256} In addition, it has been demonstrated recently that the spin polarized solution of cob(I)alamin
complexes$^{80,145}$ cannot be obtained with pure GGA functionals such as PBE$^{78}$ because they fail to converge to a spin polarized solution at the ground state singlet. Despite the fact that the hybrid B3LYP functional underestimates the strength of the Co–C bond due to the presence of HF exchange,$^{79}$ it is appropriate for the prediction of a broken symmetry solution.$^{81}$ Therefore, we applied the B3LYP functional as well for all models.$^{82}$ The relevant geometrical parameters used for all optimized structures employed in present work can be found in the Appendix (Tables, A3 and A4).

5.2.3 CASSCF/MC-XQDPT2 Calculations. The DFT computations, which are based on the Kohn–Sham formalism, cannot describe the multiconfigurational character as well as near-degeneracy problems of the electronic states of the MeCbl model system because they are restricted to a single Slater determinant description. Thus, we performed multireference CASSCF calculations, followed by quasi-degenerate perturbation theory (QDPT2)$^{92}$ calculations with a multiconfigurational self-consistent-field reference function (MC-XQDPT2) to include the dynamical correlation, as implemented in the PC GAMESS/Firefly QC package.$^{257}$ The CASSCF calculations were carried out on the DFT optimized structure of the 1e-Ox MeCbl model using the 6-31G(d) basis set.

5.3 Results and Discussion

5.3.1 Geometry of the 1e-Ox Form of MeCbl and Related Porphine Complexes. At first, the previously used structural model of MeCbl, i.e., Im-[Co$^{\text{III}}$(corrin)]-Me$^+$, was optimized employing BP86 and B3LYP functionals. The fact that this level of theory (DFT) provides axial bond lengths comparable with X-ray data$^{258}$ (Table 5.1) strengthens our confidence that the structure of 1e-Ox MeCbl species would also be well described with these functionals. The optimization of the 1e-Ox MeCbl reveals that its key
The removal of an electron has small influence on axial bond lengths; the Co–C bond elongates from 1.968 to 2.011 Å while the Co–N_{Im} bond contracts from 2.132 to 2.063 Å (Table 5.1) as noticed with the BP86 functional. A similar trend of elongation and shortening of the axial bonds was also observed with the B3LYP functional. Furthermore, a very small change of the corrin structure was obtained because the corrin ring is very robust and not influenced by the nature of the axial ligand. It is important to mention that these changes in axial bond lengths of the 1e-Ox MeCbl are more pronounced than that of the 1e-Red MeCbl species where the single electron is mainly located on the corrin ring. Although the modulation of axial bond lengths seems to be consistent with some Co(IV) character of the 1e-Ox MeCbl, the bond lengthening is not substantial enough to signify the complete removal of an electron from the Co–C \( \sigma \) bond. The \( \sim 0.05 \) Å longer Co–C bond in 1e-Ox species than in the 1e-Red MeCbl goes somehow against chemical intuition (shorter bonds are usually stronger), as the latter complex has a lower Co–C BDE (Table 5.1). This can be further explained in terms of the coordination of axial base.

Table 5.1. Relevant Co–C BDEs (kcal/mol) and Axial Bond Lengths (Å) of the Neutral, 1e-Ox, and 1e-Red MeCbl Models Computed at the BP86/6-31G* and B3LYP/6-31G* Levels of Theory

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Experiment</th>
<th>BP86</th>
<th>B3LYP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Neutral</td>
<td>Neutral 1e-Ox 1e-Red</td>
<td>Neutral 1e-Ox</td>
</tr>
<tr>
<td>Co-C</td>
<td>1.979\textsuperscript{258}</td>
<td>1.968 2.011 1.958</td>
<td>1.956 2.031</td>
</tr>
<tr>
<td>Co-N\textsubscript{ax}</td>
<td>2.162\textsuperscript{258}</td>
<td>2.132 2.063 2.112</td>
<td>2.179 2.050</td>
</tr>
<tr>
<td>BDE\textsuperscript{a}</td>
<td>37(3)\textsuperscript{65}</td>
<td>37.0 24.2 20.4\textsuperscript{b}</td>
<td>24.1 16.1\textsuperscript{b}</td>
</tr>
<tr>
<td>BDE lowering (%)</td>
<td>–</td>
<td>– 35 45</td>
<td>– 33</td>
</tr>
</tbody>
</table>

\textsuperscript{a} ZPE corrected. \textsuperscript{b} BDE of the base-off species.
Due to the lack of extensive experimental data of the 1e-Ox MeCbl species, we employed structurally related porphyrin-based complexes, B-[Co^{III}(Por)]-Me (Figure 5.4), for DFT calculations as they are typically used in experimental studies with different bases B (pyridine or imidazole).\(^{248}\) Moreover, it would also be interesting to see if any specific features of the corrin ligand, important for the Co–C bond cleavage in the 1e-Ox species, could be identified in comparison with the porphine complexes. The optimized Co–C bond of porphyrins is a few hundredths of an Ångström shorter than that observed for the corrin model (Table 5.2). The axial bond distances obtained with the BP86 functional are also shorter than the corresponding B3LYP values, as observed for the 1e-Ox MeCbl. This is mainly due to the difference in the metal–ligand bond covalency predicted by these two functionals. However, the symmetry of the axial system is lower in the BP86 than B3LYP structure because in BP86 the C–Co–N\(_B\) angle is \(\approx 170^\circ\), while it is close to \(180^\circ\) in the B3LYP optimized structure. In addition, the porphine ligand shows saddling which is considerably more extensive in the BP86 than the B3LYP optimized structures of the 1e-Ox B-[Co^{III}(Por)]-Me. This type of structural difference is absent in the 1e-Ox MeCbl where the geometry of corrin is virtually insensitive to the type of functional used.

Table 5.2. Relevant Geometrical Parameters of the 1e-Ox B-[Co^{III}(Por)]-Me (B = Im or Pyr) Optimized at the BP86/6-31G* and B3LYP/6-31G* Levels of Theory\(^{a}\)

<table>
<thead>
<tr>
<th>Axial base (B)</th>
<th>functional</th>
<th>Co-C (Å)</th>
<th>Co-N(_B) (Å)</th>
<th>C-Co-N(_B) (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Im</td>
<td>BP86</td>
<td>1.989</td>
<td>2.050</td>
<td>170.5</td>
</tr>
<tr>
<td>Im</td>
<td>B3LYP</td>
<td>1.992</td>
<td>2.063</td>
<td>178.4</td>
</tr>
<tr>
<td>Pyr</td>
<td>BP86</td>
<td>1.951</td>
<td>2.088</td>
<td>171.1</td>
</tr>
<tr>
<td>Pyr</td>
<td>B3LYP</td>
<td>1.991</td>
<td>2.108</td>
<td>178.8</td>
</tr>
</tbody>
</table>

\(^{a}\) The present work.
5.3.2 Electronic Structure of the 1e-Ox MeCbl. To investigate the electronic structure of the 1e-Ox species, the Kohn–Sham orbitals located near the HOMO–LUMO gap were extracted from the DFT calculations. The energy levels as well as the location of MOs essential for the Co–C bond rupture are shown in Figure 5.2. Within the simple three-orbital framework, the so-called strain orbital is of particular importance for predicting the properties of alkylcobalt complexes.

$$\sigma_{\text{Co-C}} \approx - c_1 \phi[N(p_z)] + c_2 \phi[\text{Co}(d_{z^2})] + c_3 \phi[C(p_z)]$$ (1)

This orbital has Co–C bonding character but is antibonding with respect to cobalt and nitrogen. In tetrapyrrolole complexes, there is an interaction of this orbital with the $\pi$ HOMO of the ligand. The ligand HOMO bears a resemblance to the $a_{2u}$ orbital of the porphine dianion of $D_{4h}$ symmetry and has large contributions not only from the $p_z$ pyrrole nitrogen but also from the $p_z$ orbitals of the bridging carbon atoms.

Figure 5.2. Orbital energy level diagram for 1e-Ox MeCbl: (a) BP86/6-31G* level of theory and (b) B3LYP/6-31G* level of theory. SOMO is indicated in green; the $\sigma^*$ energy level is indicated in blue.
The SOMO shown in Figure 5.2 is the antibonding combination of an energetically higher \( a_{2u} \)-like \( \pi \) orbital of the equatorial ligand and a small contribution from the lower energetically lying \( \sigma_{\text{Co-C}} \). Thus, the SOMO orbital has only weak Co–C bonding character:

\[
\text{SOMO} \approx \phi[\pi_{\text{corrin}}] - \sigma_{\text{Co-C}} \quad (2)
\]

The bonding combination of \( \sigma_{\text{Co-C}} \) and the \( \pi \)-ligand orbital has a large contribution from the \( \sigma_{\text{Co-C}} \). This orbital, HOMO-4 for B3LYP in Figure 5.2b, and HOMO-6 and HOMO-7 for BP86, is \(~1.5\) eV lower in energy than the SOMO. Another orbital resulting from the three-orbital approach, depicted in Figure 5.2, is the higher energy antibonding \( \sigma \) of the Co–C bond (\( \sigma_{\text{Co-C}*} \)). This orbital has double antibonding character because it is antibonding with respect to cobalt and nitrogen as well as with respect to cobalt and carbon (eq 3) and is energetically too far from the HOMO to play an important role in the Co–C bond cleavage in 1e-Ox methyl-cobalt species.

\[
\sigma_{\text{Co-C}*} \approx -c_1\phi[N(p_z)] + c_2\phi[\text{Co}(d_z^2)] - c_3\phi[C(p_z)] \quad (3)
\]

The orbital energies and the composition of most important molecular orbitals from BP86 and B3LYP calculations give a very similar picture of the SOMO orbital with respect to corrin as well as imidazole axial base contributions placing them at 72% (corrin) and 11% (Im) based on BP86, and 73% (corrin) and 9% (Im) according to B3LYP (Table 5.3). Nevertheless, the involvement of carbon and cobalt in the SOMO orbital is predicted to be somewhat different: BP86 gives more Co character (12%) and less carbon (4%), while the prediction based on the B3LYP is opposite, with a smaller contribution coming from cobalt (7%) and larger from carbon (9%).
The composition of the SOMO (Table 5.3) indicates that the 1e-Ox complex does not have a true one electron Co–C bond. The bonding electron pair in the low-energy bonding combination of $\sigma_{\text{Co-C}}$ (eq 1) and the $a_{2u}$-like orbital remains intact in the equilibrium structure of the 1e-Ox Im-[Co$^{\text{III}}$(corrin)]-Me$^+$ or its porphine 1e-Ox B-[Co$^{\text{III}}$(Por)]-Me analogues. Some depletion of the electron density from Co and the methyl group results from the absence of a $\beta$ electron in the HOMO, but this orbital has only a marginal contribution to the Co–C bonding. Therefore, the Co–C bond elongation in the equilibrium structure of the 1e-Ox system is moderate. The spin density distributions in 1e-Ox B-[Co$^{\text{III}}$(Por)]-Me (B = Im or Pyr), which are similar to that of the 1e-Ox Im-[Co$^{\text{III}}$(corrin)]-Me$^+$, suggest that the SOMO in these species can also be described as an antibonding combination of an $a_{2u}$-like porphine $\pi$ orbital and $\sigma_{\text{Co-C}}$ (eq. 2).

5.3.3 Spin Density Profiles in the 1e-Ox Models of MeCbl and Related Porphine Complexes. In the optimized structures of the 1e-Ox models, the unpaired electron is mainly distributed between cobalt and the corrin, or porphine, moiety (Figures 5.3, 5.4, and 5.5). The $\alpha$ spin density of the macrocycle is delocalized at the meso carbons as well as at the pyrrole nitrogens. Furthermore, the extent of spin distribution depends on the
functional: spin density is evenly distributed between Co and corrin in the case of BP86, while it shifts more toward the corrin ring in the case of B3LYP. However, it turns out that the spin density on cobalt is not controlled by the SOMO in BP86-based optimized models of MeCbl as well as in its porphine analogues. The BP86 computed total spin densities on cobalt: 39% in the 1e-Ox Im-[Co$^{\text{III}}$(corrin)]-Me$^+$ or 43% and 44% in the 1e-Ox B-[Co$^{\text{III}}$(Por)]-Me$^+$ (B = Pyr or Im), respectively, are larger than the composition of the SOMOs (Table 5.3). This is noticeably a case of spin polarization, which usually emerged in DFT calculations of the electronic structures of paramagnetic metal complexes with red-ox active ligands. The spin polarization occurs when at a given molecule spin contributions from $\alpha$ and $\beta$ electrons, excluding the one residing in the SOMO, do not cancel. Indeed, the close examination of unrestricted orbitals shows that the spatial parts of matching $\alpha$ and $\beta$ electrons differ significantly from each other. The spin polarization accumulates $\beta$ spin on the corrin and $\alpha$ spin on the cobalt. As a result, net $\alpha$ spin density over the corrin is reduced and subsequently transferred to cobalt. On the other hand, the B3LYP calculations of the above-discussed model systems give $\alpha$ spin density of 7%, 5%, and 5%, respectively, on cobalt and 77%, 87%, and 86% over their equatorial ligands; therefore the spin polarization is small, which means that the SOMO describes correctly the spin delocalization.
Figure 5.3. Spin density for the 1e-Ox MeCbl: (a) BP86/6-31G* level of theory and (b) B3LYP/6-31G* level of theory.
We believe that the spin polarization has little connection with the formal oxidation state of the metal; i.e., both functionals show the 1e-Ox species as a prevalently
Co(III) ligand π radical, with some admixture of the Co(IV) character. In the case of a strong spin polarization, the charge on cobalt is probably more reliable than spin density. The BP86 calculated Mulliken charge on the cobalt is +0.613 for 1e-Ox MeCbl, which is marginally larger than +0.559 of its neutral analogue. Because the B3LYP functional puts more emphasis on the ionic character of the bond due to inclusion of HF exchange, the computed charges are a bit larger (i.e., +0.696 and +0.629, respectively), but the difference caused by one-electron oxidation is similar to BP86, consistent with a small admixture of the d^5 Co(IV) state. Regardless of the methodology, the electronic charges of the metal in coordination compounds are much different from the formal oxidation states of the metal. However, the trends in atomic charges still reflect the change in the formal oxidation state of the metal in MeCbl, especially since the coordination state of the cobalt remains unchanged.

### 5.3.4 Dissociation of the Co–C Bond in the 1e-Ox MeCbl

The Co–C bond dissociation energy (BDE) computed as the energy difference between the fully optimized structure of Im-[Co^{III}(corrin)]-Me^+ and its separate fragments, i.e., Im-[Co^{II}(corrin)]^+ and the methyl radical, gives a reliable estimation of the Co–C bond strength. The computed BDE with the BP86 for the neutral cofactor has a value of 42.0 kcal/mol, while zero point energy (ZPE) correction is equal to 5.0 kcal/mol. The ZPE corrected dissociation energy, i.e., 37.0 kcal/mol (see Table 5.1), is thus in excellent agreement with experimental values for the MeCbl of 37 ± 3 based on thermolysis^{20} and 36 ± 4 kcal/mol employing calorimetric measurements^{261} respectively. On the other hand, the B3LYP functional, which is known to underestimate BDE,^{115,172,262,263} gives a
lower value of 29.5 kcal/mol, and after subtracting a ZPE correction, i.e., 5.4 kcal/mol, this value further decreases to 24.1 kcal/mol.

To analyze the cobalt–carbon bond dissociation in the 1e-Ox MeCbl model, the geometry optimization was first carried out along the Co–C stretch to verify that the Im-[Co^{III}(corrin)]^{-}-Me^{2+} complex can dissociate properly. It was assumed that the cleavage is homolytic and leads to the formation of the methyl radical. The energy optimization along the Co–C coordinate produces a smooth dissociation curve with significant lowering in energy as compared to the neutral MeCbl (Figure 5.6).

![Image](Image)

Figure 5.6. Comparison of dissociation curves for neutral and 1e-Ox forms of MeCbl calculated using the BP86/6-31G* level of theory.

The cleavage of the Co–C bond was further explored by analysis of the low-lying excited states in addition to the ground state (Figures 5.7 and 5.8). The potential energy curves representing the lowest electronic excited states do not show any state crossing for either
of the functionals, which indicates that the cleavage event involves only one electronic in contrast to the 1e-Red species.

Figure 5.7. BP86/6-31G* calculated: (a) the ground state and low-lying electronic excited states for 1e-Ox MeCbl as a function of Co–C bond length and (b) spin density changes along the Co–C bond length.
Figure 5.8. B3LYP/6-31G* calculated: (a) the ground state and low-lying electronic excited states for 1e-Ox MeCbl as a function of Co–C bond length and (b) spin density changes along the Co–C bond length.

The accurate estimation of the Co–C BDE in the 1e-Ox model of MeCbl can be obtained as the energy difference between the full complex and its fragments according to the equation:

$$\text{Im-}[\text{Co}^{III}(\text{corrin})]^*\cdot\text{Me}^{2+} \rightarrow \text{Im-}[\text{Co}^{III}(\text{corrin})]^{2+} + \cdot\text{Me}$$  \hspace{1cm} (4)$$

The Co–C BDE determined using the BP86 functional is 24.2 kcal/mol after the ZPE correction (Table 5.1), while B3LYP leads to a much lower value of 16.1 kcal/mol as
expected. Interestingly, the relative bond strength reduction in 1e-Ox models, computed as \((E_{1e-Ox} - E_{neutral})/E_{neutral}\), is essentially the same in both cases, i.e., 35% for BP86 and 33% for B3LYP. Although the absolute values are quite different, the relative changes appear independent of the type of functional used. We note that one electron reduction is more effective than one electron oxidation in weakening the Co–C bond.

The most noticeable structural change that takes place upon oxidative cleavage occurs in the Co–N_{Im} length, which is considerably shorter in Im-[Co^{III}(corrin)]^{2+} because the antibonding character of orbitals involved in Co–N_{Im} is diminished (see eq 3). A similar change in Co–N_{ax} bond length takes place in the neutral cofactor upon heterolytic cleavage of the Co–C bond, as was predicted by the previous DFT analysis.\(^{115}\) In both the neutral and the 1e-Ox species, the decreasing Co–N_{ax} distance with the Co–C stretch agrees with the observation that the presence of an axial base facilitates the homolytic Co–C bond cleavage.

5.3.5 Spin Density Changes along the Co–C Bond Stretch. The homolytic Co–C bond cleavage in the 1e-Ox model of MeCbl is associated with an intramolecular electron transfer from the Co–C bond onto the corrin cation radical. The analysis of the spin density changes provides further insight into the oxidative cleavage mechanism, particularly because these modulations are associated with the ground state properties, and can be accurately described by the DFT computations. The changes of the spin density associated with the corrin ring, cobalt, and the methyl carbon caused by the Co–C bond elongation were extracted from the DFT calculations and plotted together with the potential energy curves as shown in Figures 5.7 and 5.8 (lower panels). It is important to mention that the changes predicted on the basis of the B3LYP and BP86 functionals are
rather different. The B3LYP functional describes the 1e-Ox species as essentially Co(III) and the ligand π radical, with small spin polarization (defined as the differences of α and β spin densities for Co and the corrin as a whole, with the exclusion of the contributions from the SOMO). The modulations in spin distribution between the equatorial ligand, cobalt, and methyl along the Co–C stretch (Figure 5.8) give a straightforward description of the bond cleavage: the β electron of the Co–C bonding pair shifts to corrin via cobalt, which behaves as a conductor. As a result, the α-spin at the methyl increases while the α-spin over the equatorial ligand disappears. This clear description of the Co–C bond is rather lost in the case of the BP86 functional due to a spin polarization. The BP86 calculated α spin density on corrin remains unchanged until the Co–C bond is stretched to \( \sim 2.4 \) Å (Figure 5.7), and then it shifts to the methyl group. It appears that initially the β electron density moves to the cobalt only to remove the spin polarization, and after that point, the picture is similar to that noticed for the B3LYP. In both cases, when the cleavage is completed, the spin density is localized essentially on the carbon, with practically no spin density on the cobalt and corrin.

The spin density changes shown in Figures 5.7 and 5.8 confirm the observation based on the potential energy curves, that the mechanism of oxidative cleavage is much simpler than the reductive cleavage,\textsuperscript{115,146} because in the former only one electronic state is involved, while more electronic states, which involve \( \pi^* \) and \( \sigma^* \) orbitals, play a significant role in the latter.

5.3.6 CASSCF/MC-XQDPT2 Analysis of the 1e-Ox MeCbl. Taking into account that the BP86 functional gives the spin density that is evenly distributed between Co and the corrin ligand, whereas B3LYP shows that the spin density is shifted more toward corrin,
we turned our attention to a higher-level *ab initio* method. We used the modified version of the second-order multiconfigurational quasi-degenerate perturbation theory (MC-QDPT2) of Nakano, abbreviated as a MC-XQDPT2, employing complete active space self-consistent field (CASSCF) reference wave functions.

In CASSCF-based multireference computations, one of the most challenging steps is the selection of an active space, especially for large molecules comprised of transition metals and tetrapyrrolic ligands such as corrin. In the case of MeCbl, the process of selecting active orbitals was advocated by the active spaces that were previously employed in cobalamin’s derivatives, which were found to be the most appropriate for the excited state benchmark calculations and the cleavage of the Co–C bond. The active space includes 3d-type orbitals (d_{xz}, d_{yz}, d_{x^2-y^2}) of cobalt, π-type orbitals of the corrin ring, as well as the σ(d_{z^2}) orbital associated with the axial bonding σ(Co–C). In addition, the 4d orbitals of the cobalt as well as lowest unoccupied π* corrin orbitals were also included in the correlating orbitals in order to account for the double-shell effect. As a result, the final active space of the 1e-Ox MeCbl was comprised of four occupied d orbitals of cobalt, two occupied π-type orbitals of corrin, and six corresponding correlating orbitals (virtual), which subsequently gives 11 active electrons distributed in 12 active orbitals (Figure 5.9). The possibility of several other active spaces has also been tested; however, the CASSCF(11,12) seems to be the appropriate active space for the analysis of electronic structure of the 1e-Ox MeCbl.
To obtain an accurate description of the low-lying excited states along with the ground state, CASSCF itself is not sufficient because it does not account for the dynamical correlation effects. Therefore, in order to include the dynamical correlation, we carried out state-average SA-CASSCF calculations followed by the second order perturbation theory (QDPT2). The SA-CASSCF/MC-XQDPT2 calculations reveal that the ground state wave function of the 1e-Ox MeCbl has multireference character with three major configuration state functions (CSFs), as shown in Table 5.4. Notably, the single electron is located on $d_{yz}$ of Co and the $\pi$ orbital of corrin with 49% and 26% weight contributions,
respectively. The ground state wave function also has minor contributions (10%) from the quartet state, which corresponds to electron transfer from a $\sigma$(Co–C) bond to a $\sigma^*$ (Co–C) bond. This partial depopulation of the $\sigma$(Co–C) orbital can considerably reduce the strength of the Co–C bond and is one of the main reasons for the observed lowering of dissociation energy in DFT calculations. In addition, SA-CASSCF/MC-XQDPT2 calculations demonstrate that the first excited state is nearly degenerate with respect to ground state and has very low excited state energy (0.22 eV). The corresponding wave function also shows the same type of multireference character as that of the doublet ground state but the unpaired electron is mainly localized on $d_{xz}$ and $\pi$ orbital of the corrin (Table 5.4).

**Table 5.4. Composition of CASSCF/MC-XQDPT2 Wave Function of 1e-Ox MeCbl in terms of Major Configurations State Functions (CSFs) Computed for First Three Excited States**

<table>
<thead>
<tr>
<th>state</th>
<th>$E^b_{\text{MC-XQDPT2}}$</th>
<th>CSF$_1$ (weight%)</th>
<th>CSF$_2$ (weight%)</th>
<th>CSF$_3$ (weight%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.000 (0.000)</td>
<td>222122000000 (49)</td>
<td>221222000000 (26)</td>
<td>202122000000 (10)</td>
</tr>
<tr>
<td>1</td>
<td>0.228 (0.225)</td>
<td>122222000000 (65)</td>
<td>222221000000 (10)</td>
<td>102222000020 (9)</td>
</tr>
<tr>
<td>2</td>
<td>1.837 (1.832)</td>
<td>212222000000 (81)</td>
<td>202222000010 (6)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3.807 (3.811)</td>
<td>122122000010 (32)</td>
<td>121222000010 (17)</td>
<td>112122000020 (15)</td>
</tr>
</tbody>
</table>

$^a$ The active space occupations are described by 11 active electrons in 12 active orbitals, where 0 means unoccupied, 1 means singly occupied, and 2 means doubly occupied. The active space orbitals are shown in Figure 5.9. $^b$ The MC-XQDPT2 energies are in eV, and energy values calculated in the water solvent are shown in parentheses.

Furthermore, we verified the effect of water solvent on the low-lying electronic state using CASSCF/MC-XQDPT2 calculations, which are embedded in the polarized continuum model (PCM) as implemented in the firefly/PC-GAMMES suite of programs.
The presence of the solvent environment does not alter the multireference character of the ground state as well as the excited states. As we have discussed above, there are two major configurations (CSF$_1$ and CSF$_2$), the first one which is Co(IV) corrin with 49% weight while the second one is the Co(III) π-cation corrin radical with 26% weight (Table 5.4). However, the contribution of these two configurations in the first excited state is different than in the ground state: the weight percentage of Co(IV) corrin is increased up to 65%, whereas the Co(III) π-cation corrin radical character is reduced to 10%. In addition, it is important to mention that the calculated energies of the low-lying excited states are very similar to the those resulting from the gas phase calculations, but the solvent environment decreases the energy of the first and the second excited states by 0.003 and 0.005 eV, respectively (Table 5.4), while it increases the energy of the third excited state by 0.04 eV.

The local environment of the corrin under D$_{4h}$ pseudo-symmetry was considered to analyze interactions between low-lying electronic states from SA-CASSCF/QDPT2 calculations. Because four Co–N(corrin) bonds in corrin are nonequivalent, there should be a distortion of cobalt 3d orbitals and π-type orbitals of the corrin ligand. The Co(III) ion has a d$^6$ configuration under pseudo-D$_{4h}$ symmetry that contains three highest occupied molecular orbitals (HOMO) of e$_g$ (d$_{zx}$ and d$_{yz}$) and a$_{1g}$ (d$_{z^2}$) symmetry and two unoccupied d orbitals of b$_{1g}$ and b$_{2g}$ symmetry. When an electron is removed from the MeCbl, 3d orbitals of the Co are strongly mixed with an a$_{2u}$/a$_{1u}$ symmetry based π-type HOMO orbital of the corrin. The actual picture turns out to be more complicated upon the removal of an electron because the d$_{yz}$ and d$_{zx}$ orbitals are nearly degenerate, the resulting electronic configuration becomes unstable, and the molecular system experiences a
pseudo-Jahn–Teller (pJT) effect. This can produce more than one electronic configuration and a near-degeneracy between the ground state and low lying excited states, as 3d orbitals of Co metal are degenerate and mix at the same time with π-corrin orbitals. According to SA-CASSCF/MC-XQDPT2 calculations of the 1e-Ox MeCbl, the unpaired electron is predominantly localized on d_{yz} and corrin π in the ground state, while in the first excited state (pseudo-ground state), it is located on d_{xz}/d_{z^2} and corrin π orbital. It has been demonstrated by Bersuker\textsuperscript{265,266} that a low energy gap between vibronically coupled electronic states is the major consequence of the pJT effect. The pJT effect can only be demonstrated with multiconfigurational methods such as CASSCF, as they account for the nondynamical (near-degeneracy) correlation effects,\textsuperscript{267} which can lead to a nonadiabatic coupling between the electronic states. Indeed, the absence of the pJT effect in DFT seems to be logical for the 1e-Ox MeCbl because the conventional Kohn–Sham DFT method does not include nondynamical correlation effects. The presence of the pJT effect leads to a significant change in electronic properties of the open-shell system, i.e., cation radicals.\textsuperscript{268,269} Note that the theory of the pJT effect is well pronounced for highly symmetrical systems (i.e., metalloporphyrin) because it involves the coupling of the symmetrical states.\textsuperscript{270-272}

It can be concluded that the multireference CASSCF/MC-XQDPT2 method provides a better picture of the single electron distribution due to several effects that the conventional DFT method cannot account for. The CASSCF calculations also suggest that the single electron is not located at the Co–C bond in the ground state, which is in agreement with the DFT approach. The CASSCF/MC-XQDPT2 computations, which comprise a wave function based method, seem to resolve the confusion about the location
of the unpaired electron caused by differences observed with different density functionals. These high-level computations indicate that a single electron resides both on the cobalt and on the π-type corrin orbitals, which is qualitatively consistent with the spin polarization obtained from the BP86 functional. As a consequence, the oxidized species of MeCbl species is better described by an admixture of the π-cation corrin radical and the Co(III)/Co(IV) oxidation state.

5.4 Conclusions

The theoretical study performed in this chapter is among the first in which DFT and CASSCF/MC-XQDPT2 calculations have been employed to investigate the electronic and structural properties of the one-electron-oxidized cobalt corrinoids. In particular, the DFT/BP86 and the DFT/B3LYP levels of theory have been used to explore the mechanism of the Co–C oxidative bond cleavage in the MeCbl cofactor. The most appealing result of this study is the significant lowering of the dissociation energy (~35%), which might be due to partial depopulation of the Co–C σ orbital.

The employed DFT functionals provide a different description regarding the spin density distribution in the 1e-Ox species: the BP86 demonstrates that spin density is evenly distributed between Co and corrin while it shifts more toward the corrin ring in the case of the B3LYP functional, which is more in line with the corrin π radical and Co(III) oxidation state. It is important to note that the DFT study only supports the formation and involvement of the Co(IV) oxidation state if the results of the BP86 functional are considered. However, the state-of-the-art higher level CASSCF/MC-XQDPT2 computations also advocate the formation of the Co(IV) oxidation state of the 1e-Ox MeCbl with some contribution from corrin π-radical Co(III) character. In addition, the
CASSCF/MC-XQDPT2 calculations confirm that the unpaired electron is delocalized on different active orbitals (d orbital of Co and corrin π) where a strong mixing of electronic states has been observed upon the inclusion of a nondynamical correlation effect. A near-degeneracy between the ground and the first excited state is the result of a pseudo-Jahn–Teller (pJT) effect, as it is usually pronounced for paramagnetic species. Consequently, the oxidized species of MeCbl can be adequately described as a mixed corrin π-radical Co(III)/Co(IV) complexes.

It is important to mention that the change in Co–C bond in the 1e-Ox model of MeCbl as compared to the neutral complex is opposite the observations made for the 1e-Red MeCbl analogue as it is exemplified for data based on the BP86 functional (Table 5.1). The Co–C bond slightly lengthens upon electron removal while it shortens upon electron addition. These contrary changes can be attributed to the fact that for the 1e-Red species the electron is almost entirely located on the corrin ligand while removal of the electron partially induces the Co(IV) character. These theoretical results are further supported by the experimental study of Fanchiang et al., which suggested that the oxidative cleavage mechanism is the plausible pathway for the methyl transfer reactions between MeCbl and diaquo-Co(III)cobinamide. At present the most challenging problem is the potential involvement of the oxidized forms of cobalamins and their respective oxidative cleavage mechanism in B₁₂-catalyzed enzymatic reactions, which has yet to be confirmed.

Furthermore, the DFT computations suggest that the Pyr-[Co(Por)]⁺Me⁺ species are Co(III) ligand π radicals rather than the Co(IV) complexes, which is in agreement with experimental findings by Fukuzumi et al., for the Pyr-[Co(TPP)]⁺Me⁺. However,
the Co(III)/Co(IV) balance is quite delicate and could be shifted even by removal of the *meso* phenyls in the model structure. The calculated Co–C bonds are shorter in B-[Co(Por)]’-Me⁺ than that of the 1e-Ox MeCbl. This observation together with a comparison of the composition of the SOMOs, and the spin distribution between Co and the equatorial ligand, shows that the corrin complex has a more Co(IV) character than its porphine analogues, which seems to facilitate the bond cleavage. Thus, the corrin ligand shows some preference for the oxidative Co–C bond cleavage, as compared to porphyrin. This implies that one electron oxidation of MeCbl has a considerable weakening effect on the Co–C bond, although this effect is not as strong as in the case of one electron reduction.

The calculated DFT potential energy curves show that only one electronic state is involved in the bond rupture of the 1e-Ox complex of MeCbl, while in the 1e-Red species, a considerable electronic rearrangement occurs during state crossing. The investigation of the Co–C bond dissociation process followed by changes in spin distribution is troublesome, when BP86 is considered, which is due to the predicted strong spin polarization, whereas the hybrid B3LYP functional gives a quite clear picture for the β electron transfer from the Co–C bond to the equatorial corrin ligand with the cobalt atom playing the role of a conductor. Thus, the 1e-Ox MeCbl and B-[Co(Por)]’-Me⁺ are other examples of paramagnetic metal complexes containing a non-innocent ligand that cannot be successfully analyzed in terms of total spin distribution, unless spin polarization is accounted for.
CHAPTER 6

CHARGE SEPARATION PROPENSITY OF THE COENZYME B12–TYROSINE COMPLEX IN ADENOSYLCOBALAMIN-DEPENDENT METHYLMALONYL–COA MUTASE ENZYME

6.1 Introduction

Coenzyme B$_{12}$ (or adenosylcobalamin, AdoCbl)-dependent mutases catalyze chemical reactions with high selectivity, which makes radical-based rearrangement reactions possible to exchange a hydrogen atom with a functional group on adjacent carbon atoms.$^{25,28,31,32,218,273,274}$ In all AdoCbl-dependent enzymes, the homolysis of the cobalt–carbon bond generates the source of radicals (Ado radical and cob(II)alamin) that initiate the enzymatic catalysis.$^{57,233,275,276}$ The most remarkable aspect is the $\sim 10^{12}$ rate enhancement observed inside the enzyme,$^{1819}$ and the fundamental question is how AdoCbl-dependent enzymes accelerate the homolytic cleavage of the Co–C bond and how exactly this primary step is coupled with subsequent hydrogen atom abstraction.$^{277}$ Although several mechanisms have been suggested for the Co–C bond activation, including upward distortion of the corrin ring,$^{278}$ the trans effect of the axial ligand,$^{279}$ and steric interference with the adenosyl group,$^{280}$ the precise mechanistic base responsible for such enormous rate acceleration remains unclear.

---

$^{4}$ The material presented in this Chapter was reprinted (adapted) with permission from Kumar, N.; Liu, S.; Kozlowski, P. M. J. Phys.Chem. Lett. 2012, 3, 1035. Copyright 2012 American Chemical Society.
A well-studied member of the \( \text{B}_{12} \)-dependent family of enzymes, methylmalonyl-CoA mutase (MCM),\(^{279,281}\) is the only enzyme known to date that participates in both mammals and bacteria. Noteworthy is the X-ray crystallography studies of MCM (PDB ID: 4REQ @ 2.2 Å resolution),\(^{281}\) which have revealed that the binding of the substrate induces large conformational modifications and the active site residue that undergoes significant changes is a tyrosine (Y89) residue. The involvement of Y89 as a key residue in the active site of the MCM enzyme has been further probed by a site-directed mutation of Y89 with phenylalanine (F89).\(^{158}\) It was noticed based on the crystal structure (5REQ) that the mutation has a negligible effect on structural properties of coenzyme \( \text{B}_{12} \). However, the estimated effect of Y89F mutation\(^{158}\) on the Co–C bond homolysis was found to be \( \sim 10^4 \) of the total \( 10^{12} \) rate enhancement. Despite the extensive X-ray analysis of MCM enzyme, detailed mechanistic insight as to how the activation of the Co–C bond takes place upon substrate binding remains a subject of debate. Interestingly, the AdoCbl-based enzymes possess a common structural motif of the tyrosine (Y) residue in close proximity to the cofactor based on the X-ray crystal structures, where the phenoxyl oxygen of the Y residue is often directed toward the AdoCbl cofactor.\(^{282}\) On the basis of this observation, the feasibility of proton-coupled electron transfer (PCET) has been proposed as a possible route to AdoCbl-dependent mutases, and energetic implication regarding involvement of the corrin radical anion has been discussed.\(^{161,282,283}\)

Taking into account the experimental\(^{279,281}\) as well as theoretical studies\(^{161,283}\) of the MCM enzyme, an interesting question arises regarding the plausible role of the substrate binding to displace the Y89 active site and how precisely these structural changes induce radical-based enzymatic reactions. On the basis of the computational
modeling of the MCM enzyme, we suggest the substrate as a distinguished proton-acceptor site that can be found near the Y89 active site to trigger the electron-transfer process (PCET, Scheme 6.1).

Scheme 6.1. Proposed PCET Mechanistic Route for MCM Enzyme.
In this chapter, we report for the first time an independent proof of a PCET mechanistic pathway (Scheme 6.1) where the coupling of both proton-transfer (PT) and electron-transfer (ET) steps are taken into consideration and analyzed using a newly advocated density functional reactivity theory (DFRT). This theoretical approach advances our understanding to categorize the charge separation propensity of electron–proton-transfer reactions as radical-based enzymatic reactions are often difficult to investigate.

6.2 Computational Details

6.2.1 Gas Phase DFT Calculations and Prepared Models. The calculations of small model were performed using DFT method by employing B3LYP functional and 6-31G(d) basis set to all atoms except Co to which higher basis set 6-311G(d,P) was applied as implemented in Gaussian09 quantum chemical software. The structure model was prepared by removing all amide side chains (and the nucleotide loop of corrin ring) which were then replaced by hydrogen atoms. In addition, the adenosyl (Ado) upper axial ligand was truncated and substituted with ethyl group whereas lower axial ligand i.e His610 is replaced with the imidazole group. The Y89 active site of employed model was reasonably truncated to small groups to analyze the possibility proton coupled electron transfer. Therefore, the prepared minimal structure model which is the mimic of the active site of MCM enzyme (PDB code: 4REQ@2.2 Å resolution) was fully optimized which is subjected for single point propensity calculations. However, the possibility of charge transfer were further tested by assuming the large model which includes full Ado group as well as all amide side chains of corrin ring. Due to the added complexity, the structure model was optimized by ONIOM(B3LYP/gen:UFF) level where side chains
were treated at low level of theory (UFF) but rest of the atoms treated as same level as in the case of minimal model.

6.2.2 ONIOM-based QM/MM Calculations. An X-ray crystal structure of substrate bound Methylmalonyl CoA mutases (PDB code: 4REQ@2.2 Å resolution) was used to build a computational model. In the present contribution, we have employed chain A (with residues 38-728) which comprised of coenzyme B12, substrate and crystal water molecules near the chain A as a result of 12289 atoms in total. Hydrogen atoms were added in the enzyme model considering the normal protonation states of all titratable residues at experimental (pH 7.5) using results of PROPKA suitable programs. Furthermore, the protonation states of histidine residues were determined based on PROPKA analysis and especially with the visual inspection which turnout: HID72, HID122, HIE244, HID313, HID328, HID359, HID364, HIE423, HIP474, HIE650, HID610, and HIE661. Wherein HID and HIE are used to represent that these are neutral histidine residues with protonated $N_{\delta 1}$ and $N_{\varepsilon 2}$ atoms, respectively, and HIP is a doubly protonated histidine. Initially, the geometry optimization of hydrogen atoms and heavy atoms of prepared model were carried out using the UFF force field implemented in Gaussian 09. The atom types were explicitly defined within the molecule specifications and corresponding Merz-Kollman electrostatic potential atomic charges (in Gaussian 09) were used in mechanical embedding scheme (ONIOM-ME). In this ONIOM (B3LYP/6-31G(d):UFF) calculation, initially coenzyme B12 was defined as the model system in which higher level DFT were applied. However, the residues within 20 Å from Co and hydrogen atoms were then subjected to geometry optimization with the ONIOM-ME method, where the coenzyme was assumed in the QM region (high level). Subsequently,
active site of MCM were (substrate, Y89, R207, H932) included in the QM atoms and were optimized using ONIOM-ME. After these preliminary computations, ONIOM electronic embedding (EE) calculations were performed. Furthermore, to verify the H-bonded interaction of the Y89 active site, we also employed BP86 functional, which is considered to be suitable functional for B12-dependent coenzyme.

6.3 Results and Discussion

Inspection of available X-ray data of substrate-free as well as substrate-bound forms of MCM reveals that the displacement of active site Y89 takes place upon binding of the substrate (MCA802) and the Y89 residue is in close contact with the substrate; thereby, phenoxyl oxygen of Y89 lies at a distance of 2.62 Å from the −COO\(^-\) group of the substrate (Figure 6.1). The analysis of the active site shows that the pK\(_a\) of Y89 (\(\sim 10\)) is higher than that of the carboxylic group (−COO\(^-\), \(\sim 4\)) of the substrate in physiological environment. The displacement of the phenoxyl proton of the tyrosine residue toward the −COO\(^-\) group of the substrate initiates the ET step. The overall reaction, which involves coupling of PT and ET steps, can be summarize as follow:

$$\text{AdoCbl} + \text{YO} + \text{Sub-COO}^- \rightarrow [\text{AdoCbl}]^{a-} + \text{YO}^a + \text{Sub-COOH} \quad (1)$$

In order to investigate a Y89 active site of substrate-bound MCM enzyme (PDB ID: 4REQ @ 2.2 Å resolution), several models ranging from the minimal structural (at DFT) to the actual active site model (at QM/MM) have been used. Note that a detailed description of each computational method is given in the Chapter 3.
Figure 6.1. Active site close-up of the X-ray crystal structure of substrate-bound AdoCbl (PDB-code: 4REQ @ 2.2 Å resolution). To enhance the visualization picture of the active site, the deoxyadenosyl (Ado) moiety has not been shown.
We employ the electrophilic Fukui function \( f(r) \) and dual descriptor \( f^{(2)}(r) \) formalism from DFRT\textsuperscript{83-87} to demonstrate thermodynamic-like support of the PCET mechanism (Scheme 6.1) by analyzing the propensity of the electron to be transferred to a site other than the proton-acceptor site. Taking into account that \( f^+(r) \) or \( f^-(r) \) forms of the Fukui function measure the response of electron density change upon the addition or removal of an electron, they can be used to quantify the electrophilicity or nucleophilicity of molecular systems.\textsuperscript{88-90} Considering these two approximations, the dual descriptor can be defined as

\[
f^{(2)}(r) = f^+(r) - f^-(r) \approx \rho_{\text{LUMO}}(r) - \rho_{\text{HOMO}}(r)
\]

The \( \rho_{\text{LUMO}}(r) \) and \( \rho_{\text{HOMO}}(r) \) are the LUMO and HOMO densities, respectively. Stated differently, the dual descriptor \( f^{(2)}(r) \) provides a framework to determine electrophilic and nucleophilic regions in an acid–base (or donor–acceptor) complex, and once these regions are sufficiently separated, there will be a propensity for charge separation. For biological systems where the PCET mechanism has been generally accepted, it has been shown in the \( f^{(2)}(r) \) plot of the reactant complexes that the positive and negative \( f^{(2)}(r) \) function values indicate electrophilic and nucleophilic regions, respectively.\textsuperscript{91} When the electrophilic molecular region does not include the proton-acceptor site, then the spatial separation between the PT and ET occurs via PCET.\textsuperscript{91} It was initially anticipated that the transition state (or product) between electron-proton transfer reactions might be more relevant, but the test calculations reveal a similar conclusion to that of the reactant complex. It should be noted that the prediction of the PCET mechanism is thermodynamic in nature, not kinetic.
To characterize the charge separation propensity of AdoCbl–Tyr–Sub complex, the first analysis was performed on a minimal structural model, that is, a mimic of the MCM enzyme in which all the side chains of the cofactor were replaced by hydrogen atoms and axial ligands were simplified, as shown in Figure 6.2. The structure was first optimized using DFT, and single point calculations based on the frozen orbital approximation were further carried out to obtain approximate forms of the Fukui functions, \( f^+(\mathbf{r}) \) and \( f^-(\mathbf{r}) \), required for the dual descriptor. Analysis of the dual descriptor plot of the reactant complex (Figure 6.2) reveals that the electron-donating (or nucleophilic) region (green surface) is located on the Y89 residue, whereas the electron-accepting (or electrophilic) region (red surface) is on the \( \pi \)-system of the corrin ring of the AdoCbl cofactor. Because the nucleophilic and electrophilic parts are separated, it indicates that there would be a propensity of charge separation, assuming that the system is aligned in the direction of a PT. This gives strong evidence that the substrate is not involved in charge separation between the donor and acceptor; instead, it is exclusively susceptible for the proton shift. Importantly, a minimal structural model suggests that an electron can be transferred from aryl system of tyrosine to coenzyme B\(_{12}\), that is, leading to charge separation in the PCET route. Furthermore, to confirm the charge separation propensity between the donor and acceptor, using the finite difference approximation, we computed condensed dual descriptor values of the nucleophilic and electrophilic regions. The total condensed dual descriptor value of the Y89 residue is \(-0.508\), which shows that this site is the most nucleophilic, whereas the total dual descriptor value of AdoCbl is \(+0.735\), which indicates that AdoCbl is the most electrophilic site in the model system. It
should be noted that the sign of the condensed dual descriptor is essential to identify nucleophilic and electrophilic regions within a molecule.\textsuperscript{91}

![Dual descriptor contour surface](image)

Figure 6.2. Dual descriptor contour surface of the minimal structural model of MCM enzyme.

As a next step, a more realistic model, which contains the full deoxyadenosyl (Ado) group and all the side chains of the cofactor, was optimized, and its dual descriptor plot is shown in Figure A8 (Appendix). However, the inclusion of the full AdoCbl structure did not affect the charge separation between the cofactor and Y89, regardless of the fact that the Y89 residue showed some different orientation around the cofactor in comparison to the minimal model. Notably, the proton/charge transfer separation is observed in both models (i.e., mimic of the MCM active site), as examined in many
PCET reactions containing redox-active tyrosine, such as photosystem II, prostaglandin H synthase, and galactose oxidase.

Subsequently, we turned our attention to the actual enzyme scaffold to scrutinize whether a similar pathway (PCET) can be realized inside the enzyme. In order to incorporate the local environment of Y89 that is comprised of a H-bond network, as revealed by X-ray crystallographic analysis (Figure 6.1), several key residues in the immediate vicinity of the active site were included. The ONIOM-based QM/MM computations were performed in which the protein matrix of the substrate-bound MCM enzyme was considered as a part of the MM domain while the Y89 active site and coenzyme B₁₂ were treated as a part of the QM domain. The QM(B3LYP)/MM optimized structure of the active site reveals (Figure 6.3a) that the OH_{Y89}–OOC_{sub} bond (1.56 Å) is shorter in comparison to that in the isolated structural models (≈1.60 Å), implying that the enzyme environment tends to strengthen the H-bond interaction between the Y89 and the substrate to facilitate the PT step. Noteworthy is the solvent molecule (H932), which is engaged in H-bond interaction between the –OH end of Y89 and the side chain of the cofactor. This strong H-bond interaction of the H_{H932}–OH_{Y89} bond (≈1.62 Å) is further inclined to weaken the –OH bond of Y89 to support the consensus of swift PT. It is reasonable to point out that the network of hydrogen bonds around the Y89 helps in facilitating the efficient PT to trigger the ET process. In addition, the dual descriptor plot of the active site (Figure 6.3b) shows that electron that originate from nucleophilic Y89 will likely be transferred to the corrin (π*) ring. Nevertheless, the QM(BP86)/MM optimized structure of the MCM enzyme further confirmed the validity of the previously
optimized structure (B3LYP), which demonstrates the same kind of H-bonding environment as the Y89 active site (see Figure A9, Appendix).

Figure 6.3. (a) Active site view of the QM/MM-optimized structure of substrate-bound MCM enzyme (the Ado moiety has not been shown here). (b) Corresponding dual descriptor contour surface.

As was pointed out before, the involvement of Y89 as a key residue in the PCET mechanism is further supported by the site-directed mutation of Y89F (phenylalanine),\(^ {158}\) where the catalytic activity was depressed by one-third of its original value (\(10^{12}\)-fold). This experimental mutation study\(^ {158}\) along with the envisioned PCET mechanism implies that the Y89 active site indeed displays a crucial role in activating enzymatic reactions of the MCM enzyme. PCET-promoted enzymatic catalysis signifies that one electron-reduced form of the AdoCbl cofactor initiates the cleavage of the Co–C bond as an alternative to its neutral analogue, which can help in understanding the origin of the enormous catalytic rate enhancement in B\(_{12}\)-dependent mutases. However, it has been shown, both experimentally as well as computationally, that electrochemical
addition of an electron t/o the AdoCbl cofactor significantly reduced the bond
dissociation energy (BDE) of the Co–C bond.\textsuperscript{115,140,169}

6.4 Conclusion

The presented DFT and QM/MM computations with a dual descriptor from DFRT in this chapter provide convincing evidence that, irrespective of the structural models used, the PCET mechanism could be operational in the MCM enzyme. We characterized a charge separation propensity of the PCET mechanism based on the hydrogen bond reactant complexes in which the electron-transferred site is often distant from that of the proton-acceptor site. In particular, the active site of the MCM enzyme is inspired by the fact that the binding of the substrate displaces the hydroxyl proton of Y89, which induces the ET step to enhance the Co–C bond activation via the PCET mechanism. Thus, these promising results might settle a new paradigm for B\textsubscript{12}-based enzymatic catalysis in which the oxidation of the nearby Y residue can initiate the enzymatic rearrangement reactions.
CHAPTER 7

REDUCTIVE Co-C BOND CLEAVAGE MECHANISM IN B_{12}-DEPENDENT METHYLMALONYL CoA MUTASE: ENERGETIC IMPLICATIONS

7.1 Introduction

Adenosylcobalamin (AdoCbl or Coenzyme B_{12}, Figure 7.1) is the cofactor to a number of enzymes from a range of organisms that catalyze a wide variety of challenging reactions. In all the AdoCbl-dependent enzymatic reactions, the Co-C bond of the AdoCbl undergoes homolytic cleavage upon substrate binding leading to the formation (at least formally) of 5’-deoxy-5’-adenosyl (Ado•) radical and cob(II)alamin (Scheme 7.1). However, the generated Ado radical plays a critical role to catalyze a various difficult reactions by abstracting an H-atom from a substrate and thus generating a substrate radical that undergoes intramolecular 1,2-rearangment with the functional groups (such as alkyl, hydroxy) to adjacent carbon in order to form a product radical. The re-abstraction of hydrogen completes the catalytic cycle (Scheme 7.1).

Despite the extensive studies of AdoCbl-dependent enzymatic catalysis, there are several fundamental questions that remain unanswered. The most important question is ~10^{12}–fold rate enhancement of the Co-C bond cleavage inside the enzyme in comparison to the thermal homolysis of the free cofactor in solution. The Co-C bond
cleavage in isolated coenzyme B\textsubscript{12} is energetically unfavorable with the bond dissociation energy estimated at \(~31.0\) kcal/mol\textsuperscript{21-24}. In contrast, the homolysis of the Co-C bond inside enzyme is rapid but is not a rate-determining step\textsuperscript{57,130-132,134} implying that its strength must be significantly reduced by \(~17.0\) kcal/mol during catalytic turnover\textsuperscript{133}. The enzyme also shifts the equilibrium constant towards the homolysis products giving equilibrium constant close to unity insight the enzyme\textsuperscript{130-132,134-136}. In addition, it has been also found that homolytic cleavage of the Co-C bond is kinetically coupled to the hydrogen atom abstraction from the substrate\textsuperscript{130-132,134-139}.

Although it is generally assumed that AdoCbl cofactor is only required to initiate the radical reaction and resulting cob(II)alamin behaves like spectator (i.e. does not
participate in subsequent steps), Buckel et al.,\textsuperscript{275} challenged this traditional view and proposed that in AdoCbl-dependent mutases such as methylmalonyl CoA mutase (MCM) or glutamate mutase (GLM) (Scheme 7.1 and 7.2) plays a more active role. It was suggested that the cob(II)alamin acts as a conductor by stabilizing both Ado\textsuperscript{*} and the product-related radical intermediates. One of the way which stabilization may occur via cob(II)alamin would be through concerted mechanism involving Co-C cleavage and H-atom abstraction from substrate. Thus, the concept of cob(II)alamin by which being a
spectator can be applied to subclass of AdoCbl-dependent enzymes known as eliminases and exemplified by diol dehydratase\textsuperscript{287-290} or glycerol dehydratase.\textsuperscript{291,292}

To date several mechanistic explanations have been suggested to account for the enhanced cleavage inside the enzyme. Earlier proposals have been primarily of steric nature\textsuperscript{18,19,45,54,153-159} and emphasized importance of factors such as deformation of the corrin ring,\textsuperscript{15,134,153,154} role of the lower axial ligand,\textsuperscript{155,156} bending of upper alkyl group\textsuperscript{54,157-159} and steric strain induced by substrate binding.\textsuperscript{18,45,158} However, the recent crystallographic,\textsuperscript{41,53,293-296} spectroscopic\textsuperscript{156,297-299} as well as computational studies\textsuperscript{300-302} reveal the invalidity of steric proposals to fully explain the observed catalytic effect. Taking into account that mechanism of AdoCbl-dependent enzymatic catalysis (Scheme 7.1) has been difficult to probe experimentally several groups have employed density functional theory (DFT) calculations to acquire insight into the activation of the Co-C bond and the catalytic properties of AdoCbl cofactor. Two AdoCbl-dependent mutases such as, MCM and GLM (Scheme 2) received particular attention in which reliable crystal structures have been reported.\textsuperscript{53,303,304} The DFT calculations for the Co-C bond cleavage and subsequent hydrogen atom transfer have been reported for model systems in gas phase\textsuperscript{305-308} as well as inside protein via QM/MM approach.\textsuperscript{309-311} In particular, Jensen and Ryde performed comprehensive QM/MM calculations focusing primary on the Co-C bond cleavage inside the GLM enzyme. A brief analysis of the subsequent H-atom abstraction step was also reported.\textsuperscript{312} The significant reduction of the Co-C BDE of AdoCbl from \(~34.2\) kcal/mol in vacuum to \(~1.9\) kcal/mol was obtained inside the GLM enzyme which was attributed to several factors.\textsuperscript{312} The entire catalytic effect was made up of four contributing components with most significant coming from the geometrical
distortion of the coenzyme by the protein which was larger in Co(III) state than Co(II) and stabilization of the dissociated state via electrostatic interactions.

Our group modeled the Co-C bond homolysis and subsequent H-atom abstraction using truncated model of the GLM substrate in the gas phase calculations. The concerted pathway was predicted to be energetically more favorable than stepwise by ~7.0 kcal/mol thus confirming Buckel’s et al., conductor hypothesis. Radom and co-workers also performed similar analysis in gas phase emphasizing the importance of the ribose orientation in the substrate activation. They obtained stepwise pathway pointing out that longer value of the Co•••C distance in the TS for H-abstraction was due to use of the full GLM substrate in their calculations instead of a truncated propanol model applied previously. Their calculations also identified sizable stabilization of TS by ~7.2 kcal/mol for one distinct orientation of the ribose part of AdoCbl as a result of favorable C-H•••O3’ hydrogen bond formation with corrin macrocycle. Thus, the conductor hypothesis was also confirmed in this study with similar energetic effect but having different origin. In contrast, the DFT studies performed by Morokuma et al. predicted for the reaction involving truncated MCM substrate that in gas phase concerted route is energetically favor by 7.7 kcal/mol in comparison to stepwise.

In addition, Warshel et al., employed the empirical valence bond (EVB) methodology along with the free energy perturbation approach to investigate the Co-C bond cleavage and subsequent H-atom abstraction in the case of MCM enzyme. They showed that the concerted mechanism in MCM was more favorable and it was concluded that the catalytic effect mainly originate from the electrostatic interaction between the ribose and the protein. Morokuma et al., also investigated the origin of the catalytic
effect in MCM enzyme using the ONIOM-based QM/MM calculations where they pointed out that the cobalt-carbon bond coordinate alone is not good to study the catalytic event. The transition state for the Co-C bond cleavage was located along with conformational change in the Ado moiety predicted for the homolysis step. They also obtained significant reduction of the Co-C BDE confirming previously reported results of Jensen and Ryde\textsuperscript{312} that catalytic effect can be decomposed into several components such as cage effect or the effect of coenzyme geometrical distortion. When the formation of an Ado radical and subsequent H-atom transfer from the substrate was studied it was concluded that the initial Co-C bond cleavage and the subsequent H-atom transfer occur in a stepwise manner in the protein whereas the concerted pathway is more favorable in the gas phase by 7.7 kcal/mol as discussed above.

Very recently, our group and Yoshizawa et al., suggested a different proposal where the catalytic effect in AdoCbl-dependent enzymes was primary of the electronic nature.\textsuperscript{72,73} Analysis of crystallographic data available for AdoCbl-dependent enzymes reveal that tyrosine residue (Y) is present in the vicinity of cofactor at an average distance of \(\sim 7.5\ \text{Å}\) from the cobalt center.\textsuperscript{160,161} Moreover, a site-specific mutation of Y89 by phenylalanine (F) in case of MCM enzyme results in the significant drop of the order of \(\sim 10^3-10^4\) in the enzymatic activity which was specifically attributed to homolysis of the Co-C bond.\textsuperscript{158} It was suggested that Y residue acts as an internal redox center and enzymatic cleavage of the Co-C bond is more in line with one-electron reduced form of AdoCbl cofactor (i.e., \([\text{AdoCbl}]^-\) rather than the neutral analogue.\textsuperscript{161} In this study, the detailed electronic properties of the AdoCbl– tyrosine complex were examined by means of a wide range of quantum chemical methods including DFT, CASSCF/QDPT2, and
QM/MM framework.\textsuperscript{161} The model complexes were studied in the gas phase as well as inside enzymatic environment. Irrespective of the different theoretical method applied and varied degree of structural complexity the diradical state (i.e. \([\text{AdoCbl}]^*-\text{Y}^*\)) has been found to be the lowest electronic state of the AdoCbl-Y\textsuperscript{−} complex providing strong evidence that the electron transfer from Y residue to the cofactor could takes place.\textsuperscript{161} Furthermore, it was also suggested that substrate binding could play a critical role in the displacement of hydroxyl proton of Y residue towards –COO\textsuperscript{−} group in order to facilitate the electron transfer, hence making it a case of proton-coupled electron transfer (PCET).\textsuperscript{161} In addition, we reported an independent proof of a PCET mechanistic pathway (Scheme 6.1, in Chapter 6) where the coupling of both PT and ET steps are taken into consideration and analyzed using a newly advocated density functional reactivity theory (DFRT).\textsuperscript{317} Herein the charge separation propensity of PCET mechanism-based hydrogen bond reactant complexes has been characterized in which electron transferred site is often distant from that of the proton acceptor site.

The purpose of this study is to elaborate the PCET proposal by analyzing the reductive cleavage of the Co-C bond and subsequent hydrogen transfer inside the enzyme. In this chapter, the QM/MM calculations has been applied to understand the alternate mechanism for the activation of Co-C bond cleavage and to provide energetic implications regarding involvement of the \([\text{AdoCbl}]^{-}\) radical anion in the initial step of AdoCbl-dependent methylmalonyl CoA mutase (MCM). The energetic of the one electron reduced AdoCbl cofactor containing all intermediates as well as the transition states along the reaction pathway has been characterized and compared with its neutral analogues. In addition, we carried out first principles Car-Parrinello molecular dynamics
(MD) simulations with QM/MM framework to calculate the free energy surface for the Co-C bond and subsequent hydrogen transfer step.

7. 2 Computational Methods and Model Systems

7.2.1 Initial Set up for the System and Molecular Mechanics Calculations. To prepare a structurally reasonable model for the QM/MM calculations, we performed computational modeling based on the AdoCbl-dependent methylmalonyl CoA mutase (PDB ID: 4REQ).\textsuperscript{281} The 4REQ structure contains a mixture of the substrate and the product, together with a dissociated Ado group, which shows a questionable geometry. However, the additional structural data is available for the MCM enzyme that provides important structural insight. In particular, the substrate-free MCM structure (PDB ID: 3REQ)\textsuperscript{281} as well as structure of the Y89F mutant (PDB ID: 5REQ) was also reported.\textsuperscript{158} In addition, the detailed thermodynamic and kinetic characterization of the Co-C bond homolysis catalyzed by MCM is available in literature.\textsuperscript{18-20,25,26} Furthermore, Y89F mutation and its influence on Co-C bond homolysis were also reported.\textsuperscript{158} Thus, taking into account all the above structural information we selected 4REQ crystal structure as a primary target for the present QM/MM calculations. However, due to the existing ambiguities regarding the ribose part of Ado moiety we made additional assumptions. We assumed that ribose conformation based on the 3REQ structure would be good initial guess for the 4REQ-based structure where the Co-C bond is not cleaved but substrate is present. Consequently, the ribose part of the 3REQ structure was superimposed with 4REQ and such prepared active site was used as initial guess to perform QM/MM set up. The 4REQ structure was further simplified to prepare a reasonable QM/MM model where the residues 38-728 of the chain A, the AdoCbl cofactor, the substrate, and the
corresponding crystal water molecules closer to chain A, were included, while the remaining flexible residues 3-37 of chain A which are far from the active site as well as chain B, were not included in calculations. The prepared model in such a way contained in total 12289 atoms.

The protonation states of titratable residues at pH 7.5 were determined with the Generalized-Born method using PROPKA suite of program. The system was heated and equilibrated at the CHARMM level of theory in three steps: (i) steepest descent optimization of the system, (ii) molecular dynamics (MD) for 15 ps heating from 50 to 300 K, and (iii) equilibration for 400 ps at 300 K with a timestep of 1 fs. During the MD simulations, the coordinates of the QM-region atoms and of heavy atoms beyond 20 Å from the metal center in the substrate-binding site were kept fixed at the X-ray coordinates. The SHAKE algorithm was used to constrain bonds involving hydrogen atoms. Finally, the system was minimized with the adopted basis Newton-Raphson (ABNR) algorithm for 5000 steps. All calculations were carried out with Discovery Studio 2.0.

7.2.2 QM(DFT)/MM calculations. All QM/MM calculations were performed with the ChemShell interface using TURBOMOLE for the density functional theory (DFT) calculations at the BP86 level of theory with SV(P) and TZVP basis sets (for all atoms) were used to carry out geometry optimizations of all the intermediates as well as transition states of the reaction mechanism. The use of BP86 functional was also advocated by the fact that it gives a very good description of the Co-C bond cleavage in neutral as well as one-electron-reduced cofactors. The CHARMM force field running through the DL-POLY program was used for the MM part of the system. A standard
electronic embedding scheme was chosen where the fixed MM atomic charges were included into the one-electron Hamiltonian of the QM part, and the QM/MM electrostatic interactions were evaluated from the DFT electrostatic potential and the MM atomic charges. No cutoffs were applied for non-bonded MM and QM/MM interactions. We defined a region with 1170 atoms to be fully optimized by including all residues that have atoms within a distance of 10 Å around the corrin ring while we kept the remaining atoms fixed.

Several different partitions of QM and MM regions have been considered in QM/MM calculations but two will be only considered in details. The QM region which contains truncated AdoCbl cofactor with respect to side chains, part of substrate along with Tyr89, Arg207, Glu379 and His610 residues would constitute large QM model. On the other hand, a small QM model will be referring to truncated AdoCbl, part of substrate and His610 residue. In isolated instance where effect of specific residue will be tested the indication will be provided that the large (or small) QM model has been altered by removed or added specific residue. It should be pointed out that our small model is consistent with majority of how QM region was defined in previous QM/MM calculations including study of Jensen and Ryde\textsuperscript{312} or Warshel \textit{et al.}\textsuperscript{315} In recent studies Morokuma \textit{et al.}\textsuperscript{316} defined several models for the QM part with different structural complexity.

7.2.3 QM/MM CPMD Simulations. As a next step, the QM/MM MD simulations were carried out using Car-Parrinello Molecular Dynamics 3.1 (CPMD 3.1) and the QM/MM scheme developed by the Rothlisberger group.\textsuperscript{71} The geometry of the QM part that include the corrin ring, the 5- deoxyadenosyl group bound to the cobalt atom,
His244, and part of the substrate intermediate was optimized as a closed-shell singlet using the BP86 functional,\textsuperscript{67,68} a plane wave basis set with a 70-90 Ry kinetic energy cutoff, and Martins–Troullier pseudopotentials\textsuperscript{72} to describe the interaction between the ionic cores and the valence electrons. Core electrons for Co, C, O, and N atoms were described using semicore Martins and Trouiller pseudopotentials, with nonlocal corrections for Co. All details regarding the QM/MM calculations can be found in ref. \textsuperscript{320}. The hydrogen were used as a capping atoms to close the valences of the QM system. The QM box size was $22 \times 22 \times 22 \, \text{Å}^3$ and contained 198 atoms. Car–Parrinello QM/MM simulations\textsuperscript{71} were performed at 300 K with a Nosé-Hoover chain of thermostat\textsuperscript{321} using a coupling constant of 700 cm\textsuperscript{-1}. The electron fictitious mass was set at 400 au. After a suitable equilibration period for a structure extracted from the classical simulations in which the Co–C5′ bond was cleaved, we performed a steered QM/MM simulation to gradually form the Co–C bond over a period of 10 ps.

### 7.3 Results and Discussion

#### 7.3.1 Mechanistic Details of PCET for the Initial step of the Co-C bond Cleavage.

The X-ray crystal structure of MCM enzyme (PDB ID: 4REQ)\textsuperscript{281} has shown that the binding of substrate induces the large conformational changes and the precise active site residue that undergoes significant modification is a Y89 which moves from its original positions and forms a hydrogen-bond contacts with the carboxylic acid (COO\textsuperscript{−}) of the substrate. The hypothesis of the involvement of the Y89 active site for the B\textsubscript{12} catalysis has been further investigated by a site directed mutation studies where Y89 is mutated by phenylalanine (F) and found that the mutated crystal structure (PDB ID: 5REQ) has minor effect on structural properties of coenzyme B\textsubscript{12}.\textsuperscript{158} The estimated effect of the
Y89F mutation was found to be $\sim 10^3 - 10^4$ in the enzymatic activity which was specifically attributed to the homolysis of the Co-C bond.\textsuperscript{158} This means the substrate binding and Y89 active site play a critical role in activating the Co-C bond of the AdoCbl-dependent mutases.

However, in context to our proposed mechanism for AdoCbl-dependent enzymatic reaction, the substrate binding to MCM enzyme facilitates the proton transfer (PT) from the Y(89)OH to COO$^-$ of the substrate that is coupled with the electron transfer (ET) via PCET pathway (Scheme 6.1 in Chapter 6), which can be summarized as follow:

$$\text{AdoCbl} + \text{YOH} + \text{XCOO}^- \rightarrow [\text{AdoCbl}]^{\bullet -} + \text{YO}^* + \text{XCOOH} \quad (7.1)$$

The catalytic process is initiated by the PT, which is considered to be an energetically uphill reaction. Although the catalytic reaction described by Eq. 7.1 is thermoneutral, there is an effective energy barrier for the PCET step. Due to the nonadiabatic nature of the reaction, it is difficult to quantify the energy requirement for [AdoCbl]$^{\bullet -}$ formation. Considering there is an energy barrier associated with the PT and ET steps, it is anticipated that the energetic requirement would be lower when the PT and the ET steps will occur in a concerted fashion, because the PCET pathway avoids the generation of charged intermediates thereby minimizing the energetic demand of the reaction.

More recently, the possibility of PCET pathway was confirmed using dual descriptor and Fukui function from DFRT, which demonstrates the charge propensity of an ET site, is always different from the PT site of the AdoCbl–YH model complexes i.e., the mimic of MCM enzyme. In addition, the previous QM/MM calculations indicate that the enzyme-bound cofactor can be reduced in the presence of the deprotonated tyrosine
The diradical state (i.e. [AdoCbl]$^•-\cdot Y^*$) has been found to be the lowest electronic state of the AdoCbl–$Y^-$ complex at high-level CASSCF calculations, providing strong evidence that the electron transfer from Y residue to the cofactor could take place inside the enzyme, irrespective of the different theoretical methods used. These important findings imply that the redox chemistry of the AdoCbl–$Y^-$ complex is significantly altered by the enzymatic environment, and consequently, the diradical state, i.e., [AdoCbl]$^•-\cdot Y^*$ (instead of AdoCbl–$Y^-$), becomes the lowest electronic state of the complex.

Despite the potential relevance to the B$_{12}$ catalysis, the possibility of the reductive cleavage of Co-C bond has not been studied extensively because of the non-feasibility of outer sphere ET to the AdoCbl cofactor inside the enzyme. It was argued that species like [AdoCbl]$^•$ cannot be involved in B$_{12}$-dependent enzymatic catalysis because biological reducing agents do not have enough reducing power to overcome the very negative redox potentials of B$_{12}$ cofactors (-1.31 to -1.60 vs. saturated calomel electrode (SCE)). Although, the reduction potential of the isolated AdoCbl in the formamide/propanol solvent mixture has been determined to be -1.31 V versus SCE which is significantly negative that of the isolated tyrosine residue ( +0.51 V and +0.76 V vs. SCE), the enzyme-bound AdoCbl can be altered by the various interactions inside the protein. However, the above argument is based on the redox potentials estimated from solution-based electrochemical measurements, which cannot be directly applied to the enzymatic catalysis. In other words, the interactions among the active site residues can significantly alter the redox potential of the tyrosine as well as that of the cofactor. This has been validated more recently, in the case of photosystem II (PSII), where it has
been demonstrated that the low dielectric constant of the protein active site and hydrogen bonding interactions to other peptide residues significantly change the redox potential of the tyrosine and \( P_{680}^+ \) to balance the overall energetic of the PCET reaction. Thus, this particular study implies that the feasibility of PCET driven reductive cleavage of Co-C bond in coenzyme B\(_{12}\) dependent enzymes could be feasible. Moreover, in case of superoxide dismutase, it has been shown that hydrogen bond interaction of the active site with single a residue can significantly modulate the redox properties (\( \sim 1V \)).\(^{152,323}\) These results thus support the hypothesis that protein matrix can apply significant redox tuning through its influence to electron-coupled proton transfer and energy associated with it.

Therefore, the Y89 residue serves as an internal redox center for the activation of Co-C bond and enzymatic cleavage of the Co-C bond in MCM enzyme is more in line with one-electron reduced form of AdoCbl cofactor (i.e., \([\text{AdoCbl}]^*\)) rather than the neutral analogue.\(^{161}\) Therefore, in the present chapter we applied hybrid QM/MM calculations to understand and quantify the Co-C bond cleavage and subsequent H-abstraction from the substrate in one electron-reduced AdoCbl cofactor. In next two sections, we will extensively discuss the Co-C bond cleavage and subsequent H-abstraction in reduced and neutral form of the AdoCbl cofactor in order to compare the TS and intermediates involved in the reaction mechanism.

### 7.3.2 Reductive Co-C5’ Bond Cleavage and Subsequent H-atom abstraction from the Substrate in one Electron Reduced Form.

To investigate the energetic implication regarding the involvement of \([\text{AdoCbl}]^*\) radical anion in the initial step of AdoCbl-dependent enzymatic catalysis that involve Co-C bond cleavage and subsequent H-atom abstraction from the substrate, we carried out QM/MM calculations for the one-electron
reduced (1e-Red) enzyme-bound AdoCbl based on the crystal structure bound MCM enzyme. The reason we utilized methylmalonyl CoA substrate-bound MCM enzyme for the computational modeling is that it has a rich precedence in terms of availability of the crystal structures and is one of the best-studied enzyme among all the known AdoCbl-dependent enzymes. In hybrid QM/MM calculations, the electronic structure of the 1e-Red AdoCbl complex was investigated assuming a QM matrix which includes truncated coenzyme B$_{12}$ with respect to side chains, part of the methylmalonyl CoA substrate (MCA807) along with lower axial His610 residue, while treating the rest of the enzyme at the MM level of theory (See details in computational section). The QM region defined in the present contribution is similar to that of the previous QM/MM calculations.

The QM/MM structures of the one electron reduced AdoCbl cofactor along with the substrate were optimized considering a doublet electronic state, and the stability of the optimized different structures and transition states (TS) were confirmed by the frequency calculations at the same level of theory. It is more convenient to start the discussion with the TS for the reaction which involves the reductive Co-C5’ bond cleavage and hydrogen abstraction from the methylmalonyl CoA substrate. However, before we go into the details of the TS, we would like to highlight the structural aspects of the ribose portion of AdoCbl cofactor because there are found to be two different conformations (C2’-endo and C3’-endo) of the ribose above the corrin ring when the X-ray crystal structure glutamate mutase (GLM) was solved.$^{304}$ Therefore, in the present QM/MM calculations of the MCM enzyme we used two different conformations of the ribose part of the AdoCbl in order to understand the energetic of the reaction in terms of both conformations. The first conformation is referred to as a C2’-endo of the ribose that
has a short Co-C5’ distance in which both the hydroxyl groups of the ribose are projecting away from the corrin ring. In the second conformation i.e., a C3’-endo conformation, the Co-C5’ distance is considered to be a long with 3’-OH group of the ribose ring pointed towards the corrin ring. Thus, we optimized two different conformations of ribose ring namely C2’-endo and C3’-endo because the impeccable

Figure 7.2. The QM/MM optimized active site of MCM enzyme comprising 1e-reduced AdoCbl cofactor with C2’-endo conformation of the ribose ring: a) reactant complex (RC), b) transition state (TS) between Co-C bond cleavage and subsequent H abstraction from the substrate (Co-C5'-H), C) substrate radical (SR).
Figure 7.3. The QM/MM optimized active site of MCM enzyme comprising 1e-reduced AdoCbl cofactor with C3'-endo conformation of the ribose ring: a) reactant complex (RC), b) transition state (TS) between Co-C bond cleavage and subsequent H abstraction from the substrate (Co--C5'--H), c) substrate radical (SR).

importance of the pseudo rotation of the ribose ring of AdoCbl. In both the conformations, the TS of reductive Co-C5’ bond cleavage and hydrogen abstraction was located as shown Figure 7.2b and 7.3b, however energetically C3’-endo conformation
(Figure 7.3b) of the TS was found to be more stable by 4.26 kcal/mol as compared C2’-endo conformation (Figure 7.4). The validity of the optimized TS of both conformations was further confirmed by frequency calculations where the one imaginary frequency was found which corresponds to the hydrogen atom motion associated with C-H bond cleavage of the substrate and subsequent H-atom transfer to C5’. The stabilization of the C3’-endo conformation was attributed to the H-bond interaction between the -OH end of the ribose ring (sugar moiety) and the -HC fragment of the corrin ring (HO--HC= 2.54 Å). However, the similar interaction was not feasible in the case of C2’-endo.

Figure 7.4. Potential energy surface of 1e-reduced cofactor of the MCM enzyme calculated using QM/MM method where corresponding active sites of RC, TS\textsubscript{Co-C5’-H}, and SR are shown in Figure 7.2 and 7.3.
conformation (HO--HC= 4.52 Å) since both the hydroxyl groups of the ribose are pointed away from the corrin ring. However, such an energetic stabilization (ΔE =7.31 kcal/mol) of the TS in C3’-endo conformation relative to C2’-endo conformation has also been noticed by Radom et al., in the gas phase DFT calculations of GLM enzyme. 314 Although our QM/MM calculation are in agreement with previous study,314 the actual number of energy stabilization of C3’-endo conformation is bit higher (ΔE = 4.26 kcal/mol in present and ΔE = 7.31 kcal/mol in previous calculations) than the present QM/MM calculation, which might be due to the lack of the enzymatic environment. As a result, the most noticeable structural aspects between C3’-endo and C2’-endo conformations is the strong communication of the ribose of the AdoCbl with corrin ring which exceptionally stabilize the transition state indicating the key role of the corrin ring for AdoCbl-dependent rearrangement reactions. As a next step, the optimization of the reaction path leading to the substrate radical (SR) as well as to the reactant complex (RC) was further carried out QM/MM calculations from the TS, in both conformation of the ribose ring.

The QM/MM optimized structure of the reactant complex (RC) with C2’-endo conformation is shown in Figure 7.2a where both the hydroxyl (-OH) groups of the ribose are directed away from the corrin ring as expected based on the TS as well as previous studies.314 In particular, the distance between 3’-OH of the ribose and -HC19 of the corrin ring is found to be a 3.82 Å and the Co-C5’ distance is of 2.08 Å. The short Co-C5’ distance in the optimized geometry of the RC implies that the Co-C bond is not activated yet towards the H-abstraction step. On the other hand, the QM/MM optimized C3’-endo conformation of the RC (Figure 7.3a) also show that the 3’-OH group of the ribose part interacts with the corrin ring (-HC19) via a hydrogen bond. More specifically, the
distance between the 3’-OH of the ribose moiety and -HC19 of the corrin ring is 2.25 Å with a long Co-C5’ distance of 2.59 Å compared to C2’-endo conformation. Interestingly, there is 8.32 kcal/mol energy difference (See, Figure 7.4) for the ribose pseudorotation between the two minimum-energy conformations i.e., C2’-endo and C3’-endo of the 1e-Red reactant complex. The C2’-endo conformation (Figure 2b) of the Ado is energetically the lowest for the RC due to the six-coordinated AdoCbl cofactor in which the axial ligands are not sufficiently elongated. In contrast, the Co-C5 and Co-N(His) bonds are relatively elongated that results in the high energy of the C3’-endo conformation even though the -OH moiety of Ado is interacting with the corrin ring (HO-HC19= 2.25 Å) via H-bond. In other words, this was also due to the fact that AdoCbl cofactor in the C2’-endo conformation was better bound (Co-C5’ = 2.08 Å) than in the C3’-endo conformation (Co-C5’ = 2.59 Å). Though C2’-endo conformation was lower in energy in the resting state, but the C3’-endo conformation crosses the C2’-endo conformation en route to the TS thereby becoming the lower energy conformation in the TS. Moreover, in the C2’-endo conformation of the RC, the unpaired electron was localized on the corrin ring of the AdoCbl cofactor which is in agreement with our previous DFT gas phase calculation of 1e-Red AdoCbl cofactor. However, as the Co-C bond is elongated that electron transferred from corrin to C5’ atom of the ribose ring via Co metal. Thus, in C3’-endo conformation of the RC, there is formation of the C5• radical that will abstract H-atom from the substrate and cob(I)alamin form of the cofactor.

Therefore, it is important to conclude that the reaction catalyzed by MCM enzyme must be started from minimum energy structure i.e., C2’-endo conformation followed by pseudo rotation of the ribose ring in order to form C3’-endo conformation that will
abstract the H-atom from the substrate in the TS (See potential energy profile in Figure 7.4. The energy difference in these two conformations of the ribose of the Ado further supported by the experimental study of GLM$^{304}$ which reveal that the C2’-endo conformation is described as ground state form of the cofactor (with Co-C5’ distance relatively short) while the C3’-endo conformation containing long Co-C5’ distance that results in an H-abstraction step from substrate.

As a next step, similar QM/MM calculations were carried out to connect the TS to the product complex that is to form substrate radical for the both conformation of the ribose (Figure 7.2c and 7.3c). The most important finding, regardless of the ribose conformation of the Ado, is the energetic of the substrate radical which was found to be essentially similar (Figure 7.4) i.e., the C3’-endo conformation was found to be slightly stabilized than the C2’-endo one ($\Delta E = -1.69$ kcal/mol). The closer inspection of the QM/MM optimized structures of the substrate radical implies that the HO--HC19 interaction is noticeable in the C3’-endo conformation (2.74 Å) while it is completely ruptured in the C2’-endo (HO--HC19 = 4.84 Å) in which both the hydroxyl groups are pointed away from the corrin ring, thus further providing insight into the stabilization of C3’-endo conformation.

The QM/MM calculations reveal that the product complex with both conformation (C3’-endo or C2’-endo) is the substrate radical containing cob(II)alamin form of the cofactor, which results from reductive cleavage of the Co-C5’ in 1e-Red AdoCbl cofactor. However, the strong communication between the -OH end of the ribose and -HC19 of corrin ring leads to the stabilization of the C3’-endo conformation in the TS. On the reactant side, when the QM/MM calculations smoothly connects the TS with
the reactant complex (RC), it shows the absence of any intermediate that would be indicative of a stepwise mechanism. However, C2’-endo conformation is more stabilized in the RC that of the C3’-endo conformation which is due to the fact that the Co-C5’ (2.05 Å) bond of the AdoCbl cofactor was coordinated better than in the C3’-endo conformation (Co-C5’ = 2.59 Å). Although C2’-endo conformation was lower in energy in the resting state of MCM bound enzyme, but the C3’-endo conformation crosses the C2’-endo conformation when connect to the TS, thereby becoming the lower energy conformation in the TS. Thus, the energy profile calculated with the QM/MM calculations in 1e-Red AdoCbl (Figure 7.4) containing two different conformation of the ribose confirms that the located TS is indeed associated with a concerted mechanism. The overall activation energy barrier require for concerted mechanism that involve the reductive Co- C5’ cleavage followed by the hydrogen abstraction step calculated to be 18.07 kcal/mol.

The concerted pathway for the Co-C bond cleavage and subsequent H-abstraction in MCM observed in our current work is qualitatively similar to the two previous studies carried out for the B_{12}-dependent mutases, namely, MCM by Warshel et al., and GLM by our group. In the Warshel’s EVB study on MCM, the concerted pathway was calculated to be more favorable than the stepwise Co-C cleavage where 10 kcal/mol of driving force for the hydrogen transfer step was introduced for the concerted pathway. In addition, in our previous DFT studies of GLM, the concerted pathway was predicted to be energetically more favorable than stepwise by ~7.0 kcal/mol.

7.3.3 Co-C5’ Bond Cleavage and Subsequent H-atom abstraction from the Substrate in Neutral AdoCbl. To further elaborate the reductive Co-C5 cleavage
mechanism and subsequent hydrogen abstraction step from the methylmalonyl CoA substrate, we further carried out QM/MM calculations of the neutral AdoCbl cofactor. Although, the reaction mechanism and energetic of such reaction catalyzed by MCM has been studied previously, the various conformation of the ribose part that results from the pseudo rotation glycosidic bond has not discussed extensively. In addition, we would also like to make a comparison of the present results with a similar mechanistic pathway investigated during the earlier computational studies in particular, by Jensen, Morokuma and Warshel et al.
Figure 7.5. The QM/MM optimized active site of MCM enzyme comprising neutral AdoCbl cofactor with C2’-endo conformation of the ribose ring: a) reactant complex (RC), b) transition state (TS) of the H abstraction from the substrate (C5’–H), C) substrate radical (SR).

Figure 7.6. The QM/MM optimized active site of MCM enzyme comprising neutral AdoCbl cofactor with C3’-endo conformation of the ribose ring: a) reactant complex (RC), b) transition state (TS) of the H abstraction from the substrate (C5’–H), C) substrate radical (SR).

To begin the extensive analysis of the reaction catalyzed by MCM, we searched the TS for the Co-C5’ bond cleavage in the MCM-bound cofactor and subsequent H-abstraction from the substrate and thus connected it with the reactant complex (RC) and substrate radical (SR). As discussed above in one electron reduced form of the AdoCbl, we
optimized TS in two different conformation of the ribose (Ado) i.e., C3’-endo and C2’-endo using QM/MM calculations which display only one imaginary frequency along the hydrogen abstraction step from the substrate (C5’-H-Sub). It appears that the energy of the C3’-endo conformation (Figure 7.6b) is slightly lower by 0.90 kcal/mol with respect to C2’-endo (Figure 7.5b) conformation which might be arising due to the different structural interaction between the corrin and ribose part. However, when we optimized the reactant complex from the TS, we reach the energy minimum that appears to be a radical intermediate (RI) and energetically both conformation of the ribose are comparable (Figure 7.7). This radical state is connected to the reactant complex via a high-energy state (1.64 kcal/mol), which can be called as a TS_{Co-C} of the Co-C bond cleavage; nevertheless at the same time, it should be characterized as a second order saddle point on the potential energy surface. Most interestingly, the subsequent H-abstraction step from the radical intermediate (RI) was found to be a thermoneutral reaction with an activation barrier of 8.74 kcal/mol. On the other hand, there is substantial energy difference (5.63 kcal/mol) in the optimized minimum energy C3’-endo (Figure 7.6a) and C2’-endo conformation (Figure 7.5a) of the RC as noticed in the case of 1e-Red analogues. The Co-C bond is ruptured in C3’-endo conformation therefore we further optimized Ado-bound state of AdoCbl coenzyme which is referred to as a C3’-exo conformation of the ribose ring (Figure 7.8) because the distance between ribose part and corrin ring is a very long (5.20 Å).

The overall reaction for the Co-C bond cleavage and subsequent hydrogen abstraction appear to take place in a stepwise mechanism, which starts with the homolytic cleavage of the Co-C bond leading to the formation of a cob(II)alamin and Ado• radical
in a first step, and then subsequently abstracts a hydrogen from the substrate in a second step. In the present study, the activation barrier for the stepwise hydrogen transfer via TS

Figure 7.7. Potential energy surface of the neutral AdoCbl cofactor of MCM enzyme calculated using hybrid QM/MM method where corresponding active sites of the RC, TS$_{C5'-H}$, and SR are shown in Figure 7.5, Figure 7.6 and Figure 7.8.
containing the hydrogen abstraction step from the substrate is computed to be 8.89 kcal/mol relative to the RI complex (C3′-endo conformation) as shown in Figure 7.7, and 19.11 kcal/mol relative to the C3′-exo conformation which is lowest energy state of the reactant complex. The stepwise route initiated from C3′-exo conformation with the Co-C bond cleavage followed by the pseudo rotation of the ribose part to form C3′-endo conformation (RI) that results in the formation the metal-centered cob(II)alamin radical and the C5′-centered ribose radical. The formation of radical intermediate (RI) is critical in the reaction, that tend to abstracts an hydrogen atom from the substrate to form a substrate radical (SR) via high energy TS (C5-H). Interestingly, the SR complex in is isoenergetic (with difference of 0.15 kcal/mol) with the RI complex (Figure 7.7) which is one of the most important finding of the QM/MM calculations where the entire protein was considered for the optimization. However, such detailed information was missing in the previous gas phase DFT calculations, indicating the possibility of concerted mechanism. It is important to note that our computed potential energy profile of MCM-bound AdoCbl is consistent with the previous computational study of Morokuma et al.,\textsuperscript{316} where they calculated the similar energetic information for the different intermediates and transition states involved in the reaction catalyzed by MCM. In particular, the activation barrier for stepwise hydrogen transfer was computed to be 8.9 kcal/mol relative to the intermediate (I), and 18.6 kcal/mol relative to the most stable Co\textsuperscript{III} state of the reactant complex.\textsuperscript{316} Although, this particular QM/MM study is consistent with the
energetic implications of the present finding, however they did not utilized the extensive analysis of the ribose part of the Ado moiety. In addition, the result of the earlier QM/MM study of Co-C cleavage and hydrogen abstraction in GLM by Jensen et al. were also consistent with the stepwise pathway in which they concluded the various energetic contribution that is responsible for the catalytic reduction of the Co-C bond cleavage.

7.3.4 Free energy surface of the Reductive Co-C5’ Bond Cleavage and Subsequent hydrogen abstraction step in one Electron Reduced AdoCbl. Taken together the reduced and neutral form of the AdoCbl, it appears that there are two different mechanisms, namely, concerted and stepwise Co-C homolysis and H-atom abstraction pathway, respectively, for the B_{12} catalysis. In addition, the activation barrier for both the reactions is almost comparable. However, the major limitation of the potential energy surface calculated with QM/MM calculations is the absence of configurational sampling (the dynamics of the intermediates involved in the reactions) and finite temperature, which can be included with the free energy dynamical simulations. Therefore, we carried out first-principles (Car-Parrinello) molecular dynamics (MD) simulations within a QM/MM framework for the 1e-Red AdoCbl cofactor to further understand the reductive Co-C5’ cleavage mechanism and subsequent hydrogen abstraction step from the methylmalonyl CoA substrate.

For the free energy simulations, we employed metadynamics (MTD) to investigate both Co–C5’ bond cleavage and subsequent hydrogen abstraction step from the substrate in a single trajectory. This would provide information regarding the overall free energy of the reaction catalyzed by AdoCbl-bound MCM enzyme. For the H-
abstraction step we used an epsilon coordinate (ε), that is defined as ε = \[r(\text{Sub:C} - \text{Sub:H})] - [r(\text{Ado:C5'} - \text{Sub:H})]\), where Sub is the substrate and Ado is the 5'-deoxyadenosyl (Ado) radical. Therefore, the MTD simulations was carried out in two dimensions using ε (see above) and Co–C5' as collective variables and gaussian functions were added to the potential every 50 MD steps (<5 fs), where 0.3 Å and 3.7 kJ mol⁻¹ were selected as the half-width and depth of the gaussians, respectively. The main advantage of using the two mutually dependent coordinates of the reaction catalyzed by MCM is to identify whether the Co–C5' bond cleavage and subsequent hydrogen abstraction is stepwise or concerted in one electron reduced form of the AdoCbl cofactor.

We first equilibrate the initial structure model of MCM enzyme with CPMD simulations at QM/MM framework simulations that incorporates an intact Co–C5' bond.

![Figure 7.9](image)

**Figure 7.9.** 2-D Free energy surface (KJ/mol) for reductive cleavage of the Co-C5 bond and subsequent hydrogen abstraction step (epsilon) computed with the metadynamics simulations.
with an equilibrium distance of 2.05 Å in neutral as well as one electron-reduced form of the AdoCbl. In addition, the Co–C5' cleavage was evaluated from a series of CPMD QM/MM simulations with constrained values of the Co–C5' coordinate. After the equilibration, we carried out MTD simulations of the Co–C5' cleavage and H-atom abstraction reactions from a model of MCM to compute the free energy surface with two independent coordinates (Co–C5' and epsilon) known as collective variable. The free energy surface for the one electron reduced (1e-Red) form of the AdoCbl (Figure 9) was obtained from 20 ps of sampling. The proper selection of these two coordinates, and the sampling times used, were determined based on the previous MTD simulations by Radom et al.,\textsuperscript{320} where they carried out extensive analysis and concluded that 20ps sampling is enough for the MCM reaction. As shown in Figure 7.9, the lowest energy reaction pathway involving the reductive Co-C5' bond cleavage and subsequent hydrogen abstraction step takes place in a concerted fashion. The reaction starts from the lower right hand side corner with Co-C bond 2.05 Å then proceed diagonally towards the upper left hand side corner where both the Co-C and ε coordinates change concurrently in the plotted free energy surface. Thus, the MTD results are consistent with a concerted pathway indicating the free energy barrier of 80.7 ±3.2  kJ/mol (19.2 ±0.76 kcal/mol) which do not involve the creation of a Ado• radical intermediate as it does in stepwise pathway. However, the stepwise pathway involving the reductive Co-C5' bond cleavage and subsequent hydrogen abstraction step were found to be slightly higher in energy than the concerted pathway on MTD free energy surface. It should be noted that concerted nature of the MCM reaction with the free energy barrier 19.2 ± 0.76 kcal/mol is
consistent with the potential energy surface calculated with the QM/MM calculations given in section 3.2.

4 Conclusion

Taking into account the feasibility of PCET mechanism, the possibility of an alternative mechanism for the activation of Co-C bond cleavage as a initial step in the AdoCbl-dependent enzyme methylmalonyl CoA mutase (MCM ) has been investigated in this chapter. The hybrid QM/MM calculations has been employed to elaborate the reductive cleavage of the Co-C bond and hydrogen abstraction step in coenzyme B\textsubscript{12}-mediated reactions. Most importantly, two distinct conformations (C3’-endo or C2’-endo) of the ribose part of the deoxyadenosyl (Ado) which are considered to be important for the B\textsubscript{12} catalysis are exploited in QM/MM calculations. The potential energy profile of the one electron reduced (1e-Red) form of AdoCbl cofactor containing all intermediates as well as the transition states along the reaction pathway has been characterized and compared with its neutral analogues. The energetic implications regarding the involvement of the [AdoCbl]•- radical anion in the initial step of AdoCbl-dependent has been further discussed. In addition, we calculated the free energy surface for the overall reaction with first principles Car-Parrinello molecular dynamics (MD) QM/MM simulations to further explore initial step of the MCM reaction that comprised of the Co-C cleavage and subsequent hydrogen abstraction step.

The QM/MM calculations of 1e-Red form of the AdoCbl cofactor with two distinct orientations (C3’-endo or C2’-endo) of the ribose portion reveal that Co-C bond cleavage and hydrogen abstraction step is associated with a concerted mechanism. The concerted route of the reaction is further supported by the calculated transition state for
the H-abstraction step and moreover there was no intermediate found along the reaction coordinates when we connected the TS state with the reactant complex and product complex (substrate radical). The overall activation energy barrier require for the concerted mechanism that involve the reductive Co-C5’ cleavage followed by the hydrogen abstraction step found to be 18.07 kcal/mol. In particular, the strong communication between the -OH end of the ribose and -HC19 of corrin ring leads to the stabilization of the C3’-endo conformation in the TS. On the reactant side, when the QM/MM calculations smoothly connect the TS with the reactant complex (RC), it shows the absence of any intermediate that would be indicative of a stepwise mechanism. However, C2’-endo conformation is more stabilized in the RC that of the C3’-endo conformation which is due to the fact that the Co-C5’ (2.05 Å) bond of the AdoCbl cofactor was coordinated better than in the C3’-endo conformation (Co-C5’ = 2.59 Å). Although C2’-endo conformation was lower in energy in the resting state of MCM bound enzyme, but the C3’-endo conformation crosses the C2’-endo conformation when connect to the TS, thereby becoming the lower energy conformation in the TS. The concerted pathway for the reductive Co-C bond cleavage and subsequent H-abstraction in MCM observed in our current work is qualitatively similar to the previous studies carried out for the B_{12}-dependent mutases, namely, MCM by Warshel et al.,^{315} and GLM by our group.^{313}

In the neutral form of AdoCbl cofactor, the overall reaction catalyzed by MCM enzyme found to be occurred in a stepwise mechanism, which starts with the homolytic cleavage of the Co-C bond leading to the formation of a cob(II)alamin and Ado• radical in a first step, and then subsequently abstracts a hydrogen from the substrate in a second
step. The activation barrier for the stepwise hydrogen transfer via TS containing the hydrogen abstraction step from the substrate is computed to be 8.89 kcal/mol relative to the RI complex (C3’-endo conformation) and 19.11 kcal/mol relative to the C3’-exo conformation which is lowest energy state of the reactant complex. The stepwise route initiated from C3’-exo conformation with the Co-C bond cleavage followed by the pseudo rotation of the ribose part to form C3’-endo conformation (RI) that results in the formation the metal-centered cob(II)alamin radical and the C5’-centered ribose radical. As a result, the comparative analysis of potential energy profiles of 1e-Red and neutral AdoCbl cofactor indicate that the structure of the 1e-Red cofactor is not different from that of the neutral complexes, however, the added electron changes the overall reaction pathway from stepwise to concerted.

To further explore the concerted nature of the reductive Co-C cleavage and subsequent hydrogen abstraction step in MCM enzyme, we performed the free energy simulations with two different variables (Co-C and epsilon) representing Co-C cleavage and H-atom abstraction using metadynamics. The lowest energy reaction pathway from metadynamics certainly consistent with a concerted pathway with the free energy barrier of 80.7 ±3.2 kJ/mol (19.2 ±0.76 kcal/mol) which do not involve the creation of a Ado• radical intermediate as it does in stepwise pathway. Most importantly, the concerted nature of the MCM reaction with the free energy barrier 19.2 ± 0.76 kcal/mol is consistent with the potential energy surface calculated with the QM/MM calculations.
CHAPTER 8

SUMMARY AND CONCLUSIONS

In this dissertation, the details of chemical reactions catalyzed by B$_{12}$-dependent cofactors have been investigated using a wide variety of computational methodology. The primary objective of this research work is to understand the B$_{12}$-based biochemical reactions at the atomistic level using quantum chemical calculations and hybrid QM/MM approaches to complement the experimental studies. The bigger pictures of the Ph.D. dissertation is that the computational tools can be applied to predict the electronic structures of different intermediates and reaction complex involved in the enzymatic catalysis by modeling their local electronic events of modular proteins. This would subsequently help in identifying the mechanistic details of how enzymatic environment can catalyzes some challenging chemical reactions so efficiently. In particular, two enzymatic systems are specifically studied, 1) the mechanistic details of various methyl transfer reactions in the catalytic cycle of methionine synthase (MetH), 2) the activation of Co-C bond in AdoCbl-dependent enzymes, particularly in methylmalonyl-CoA mutase (MCM).

In the first part, the electronic structures of different intermediates and reaction complexes involved in the catalytic cycle of MetH and the mechanistic details for the methyl transfer reactions are extensively studied using hybrid quantum
mechanics/molecular mechanics (QM/MM), density functional theory (DFT), and complete active space self-consistent field with second order perturbation theory, computational methodologies. The detailed computational study in chapter 2 revealed the unprecedented role of the Co(I) and the coordination of enzyme-bound cob(I)alamin intermediate involved in the first half catalytic cycle of MetH. Because of its high reactivity and supernucleophilic nature, there is no X-ray crystal structure available for the enzyme-bound cob(I)alamin intermediate, therefore the enzyme-bound structure was solved computationally using mixed QM/MM simulations. The observed displacement of the histidine (His759) residue from the Co(I) center is in agreement with the free cob(I)alamin in solution, indicating the Co(I) is not axially coordinated inside the MetH enzyme. This theoretical study further indicated that the transfer of the methyl group to the substrate is associated with the displacement of the (His759) axial base of the protein scaffold. In addition, a gas phase DFT and high-level ab initio CASSCF/QDPT2 calculations imply that the ground state of the cob(I)alamin is multiconfigurational where the diradical Co(II)-corrin radical configuration (formed by electron transfer from the cobalt to the corrin ring) contributes to the electronic structure of the cob(I)alamin intermediate, indicating a non-innocent behavior of the corrin ring. The weight of these two configurations depends upon the distance of the axial base His from the Co center. The main contribution to the ground state wave function at short Co–N_{Im} distances is the diradical configuration, whereas at long distances, it is the closed-shell. Therefore, our results suggest that (I) the standard description of the Co(I) nucleophile is not appropriate for cob(I)alamin intermediate, due to the non-innocent character of the corrin ring, and
(II) the distance between cob(I)alamin and the axial His plays an important role in modulating the nucleophilicity of Co(I).

In view of our results, we further suggested that (chapter 3) based on the docking of 5-methyltetrahydrofolate at the β-site of the enzyme-bound cob(I)alamin domain followed by QM/MM calculations that the methylation reaction in the first half catalytic reaction of MetH involve not only the metal (i.e., the closed-shell Co(I) configuration) but also the corrin ring (i.e., the diradical Co(II)-corrin radical configuration). This computational study reveals the traditionally assumed S_N2 mechanism for the formation of CH_3-cob(III)alamin and the activation energy barrier for S_N2 reaction is found to be comparable with respect to the determined experimental rate constant. However, the possibility of ET-based radical mechanism consistent with the close-lying diradical states has also been suggested, where an electron transfer (ET) from His-on cob(I)alamin to pterin ring of the protonated CH_3-H_4Folate takes place, indicating Co^{II}(d^7)-pterin radical (π*) diradical state, followed by a methyl radical transfer. In other words, in addition to the traditionally assumed S_N2 mechanism, our findings suggest the possibility of an alternative radical mechanism, in which an electron is transferred from the corrin to the CH_3-H_4folate in order to generate Co(II) and [CH_3-H_4Folate]^-. It should be noted that this ET does not require the presence of any strong reducing agent near the CH_3-H_4folate substrate but rather the cofactor induced formation of anion-radical-like species within the cob(I)alamin:[CH_3-H_4Folate] reactant complex.

As a next step, the mechanistic details of the second-half catalytic cycle of MetH are investigated (chapter 4), where the displacement of a methyl group takes place from the MeCbl form of the cofactor to the homocysteine (Hcy) substrate. The electronic
structure of the reactant complex comprised in second half-catalytic cycle has been studied by means of complete active space self-consistent field (CASSCF) calculation followed by second order perturbation theory computational methods using two different structural models. In spite of the structural differences, both the employed models representing reactant complex, i.e., a mimic of cofactor:substrate provide the same conclusion regarding the modulation of two electronic states (with C···S distances), i.e., ionic and diradical states, indicating S_N2 and ET-based radical mechanisms, respectively. At longer C···S distances the diradical state of the complex is the lowest electronic state which implies that ET-based reductive cleavage is more in line at 4.0 Å C···S distance than that of S_N2-type. In addition, the major contribution of the electronic state represented their weight percentage also strengthened our hypothesis, as one state is always dominant over the other for a particular C···S distance. Thus, the results of the first part of the dissertation could also help in understanding some of the critical aspects related to the enzymatic methyl transfer reactions catalyzed by MetH. However, the key question concerned with the methyl transfer reactions is at what distance the Co–C bond is cleaved and at what distances the C···S bond is formed, i.e., how closely the two reacting modules interact with each other to form the reactant complex. Unfortunately, such structural information is not available from existing X-ray crystallographic data.

To gain further insight for the Co-C bond cleavage we investigated the electronic and structural properties of MeCbl and its one-electron-oxidized form to understand the oxidative mechanism of the Co-C bond, which remains unclear from experimental point of view (chapter 5). The results based on the two different density functionals indicate that the oxidized species have a mixed π-cation corrin radical and Co(III)/Co(IV)
character. The mixed character is further supported by high-level ab initio CASSCF/QDPT2 calculations, which reveal the strong mixing of the electronic states due to nondynamical correlation effects. The near degeneracy, which takes place between the ground and first excited state, is consistent with the presence of a pseudo-Jahn–Teller (pJT) effect in the oxidized form of MeCbl. In addition, the DFT-based investigation of the structurally related porphyrin complexes gives a description consistent with corrin-based analogues and reveals that the corrin species have more Co(IV) character. The most noteworthy result of this study, regardless of the type of method used, is the significant lowering of the Co-C bond dissociation energy (~35 %) in the oxidized form.

In the second part of the dissertation, the alternative mechanism for the Co-C bond activation as well as initial step of the B_{12} catalysis that comprised of the Co-C cleavage followed by hydrogen abstraction from the substrate have been investigated by means of Fukui function and dual descriptor from density functional reactivity theory, hybrid quantum mechanics/molecular mechanics calculations (QM/MM) as well as Car-Parinello molecular dynamics (CPMD) in mixed QM/MM framework to calculate the free energy surface for the enzymatic reaction.

The detailed theoretical study of chapter 6 along with the analysis of the tyrosine active site demonstrates that the electron transfer site is often different than the proton acceptor site in Tyr-AdoCbl reactant complexes, thus indicating the possibility of the proton coupled electron transfer (PCET) mechanism that cause the one electron reduction of the AdoCbl cofactor. The DFT and hybrid QM/MM computations with a dual descriptor from DFRT provide convincing evidence that, irrespective of the structural
models used, the PCET mechanism could be operational in the methylmalonyl CoA mutase (MCM). In particular, the active site of the MCM enzyme is inspired by the fact that the binding of the substrate displaces the hydroxyl proton of Y89, which induces the ET step to enhance the Co-C bond activation via the PCET mechanism. PCET-initiated enzymatic catalysis implies that one electron-reduced form of the AdoCbl cofactor initiates the cleavage of the Co–C bond as an alternative to its neutral analogue, which can help in understanding the origin of the enormous catalytic rate enhancement in B$_{12}$-dependent mutases.

As a next step, the energetic implications of the Co-C bond cleavage followed by hydrogen abstraction from the methylmalonyl substrate in a reduced AdoCbl cofactor has been investigated (chapter 7), by applying QM/MM calculation that considered entire protein scaffold. The QM/MM calculations reveal that the pseudorotation of the ribose ring of Ado moiety plays a critical role that results in a formation of the H-bond interaction between the -O3H part of the ribose and -HC(19) of the corrin ring, thus stabilize the transition state of the H abstraction step. The QM/MM optimizations of various intermediates and transition states along with their predicted energy profile indicate that the reaction consisting of Co-C bond cleavage and subsequent H-atom abstraction occurs in a concerted fashion if the reduced form of AdoCbl cofactor is involved in the reaction. Moreover, a pseudorotation of the ribose ring is important to facilitate the Co-C bond cleavage and subsequent hydrogen abstraction step. The calculated activation energy barrier of the MCM reaction in the reduced model was slightly lower in comparison to the reaction involving the neutral cofactor model that observation further support the PCET-based mechanistic pathway for the origin of
catalysis in AdoCbl-dependent mutases. In addition, the concerted pathway was further supported by the free energy metadynamics simulations that predict the \( \sim 19.2 \pm 0.76 \) kcal/mol free energy barrier for the concerted reaction of the Co-C bond cleavage and subsequent hydrogen abstraction step from the MCM substrate.
REFERENCES


(63) Frisch, M. J.; et al. Gaussian 09, Revision B.1; Gaussian, Inc.: Wallingford, CT, 2009.


        2007, 282, 6609.

(110)  Zydowsky, T. M.; Courtney, L. F.; Frasca, V.; Kobayashi, K.; Shimizu, 


(113)  Datta, S.; Koutmos, M.; Pattridge, K. A.; Ludwig, M. L.; Matthews, R. G. 


        7638.

(116)  Alfonso-Prieto, M.; Biarnes, X.; Kumar, M.; Rovira, C.; Kozlowski, P. M. 


        Chem. 1985, 279, 263.

        34, 5973.


(127)  Fanchiang, Y. T. Organomet. 1983, 2, 121.


(136) Bandarian, V.; Reed, G. H. Biochem. 2000, 39, 12069.


(205) Banerjee, R. V.; Johnston, N. L.; Sobeski, J. K.; Datta, P.; Matthews, R. 


(207) Matthews, R. G.; Smith, A. E.; Zhou, Z. S.; Taurog, R. E.; Bandarian, V.; 

176, 6127.

Biochem. 1995, 228, 640.

(210) Vannelli, T.; Messmer, M.; Studer, A.; Vuilleumier, S.; Leisinger, T. 


2009, 11, 1.

(216) Drennan, C. L.; Huang, S.; Drummond, J. T.; Matthews, R. G.; Ludwig, 


(220) Koutmos, M.; Pejchal, R.; Bomer, T. M.; Matthews, R. G.; Smith, J. L.; 

(221) Bandarian, V.; Pattridge, K. A.; Lennon, B. W.; Huddler, D. P.; Matthews, 


(227) BP86 functional predicts TS which has Co-C = 2.082 Å and C---S = 2.852 Å.


(230) Without inclusion of dynamical correlations the ionic state lies energetically much higher than diradical.


(235) Corrinoids are a family of cofactors with structures similar to cobalamin. In corrinoids, the dimethylbenzimidazole (DBI) that is present in cobalamin is modified or replaced by an alternate ligand.


(244) Halpern, J.; Chan, M. S.; Hanson, J.; Roche, T. S.; Topich, J. A. J. Am. Chem. Soc. 1975, 97, 1606.


(247) Methylcobaloxime = methyl bis(dimethylglyoximato)Co(III).


(253) Fanchiang, Y. T. Organomet. 1983, 2, 121.


(257) Granovsky, A. A. In version 7.1.G; PC GAMESS/Firefly.
(259) Strain orbital.
(281) Mancia, F.; Evans, P. R. Structure 1998, 6, 711.

218
(294) Tollinger, M. K., R.; Hilbert, B. H.; Marsh, E. N. G.; Kräutler, B.


APPENDIX
Table A1. Relative Energy and Composition of the CASSCF/MC-XQDPT2 Wave Function of the Ground State Singlet and three-low Lying Excited States for each Co-Nim Distances. The Wave Function is given in Terms of the corresponding Three Major Configurations State Functions (CSFa, CSFb and CSFc) of each State.*

<table>
<thead>
<tr>
<th>Co-Nim</th>
<th>state</th>
<th>(E_{MC-XQDPT2})</th>
<th>CSFa (weight)</th>
<th>CSFb (weight)</th>
<th>CSFc (weight)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>1</td>
<td>0.00</td>
<td>22221100000 (80%)</td>
<td>22222000000 (6%)</td>
<td>22212100000 (3%)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.42</td>
<td>22221200000 (48%)</td>
<td>22222000000 (21%)</td>
<td>22211100000 (12%)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.81</td>
<td>22122100000 (79%)</td>
<td>22212100000 (6%)</td>
<td>22211101000 (4%)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.44</td>
<td>21221200000 (68%)</td>
<td>22112101000 (16%)</td>
<td>22122001000 (4%)</td>
</tr>
<tr>
<td>2.2</td>
<td>1</td>
<td>0.00</td>
<td>22221100000 (76%)</td>
<td>22222000000 (21%)</td>
<td>22221100000 (12%)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.49</td>
<td>22221200000 (48%)</td>
<td>22221100000 (19%)</td>
<td>22222000000 (18%)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.92</td>
<td>22122100000 (81%)</td>
<td>22212100000 (6%)</td>
<td>22211101000 (4%)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.46</td>
<td>21221200000 (75%)</td>
<td>22112101000 (12%)</td>
<td>22122001000 (3%)</td>
</tr>
<tr>
<td>2.3</td>
<td>1</td>
<td>0.00</td>
<td>22221100000 (72%)</td>
<td>22222000000 (11%)</td>
<td>22221100000 (5%)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.72</td>
<td>22221200000 (46%)</td>
<td>22221100000 (19%)</td>
<td>22222000000 (19%)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.04</td>
<td>22122100000 (82%)</td>
<td>22212100000 (4%)</td>
<td>22211101000 (3%)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.50</td>
<td>21221200000 (69%)</td>
<td>22112101000 (16%)</td>
<td>22122001000 (4%)</td>
</tr>
<tr>
<td>2.4</td>
<td>1</td>
<td>0.00</td>
<td>22221100000 (60%)</td>
<td>22222000000 (19%)</td>
<td>22212100000 (8%)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.74</td>
<td>22221200000 (39%)</td>
<td>22221100000 (31%)</td>
<td>22222000000 (16%)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.12</td>
<td>22122100000 (83%)</td>
<td>22212100000 (3%)</td>
<td>22211101000 (3%)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.52</td>
<td>21221200000 (70%)</td>
<td>22112101000 (16%)</td>
<td>22122001000 (4%)</td>
</tr>
<tr>
<td>2.5</td>
<td>1</td>
<td>0.00</td>
<td>22221100000 (36%)</td>
<td>22222000000 (35%)</td>
<td>22212100000 (15%)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.70</td>
<td>22221200000 (56%)</td>
<td>22221100000 (24%)</td>
<td>22222000000 (10%)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.12</td>
<td>22122100000 (84%)</td>
<td>22211101000 (3%)</td>
<td>22212100000 (2%)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.48</td>
<td>21221200000 (84%)</td>
<td>22112101000 (16%)</td>
<td>22122001000 (4%)</td>
</tr>
<tr>
<td>2.6</td>
<td>1</td>
<td>0.00</td>
<td>22222000000 (51%)</td>
<td>22212100000 (21%)</td>
<td>22211100000 (13%)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.63</td>
<td>22222000000 (79%)</td>
<td>22212100000 (9%)</td>
<td>22222000000 (3%)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.02</td>
<td>22122100000 (85%)</td>
<td>22211101000 (3%)</td>
<td>22212100000 (2%)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.40</td>
<td>21221200000 (70%)</td>
<td>22112101000 (16%)</td>
<td>22122001000 (4%)</td>
</tr>
<tr>
<td>2.7</td>
<td>1</td>
<td>0.00</td>
<td>22222000000 (57%)</td>
<td>22212100000 (24%)</td>
<td>22211100000 (4%)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.60</td>
<td>22221100000 (88%)</td>
<td>22212100000 (3%)</td>
<td>22211100000 (3%)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.94</td>
<td>22122100000 (85%)</td>
<td>22111101000 (3%)</td>
<td>22112200000 (3%)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.35</td>
<td>21222100000 (70%)</td>
<td>22112101000 (16%)</td>
<td>22122001000 (4%)</td>
</tr>
<tr>
<td>2.8</td>
<td>1</td>
<td>0.00</td>
<td>22222000000 (59%)</td>
<td>22212100000 (24%)</td>
<td>22202200000 (5%)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.66</td>
<td>22221100000 (91%)</td>
<td>22211200000 (3%)</td>
<td>22212100000 (1%)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.86</td>
<td>22122100000 (86%)</td>
<td>22111101000 (3%)</td>
<td>22112200000 (1%)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.32</td>
<td>21222100000 (71%)</td>
<td>22112101000 (16%)</td>
<td>22122001000 (4%)</td>
</tr>
</tbody>
</table>

*The MC-XQDPT2 energies are in eV. The active space occupations are described by eleven active orbitals, where 0 means unoccupied, 1 means singly occupied and 2 means doubly occupied. Note that the notation of CSFs used here are in decreasing order of weight % of most
dominant terms (CSFa>CSFb>CSFc) of ground as well as excited states and are different from figure 8. The results of Table S1 show that for each Co-NIm distance, the ground state is separated from the lowest excited state within the range of 0.4 to 0.75 eV and all three excited states are within 1.5 eV of the ground state. Moreover, each excited states has significant diradical character showing metal-to-ligand charge transfer.

Table A2. Weight of CASSCF/MS-CASPT2 calculations in $\Psi_{\text{dirad}}^{\|}$, $\Psi_{\text{dirad}}^{\perp}$ Wave Functions for Small Model Computed at Different C-S Distance.$^b$

<table>
<thead>
<tr>
<th>C-S distance</th>
<th>CASSCF State</th>
<th>CASSCF Energy $^a$(a.u.)</th>
<th>Weight of CASSCF state in MS-CASPT2</th>
<th>Weight of diradical state (CSFs)</th>
<th>CASSCF State</th>
<th>CASSCF Energy $^a$(a.u.)</th>
<th>Weight of CASSCF state in MS-CASPT2</th>
<th>Weight of diradical state (CSFs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.6</td>
<td>1</td>
<td>-0.2993</td>
<td>1</td>
<td>0.65</td>
<td>1</td>
<td>-0.3158</td>
<td>1</td>
<td>0.74</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>-0.3221</td>
<td>1</td>
<td>0.71</td>
<td>1</td>
<td>-0.3281</td>
<td>1</td>
<td>0.73</td>
</tr>
<tr>
<td>3.4</td>
<td>1</td>
<td>-0.3322</td>
<td>1</td>
<td>0.76</td>
<td>1</td>
<td>-0.3324</td>
<td>1</td>
<td>0.79</td>
</tr>
<tr>
<td>3.8</td>
<td>1</td>
<td>-0.3328</td>
<td>1</td>
<td>0.79</td>
<td>1</td>
<td>-0.3336</td>
<td>1</td>
<td>0.81</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>-0.3332</td>
<td>1</td>
<td>0.8</td>
<td>1</td>
<td>-0.3336</td>
<td>1</td>
<td>0.81</td>
</tr>
<tr>
<td>4.2</td>
<td>1</td>
<td>-0.3335</td>
<td>1</td>
<td>0.8</td>
<td>1</td>
<td>-0.3337</td>
<td>1</td>
<td>0.81</td>
</tr>
<tr>
<td>4.6</td>
<td>1</td>
<td>-0.3336</td>
<td>1</td>
<td>0.81</td>
<td>1</td>
<td>-0.3334</td>
<td>1</td>
<td>0.83</td>
</tr>
</tbody>
</table>

$^a$Computed energy of each C-S distance is equal to CASSCF Energy+2423 a.u. $^b$Note that first four electronic state has been used for State Average CASSCF/MS-CASPT2 calculations as there is no mixing of CASSCF states at CASPT2 level.
Table A3. Relevant energetical and geometrical parameters of the neutral and 1e-Ox MeCbl optimized at the BP86/6-31G* level of theory.

<table>
<thead>
<tr>
<th>System</th>
<th>Co-C</th>
<th>Co-Nax</th>
<th>Total energy</th>
<th>Total energy + ZPE</th>
<th>Co-C BDE*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Im-[Co$^{III}$ corrin]-Me$^+$</td>
<td>1.968</td>
<td>2.132</td>
<td>-2604.594802</td>
<td>-2604.120053</td>
<td>42.0(37.0)</td>
</tr>
<tr>
<td>Im-[Co$^{II}$ corrin]$^+$</td>
<td>2.094</td>
<td></td>
<td>-2564.697421</td>
<td>-2564.259879</td>
<td></td>
</tr>
<tr>
<td>Im-[Co$^{III}$ corrin]-Me$^{3+}$</td>
<td>2.011</td>
<td>2.063</td>
<td>-2604.259631</td>
<td>-2603.785176</td>
<td>28.0 (24.2)</td>
</tr>
<tr>
<td>Im-[Co$^{I}$ corrin]$^{2+}$</td>
<td>1.877</td>
<td></td>
<td>-2564.384554</td>
<td>-2563.945404</td>
<td></td>
</tr>
<tr>
<td>Me radical</td>
<td></td>
<td></td>
<td>-39.830409</td>
<td>-39.801255</td>
<td></td>
</tr>
</tbody>
</table>

Table A4. Relevant energetical and geometrical parameters of the neutral and 1e-Ox MeCbl optimized at the B3LYP/6-31G* level of theory.

<table>
<thead>
<tr>
<th>System</th>
<th>Co-C</th>
<th>Co-Nax</th>
<th>Total energy</th>
<th>Total energy + ZPE</th>
<th>Co-C BDE*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Im-[Co$^{III}$ corrin]-Me$^+$</td>
<td>1.956</td>
<td>2.179</td>
<td>-2604.372425</td>
<td>-2603.883238</td>
<td>29.5(24.1)</td>
</tr>
<tr>
<td>Im-[Co$^{II}$ corrin]$^+$</td>
<td>2.181</td>
<td></td>
<td>-2564.487893</td>
<td>-2564.037171</td>
<td></td>
</tr>
<tr>
<td>Im-[Co$^{III}$ corrin]-Me$^{3+}$</td>
<td>2.031</td>
<td>2.050</td>
<td>-2604.033787</td>
<td>-2603.546132</td>
<td>19.2(16.1)</td>
</tr>
<tr>
<td>Im-[Co$^{I}$ corrin]$^{2+}$</td>
<td>1.888</td>
<td></td>
<td>-2564.165721</td>
<td>-2563.712798</td>
<td></td>
</tr>
<tr>
<td>Me radical</td>
<td></td>
<td></td>
<td>-39.837517</td>
<td>-39.807691</td>
<td></td>
</tr>
</tbody>
</table>
Figure A1. Evolution of the spin density of the truncated Im$^{2-}$[Cob(II)alamin] model system with the Co-N_Im distance (Å). Left: spin populations; Right: alpha and beta spin density distributions, coloured in green and magenta, respectively.
Figure A2. CASSCF Active Space Orbitals used in the calculations of the free cob(I)alamin model system.

Figure A3. Five major configurations (CSFs) contributing to the CASSCF ground state singlet of the Im⁻[Cob(I)alamin] model system before localization and their weight (%) as a function of the Co-N₃ distance. (Note that the notation of the CSFs described here are different from the rest of the discussion).
Figure A4. The quantum mechanics (QM) region of the MetH-cob(I)alamin:CH$_3$-H$_4$Folate(THF) reaction complex that was used in the ONIOM-based quantum mechanics /molecular mechanics (QM/MM) calculations is shown in the sticks framework and rest of the protein is treated as the part of the MM region.
Figure A5. The ONIOM-EE based QM(BP86/6-31G(d)/MM optimized structure of the methionine synthase (MetH)-bound methylcobalamin (MeCbl) cofactor.

Figure A6. The ONIOM-ME based QM(BP86/6-31G(d)/MM optimized structure of the methionine synthase (MetH)-bound cob(I)alamin cofactor which were generated by removing the methyl group of MeCbl and adjusting the number of electrons consistent with Co$^{+1}$ oxidation state.
Figure A7. The ONIOM-based QM(BP86/6-31G(d)/MM-ME optimized structure of the unprotonated (N5) MetH-cob(I)alamin:CH$_3$-H$_4$Folate reaction complex of: (a) reactant complex (RC$_{N5}$), (b) transition state (TS$_{N5}$) in closed shell singlet calculations.
Figure A8. Larger Model of B12-dependent coenzyme containing full AdoCbl,Y89 and truncated substrate that shows dual descriptor contour surface for the reactant complex.

Figure A9. Active site view of the ONIOM-based QM/MM optimized structure of substrate-bound MCM enzyme (BP86 Functional). For viewpoint purpose Ado group has not been shown.
## LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vitamin B12</td>
<td>CNCbl</td>
</tr>
<tr>
<td>CNCbl</td>
<td>cyanocobalamin</td>
</tr>
<tr>
<td>XCbl</td>
<td>general cobalamin</td>
</tr>
<tr>
<td>RCbl</td>
<td>alkylcobalamin</td>
</tr>
<tr>
<td>AdoCbl</td>
<td>adenosylcobalamin</td>
</tr>
<tr>
<td>MeCbl</td>
<td>methylcobalamin</td>
</tr>
<tr>
<td>GSCbl</td>
<td>glutathionyl</td>
</tr>
<tr>
<td>OHCbl</td>
<td>hydroxycobalmin</td>
</tr>
<tr>
<td>H2O</td>
<td>aquaocobalamin</td>
</tr>
<tr>
<td>Me</td>
<td>methyl group</td>
</tr>
<tr>
<td>Ado</td>
<td>adenosyl group</td>
</tr>
<tr>
<td>Ado•</td>
<td>adenosyl radical</td>
</tr>
<tr>
<td>B12 cofactors</td>
<td>MeCbl and AdoCbl</td>
</tr>
<tr>
<td>Coenzyme B12</td>
<td>AdoCbl</td>
</tr>
<tr>
<td>α-face</td>
<td>lower face</td>
</tr>
<tr>
<td>β-face</td>
<td>upper face</td>
</tr>
<tr>
<td>DBI</td>
<td>dimethylbenzimidazole</td>
</tr>
<tr>
<td>BDE</td>
<td>bond dissociation energy</td>
</tr>
<tr>
<td>MetH</td>
<td>methionine synthase</td>
</tr>
<tr>
<td>Met</td>
<td>methionine</td>
</tr>
<tr>
<td>Hcy</td>
<td>homocysteine</td>
</tr>
<tr>
<td>His</td>
<td>histidine</td>
</tr>
<tr>
<td>CH₃-H₄Folate</td>
<td>methyltetrahydofolate</td>
</tr>
<tr>
<td>H₄Folate</td>
<td>tetrahydrofolate</td>
</tr>
<tr>
<td>AdoMet</td>
<td>adenosylmethionine</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>MCM</td>
<td>methylmalonyl coA mutase</td>
</tr>
<tr>
<td>RNR</td>
<td>ribonucleotide reductase</td>
</tr>
<tr>
<td>GLM</td>
<td>glutamate mutase</td>
</tr>
<tr>
<td>QM/MM</td>
<td>quantum mechanics/molecular mechanics</td>
</tr>
<tr>
<td>DFT</td>
<td>density functional theory</td>
</tr>
<tr>
<td>CASSCF</td>
<td>complete active space self-consistent field</td>
</tr>
<tr>
<td>MC-QDPT2</td>
<td>multiconfigurational quasi-degenerate second order perturbation theory</td>
</tr>
<tr>
<td>QDPT2</td>
<td>quasi-degenerate second order perturbation theory</td>
</tr>
<tr>
<td>MetH</td>
<td>methionine synthase</td>
</tr>
<tr>
<td>CFeSP</td>
<td>corrinoid iron-sulfur protein</td>
</tr>
<tr>
<td>Y (Tyr)</td>
<td>tyrosine residue</td>
</tr>
<tr>
<td>F</td>
<td>phenylalanine</td>
</tr>
<tr>
<td>Y*</td>
<td>tyrosyl radical</td>
</tr>
<tr>
<td>ET</td>
<td>electron transfer</td>
</tr>
<tr>
<td>PT</td>
<td>proton transfer</td>
</tr>
<tr>
<td>PCET</td>
<td>proton-coupled electron transfer</td>
</tr>
<tr>
<td>H-atom</td>
<td>hydrogen atom</td>
</tr>
<tr>
<td>TS</td>
<td>transition state</td>
</tr>
<tr>
<td>SCE</td>
<td>saturated calomel electrode</td>
</tr>
<tr>
<td>Gln</td>
<td>glutamine residue</td>
</tr>
<tr>
<td>SOD</td>
<td>superoxide dismutase</td>
</tr>
<tr>
<td>PS-II</td>
<td>photosystem-II</td>
</tr>
<tr>
<td>PDB</td>
<td>protein data bank</td>
</tr>
<tr>
<td>AdoMet</td>
<td>adenosyl-L-methionine</td>
</tr>
<tr>
<td>SHE</td>
<td>standard hydrogen electrode</td>
</tr>
<tr>
<td>PhOH</td>
<td>phenol</td>
</tr>
<tr>
<td>1el-Red</td>
<td>one-electron reduced</td>
</tr>
<tr>
<td>1el-Ox</td>
<td>one-electron oxidized</td>
</tr>
<tr>
<td>[AdoCbl]**</td>
<td>one electron red (1e-Red) AdoCbl</td>
</tr>
<tr>
<td>UHF</td>
<td>unrestricted hartree Fock</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>( &lt;S^{**2} &gt; )</td>
<td>expectation value of the square of spin</td>
</tr>
<tr>
<td>His</td>
<td>histidine residue</td>
</tr>
<tr>
<td>-COO(^-)</td>
<td>carboxylate anion</td>
</tr>
<tr>
<td>DMF</td>
<td>dimethylformamide</td>
</tr>
<tr>
<td>Im</td>
<td>imidazole</td>
</tr>
<tr>
<td>CPMD</td>
<td>Car-Parrinello Molecular Dynamics</td>
</tr>
<tr>
<td>MTD</td>
<td>Metadynamics</td>
</tr>
<tr>
<td>ZPE</td>
<td>zero point energy</td>
</tr>
<tr>
<td>( N_{ax} )</td>
<td>axial nitrogen</td>
</tr>
<tr>
<td>SET</td>
<td>single electron transfer</td>
</tr>
<tr>
<td>RC</td>
<td>reactant complex</td>
</tr>
<tr>
<td>TS</td>
<td>transition state</td>
</tr>
<tr>
<td>SR</td>
<td>substrate radical</td>
</tr>
<tr>
<td>PC</td>
<td>product complex</td>
</tr>
<tr>
<td>IC</td>
<td>intermediate complex</td>
</tr>
</tbody>
</table>
CURRICULUM VITAE

Neeraj Kumar
Department of Chemistry
n0kuma05@louisville.edu
University of Louisville
1800 South Second Street
Louisville, KY 40208

502-852-2330 (office)
502-507-3612 (mobile)
502-852-8149 (fax)

OBJECTIVE
Computational Chemistry and Biology, Catalysis, Bioenergy and Biofuels

EDUCATION AND TRAINING:

Ph.D. candidate in Physical Chemistry
Department of Chemistry
University of Louisville, Louisville, KY
Ph.D. candidate Aug 2008 – Summer 2013
GPA 3.98/4.00

M. S. in Chemistry
With Honors, Panjab University
Chandigarh, INDIA
GPA 3.50/4.00

B. S. in Chemistry, Physics and Mathematics
B.A.M. Khalsa College, Panjab University
Chandigarh, INDIA
Aug 2002 – May 2005
GPA 3.70/4.00

RESEARCH EXPERIENCE

Graduate Research Assistant (Ph.D.):
Aug. 2008 - Aug. 2013 (anticipated), Graduate Studies with Prof. Pawel M. Kozlowski
University of Louisville, Louisville, KY
Dissertation title: “Computational modeling of chemical reactions catalyzed by cobalamin (B12)-dependent enzymes using a wide variety of computational techniques”

Summer Research:
(June, 2012 - Aug, 2012), Summer Student with Dr. Felice Lightstone, a senior scientist and
group leader of the Bioscience and Biotechnology Division (BBTD) at Lawrence Livermore
National Laboratory (LLNL), Livermore, CA
Summer Project: “Protein-Ligand interactions to inhibit the function of MurB: Non covalent
interactions for drug design”

Master Research:
(July, 2006- May, 2007) Master Research Project with Dr. Neetu Goel
Panjab University, Chandigarh, INDIA
Dissertation title: “DFT analysis of structure and stability of sodium and beryllium metal clusters”

Experiences:

Computational modeling of B\textsubscript{12}-dependent enzymatic reactions and probed their mechanistic details using supercomputing approaches including docking, DFT, CASSCF/MC-XQDPT2, hybrid QM/MM and \textit{ab initio} CPMD (free energy) simulations.

Biomolecular modeling of active species of methionine synthase and methylmalonyl-CoA mutase to understand the catalytic effect of methyl transfer reactions and rearrangement reactions, respectively.

Structural and Spectroscopic properties of B\textsubscript{12}-based cofactors (CNCbl, MeCbl, AdoCbl) using time dependent TD-DFT, CAASCF/MC-XQDPT2, coupled cluster (CCSD) methods and benchmarked with experimental vibrational spectra.

Spectroscopic characteristics of heme derivatives (of catalytic cycle) in myoglobin protein: Direct comparison between the simulated and experimental resonance Raman (RR) spectra.

Worked on MurB enzyme and incorporated antibacterial agents (ligands) in the active site of MurB to inhibit its function with the aid of docking, molecular dynamics (MD) and QM/MM simulations.

Quantum mechanics, classical and \textit{ab initio} molecular dynamics, free energy simulations, reduction potential related calculations, in depth knowledge of multireference MCSCF and CCSD calculations.

Presenting theoretical topics and scientific research at group meetings, chemistry department, regional and national meetings.

Personal skills include collaboration, organization, good verbal and written communication skills.

Computational Chemistry Codes:

Quantum Chemistry Packages: NWChem, Gaussian, Firefly/PC GAMESS, GAMESS, Q-Chem, MOLPRO.

Molecular Dynamics: AutoDock, AMBER, CHARMM, GROMOS and \textit{ab initio} Car-Parinello Molecular Dynamics (CPMD), VASP.

Computer Skills:

Linux Shell Scripting.

Computer Languages: FORTRAN, Limited Knowledge of C and python.

Operating Systems: Unix/Linux, Windows and Mac.

Visualization softwares: VMD, Chimera, GaussView, Chemcraft, Molden, Maestro, and Jmol.

Plotting softwares: Grace/Xmgr, Gnuplot, Origin, Hyperchem, Chemdraw and Matlab.

Systems: Experience in using IBM, Beowulf, and high-performance distributed-memory clusters, Cardinal Research Cluster (CRC) and Livermore Computing (LC) supercomputer.

AWARDS AND ACHIEVEMENTS
• Lawrence Graduate Student Award from Associate Director, Dr. Art Wong for the Summer Student Poster Symposium held in Aug 9th 2012, Lawrence Livermore National Laboratory (LLNL), Livermore, CA.

• Travel award from Graduate Student Council (GSC) for National Meeting, San Diego, CA, March 2012-Poster presentation, National Meeting of the American Chemical Society (ACS).

• University of Louisville Arts and Sciences Graduate Student Union (GSU) Fellowship-2011 for purchasing electronic devices, February 2011.

• University of Louisville Graduate Student Association Travel Award for attending Q-Chem Workshop, held at Pittsburgh Supercomputer Center, Pittsburgh, 19-20th March 2009.

• Punjab State Merit prestigious Scholarship, D.P.I, Panjab, for the distinct performance in B.S., May 2005 to March 2006.

RESEARCH PUBLICATIONS


**In Preparation:**

9. Kumar, N.; Kozlowski, P. M. “Mechanistic Insights for the Methylation of cob(I)alamin by tetrahydrofolate substrate” (ready for submission)

10. Kumar, N.; Bucher, D.; Kozlowski, P. M. “Initial Step of B12-Dependent Methylmalonyl CoA Mutase” (ready for submission)


**PRESENTATIONS/POSTERS**


2. Neeraj Kumar, Lucas Koziol and Felice Lightstone. “Understanding Ligand-Protein interactions in MurB: QM/MM calculations.” Graduate Student Symposium in Lawrence Livermore National Laboratory, Livermore, CA, Aug. 2012-Poster presentation. (Received best poster Award)


5. Neeraj Kumar and Pawel M. Kozlowski. “Methyl transfer reaction mediated by electron transfer in the catalytic cycle of cobalamin dependent methionine synthase.” Central Regional Meeting of the American Chemical Society, Indianapolis, IN, June 2011 – Oral Presentation.

6. Pawel M. Kozlowski, M. Kumar, and Neeraj Kumar, “Computational modeling of methyl transfer reactions catalyzed by methionine synthase enzyme”, 7th Congress of the International Society for Theoretical Chemical Physics, Waseda University, Tokyo, Japan, September, 2011 (Oral).


**PROFESSIONAL MEMBERSHIPS:**
American Chemical Society (ACS): Journal of Physical Chemistry (JPC) and Computer in Chemistry.