

**Computational Analysis on the Formation of Metal-Binding  
Pterin (MPT) from cyclic Pyranopterin Monophosphate (cPMP)  
in Molybdenum Cofactor Biosynthesis**



**By**

**Khifsa Haroon**

**NUST2016172870**

**Supervisor**

**Dr. Uzma Habib**

**MS Computational Science and Engineering  
Research Center for Modeling and Simulation  
National University of Sciences and Technology  
Islamabad, Pakistan**

**October, 2019**

**Computational Analysis on the Formation of Metal-Binding  
Pterin (MPT) from cyclic Pyranopterin Monophosphate (cPMP)  
in Molybdenum Cofactor Biosynthesis**

By

**Khifsa Haroon**  
**NUST2016172870**

**Supervisor**

**Dr. Uzma Habib**

A thesis submitted in partial fulfillment of the requirement for the  
degree of Masters of Science

in

**Computational Science and Engineering**

**Research Center for Modeling and Simulation**  
**National University of Sciences and Technology**

**Islamabad, Pakistan**

**October, 2019**

# National University of Sciences & Technology

## MASTER THESIS WORK

We hereby recommend that the dissertation prepared under our supervision by Khifsa Haroon & Regn No. NUST2016172870, Titled: Computational Analysis on the Formation of Metal-Binding Pterin (MPT) from cyclic Pyranopterin Monophosphate (cPMP) in Molybdenum Cofactor Biosynthesis be accepted in partial fulfillment of the requirements for the award of MS Computational Science and Engineering degree with (\_\_\_\_\_Grade).

### Examination Committee Members

- |   |                  |
|---|------------------|
| 1. Name: <u>Dr. Rehan Zafar Paracha</u> | Signature: _____ |
| 2. Name: <u>Dr. Mehak Rafiq</u>         | Signature: _____ |

Supervisor's name: Dr. Uzma Habib

Signature: \_\_\_\_\_

Date: \_\_\_\_\_

\_\_\_\_\_  
Head of Department

\_\_\_\_\_  
Date

### COUNTERSIGNED

Date: \_\_\_\_\_

\_\_\_\_\_  
Dean/Principal

**THESIS ACCEPTANCE CERTIFICATE**

Certified that final copy of MS/MPhil thesis written by Ms **Khifsa Haroon** Registration No. **00000172870** of **RCMS** has been vetted by undersigned, found complete in all aspects as per NUST Statutes/Regulations, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for award of MS/MPhil degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the said thesis.

Signature with stamp: \_\_\_\_\_

Name of Supervisor: Dr. Uzma Habib

Date: \_\_\_\_\_

Signature of HoD with stamp: \_\_\_\_\_

Date: \_\_\_\_\_

**Countersign by**

Signature (Dean/Principal): \_\_\_\_\_

Date: \_\_\_\_\_

## APPROVAL

It is certified that the contents and form of the thesis entitled “Computational Analysis on the Formation of Metal-Binding Pterin (MPT) from cyclic Pyranopterin Monophosphate (cPMP) in Molybdenum Cofactor Biosynthesis” submitted by **Khifsa Haroon** have been found satisfactory for the requirement of the degree.

Advisor: **Dr. Uzma Habib**

Signature: \_\_\_\_\_

Date: \_\_\_\_\_

Committee Member 1: **Dr. Rehan Zafar Paracha**

Signature: \_\_\_\_\_

Date: \_\_\_\_\_

Committee Member 2: **Dr. Mehak Rafiq**

Signature: \_\_\_\_\_

Date: \_\_\_\_\_

## **DECLARATION**

I hereby declare that my thesis work titled — Computational Analysis on the Formation of Metal-Binding Pterin (MPT) from cyclic Pyranopterin Monophosphate (cPMP) in Molybdenum Cofactor Biosynthesis is carried out by me under the supervision of Dr. Uzma Habib at Research Center for Modeling and Simulation (RCMS) in National University of Sciences and Technology (NUST). I solemnly declare that to the best of my knowledge, this is my original work and it contains no material which has been accepted for the award of other degree in my name, in any other university. Also no material previously published or written by any other person has been included in this thesis except where due reference has been made to the previously published work.

**Khifsa Haroon**

MS Computational Science & Engineering

**Dedicated to...**  
My family

## **ACKNOWLEDGEMENTS**

I am very thankful to the Almighty Allah for giving good strength and health to complete this work. I would also like to say special thanks to AVM Dr. Rizwan Riaz, Principal Research Center for Modeling and Simulation (RCMS), National University of Sciences and Technology (NUST) for giving me opportunity to work on RCMS Super Computer facility and institutional support to complete my work.

I am very thankful to my supervisor Dr. Uzma Habib Assistant Professor, Research Center for Modeling and Simulation (RCMS), National University of Sciences and Technology (NUST). She guided and supported me a lot during the whole research work. I am thankful to my GEC members Dr. Mehak Rafiq and Dr. Rehan Zafar Paracha Assistant Professors at Research Center for Modeling and Simulation (RCMS), National University of Sciences and Technology (NUST) for their persistent guidance. I am also very thankful to Engineer Muhammad Usman and Engineer Hassan for providing me IT support during my computational work and helping me finding solution whenever I stuck during my simulations.

In the end I would like to thank the persons whom I owe the most, my friends, my family without them beside me I am sure of the fact that I wouldn't have been able to complete my master's degree. Thank you all.

Khifsa Haroon



# Table of Contents

List of Tables

List of Plots

List of Figures

|   |    |
|---|----|
| Abstract.....                                 | 1  |
| 1 Introduction to Molybdenum.....             | 2  |
| 1.1 Molybdenum Enzymes.....                   | 3  |
| 1.1.1 Sulfite Oxidase Family.....             | 3  |
| 1.1.2 Xanthine Oxidase Family.....            | 4  |
| 1.2 Molybdenum Cofactor.....                  | 5  |
| 1.3 Molybdenum Cofactor Biosynthesis.....     | 6  |
| 1.4 Molybdenum Cofactor Deficiency(MoCD)..... | 8  |
| 1.5 Computational Analysis.....               | 8  |
| 1.5.1 Classical Mechanical Methods.....       | 9  |
| 1.5.2 Quantum Mechanical Method.....          | 10 |
| 1.5.3 Molecular Docking.....                  | 12 |
| 2 Literature Review.....                      | 13 |
| 2.1 Molybdenum Cofactor.....                  | 13 |
| 2.2 Biosynthesis of Molybdenum Cofactor.....  | 13 |
| 2.2.1 Synthesis of MPT Dithiolate.....        | 14 |
| 2.3 Moco Deficiency.....                      | 15 |

|   |    |
|---|----|
| Problem Statement .....                                   | 17 |
| Objectives.....   | 18 |
| 3 Methodology .....                                       | 19 |
| 3.1 Active Site Search.....                               | 20 |
| 3.1.1 Crystal Structure from PDB (Protein Data Bank)..... | 20 |
| 3.1.2 Binding Pocket Prediction .....                     | 21 |
| 3.1.2.1 DoGsite Scorer.....                               | 21 |
| 3.1.2.2 RaptorX.....                                      | 21 |
| 3.2 Molecular Docking .....                               | 22 |
| 3.2.1 PyRx.....   | 22 |
| 3.3 Active Site Extraction / Model .....                  | 23 |
| 3.3.2 Gauss view .....                                    | 23 |
| 3.4 Quantum Studies/Calculations- Reaction Mechanism..... | 24 |
| 3.4.1 Gaussian-09.....                                    | 25 |
| 3.4.2 Density functional theory (DFT) method .....        | 25 |
| 3.4.3 Geometry Optimization/Energy Optimization.....      | 25 |
| 3.5 Molden .....  | 26 |
| 4 Results.....  | 27 |
| 4.1 Active Sites/Binding Pocket Search .....              | 27 |
| 4.1.1 Results obtained from DoGsite Scorer.....           | 28 |
| 4.1.2 Results obtained from RaptorX.....                  | 33 |
| 4.2 Molecular Docking .....                               | 35 |

|   |    |
|---|----|
| 4.3 Active Site Extraction/Model .....                    | 37 |
| 4.4 Quantum Studies/Calculations- Reaction Mechanism..... | 38 |
| 4.4.1  Educt Complex Geometry (E).....                    | 38 |
| 4.4.2  Ligand Geometry (L) .....                          | 40 |
| 4.4.3  Educt-Ligand Complex Geometry (EL).....            | 40 |
| 4.5 Role of copper.....                                   | 65 |
| 5 Discussion .....  | 68 |
| 6 Conclusion .....  | 76 |
| References .....  | 77 |

## List of Tables

|  |    |
|--|----|
| Tabel 4.1: DoGsite Scorer generated binding pocket residues along with chains and Drug Scores  | 30 |
| Tabel 4.2: Binding pocket residues from RaptorX of MPT synthase for cPMP ligand .....  | 33 |
| Tabel 4.3: Docking results generated by PyRx.....  | 36 |
| Tabel 4.4: Relative energies (kcal/mol) of optimized geometries in solvent phase of Educt-Ligand Complex (EL <sub>01</sub> L) and Educt-Ligand Complex (EL <sub>01</sub> S) .....  | 53 |
| Tabel 4.5: Relative energies (kcal/mol) of optimized geometries in solvent phase of Educt-Ligand Complex (EL <sub>-11</sub> L) and Educt-Ligand Complex (EL <sub>-11</sub> S)..... | 64 |

## List of Figures

|  |    |
|--|----|
| Figure 1.1: Transport of molybdate into cell .....   | 3  |
| Figure 1.2: Sulfite Oxidase Family .....   | 4  |
| Figure 1.3: Xanthine Oxidase Family .....  | 4  |
| Figure 1.4: Molybdenum Cofactor .....  | 5  |
| Figure 1.5: Structure of nitrogenases .....  | 6  |
| Figure 1.6: Pathway of molybdenum cofactor biosynthesis .....                                      | 7  |
| Figure 1.7: Molecular Mechanics Model .....  | 9  |
| Figure 2.1: Pathway for MPT synthesis from cPMP .....  | 15 |
| Figure 2.2: Normal/deficient Pathways for Moco Biosynthesis.....                                   | 16 |
| Figure 3.1: Web-Server-GUI of DoGsite Scorer .....   | 21 |
| Figure 3.2: Toggled residues in PyRx(left), different confirmation generated by PyRx (right) ..... | 23 |
| Figure 3.3: Gauss View built and edited model structure.....                                       | 24 |
| Figure 4.1: Suggested active site for ligand binding, given in uniprot.....                        | 28 |
| Figure 4.2: Results generated from DoGsite Scorer .....  | 29 |
| Figure 4.3: PyMOL view of first binding pocket (P_0).....  | 31 |
| Figure 4.4: PyMOL view of second binding pocket (P_1) .....  | 31 |
| Figure 4.5: PyMOL view of third binding pocket (P_10) .....  | 32 |
| Figure 4.6: PyMOL view of fourth binding pocket (P_11) .....                                       | 32 |
| Figure 4.7: First Binding pocket (1) for cPMP .....  | 34 |
| Figure 4.8: Second Binding pocket (2) for cPMP.....  | 34 |
| Figure 4.9: Protein-Ligand Interactions after docking (PLIP view) .....                            | 37 |
| Figure 4.10: Active site of MPT synthase extracted by Swiss-PDB Viewer.....                        | 38 |
| Figure 4.11: Proposed reaction mechanism .....   | 39 |

|  |    |
|--|----|
| Figure 4.12: Protonated and unprotonated cPMP.....   | 40 |
| Figure 4.13: Stepwise formation of MTP from EL <sub>01</sub> to P <sub>1,2-5</sub> .....   | 41 |
| Figure 4.14: Stepwise formation of MTP from EL <sub>11</sub> to P <sub>1,2-5</sub> .....   | 42 |
| Figure 4.15: Structure of EL <sub>01</sub> geometry with two hydroxyl groups at C <sub>2</sub> .....                               | 43 |
| Figure 4.16: Optimized structure of EL <sub>01</sub> S .....   | 43 |
| Figure 4.17: Optimized structure of EL <sub>01</sub> L .....   | 44 |
| Figure 4.18: Structure of EL <sub>01</sub> -Int-1 geometry with carbon (C <sub>2</sub> ) oxygen (O <sub>2</sub> ) double bond..... | 45 |
| Figure 4.19: Optimized structure of EL <sub>01</sub> -Int-1-S .....  | 45 |
| Figure 4.20: Optimized structure of EL <sub>01</sub> -Int-1-L .....  | 46 |
| Figure 4.21: Structure of EL <sub>01</sub> -Int-2 geometry with S <sub>1</sub> H <sub>5</sub> group at C <sub>1</sub> .....        | 47 |
| Figure 4.22: Optimized structure of EL <sub>01</sub> -Int-2-S .....  | 47 |
| Figure 4.23: Optimized structure of EL <sub>01</sub> -Int-2-L .....  | 48 |
| Figure 4.24: Structure of EL <sub>01</sub> -Int-3 geometry with SH group at C <sub>1</sub> .....                                   | 49 |
| Figure 4.25: Structure of EL <sub>11</sub> -Int-3 geometry with S <sub>1</sub> H <sub>1</sub> group at C <sub>1</sub> .....        | 49 |
| Figure 4.26: Optimized structure of EL <sub>11</sub> -Int-3-S .....  | 50 |
| Figure 4.27: Optimized structure of EL <sub>11</sub> -Int-3-L .....  | 50 |
| Figure 4.28 Structure of EL <sub>01</sub> -Int-4 geometry with SH groups at C <sub>1</sub> and C <sub>2</sub> .....                | 51 |
| Figure 4.29: Optimized structure of EL <sub>01</sub> -Int-4-S .....  | 52 |
| Figure 4.30: Optimized structure of EL <sub>01</sub> -Int-4 -L .....   | 52 |
| Figure 4.31: Relative energies (kcal/mol) graph in solvent phase of Educt-Ligand Complex (EL <sub>01</sub> L).....                 | 53 |
| Figure 4.32: Relative energies (kcal/mol) graph in solvent phase of Educt-Ligand Complex (EL <sub>01</sub> S).....                 | 54 |
| Figure 4.33: Structure of EL <sub>11</sub> geometry with two hydroxyl groups .....   | 55 |

|  |    |
|--|----|
| Figure 4.34: Optimized geometry of EL <sub>-11</sub> -S .....  | 55 |
| Figure 4.35: Optimized geometry of EL <sub>-11</sub> -L .....  | 56 |
| Figure 4.36: Structure of EL <sub>-11</sub> -Int-1 geometry with carbon oxygen double bond at C <sub>2</sub> .....   | 56 |
| Figure 4.37: Optimized geometry of EL <sub>-11</sub> -Int-1-S .....  | 57 |
| Figure 4.38: Optimized geometry of EL <sub>-11</sub> -Int-1-L .....  | 57 |
| Figure 4.39: Structure of EL <sub>01</sub> -Int-2 geometry with SH group at C <sub>1</sub> .....   | 58 |
| Figure 4.40: Optimized geometry of EL <sub>-11</sub> -Int-2-S.....   | 59 |
| Figure 4.41: Optimized geometry of EL <sub>-11</sub> -Int-2-L.....   | 59 |
| Figure 4.42: Structure of EL <sub>-11</sub> -Int-3 geometry with SH group at C <sub>1</sub> .....  | 60 |
| Figure 4.43: Optimized geometry of EL <sub>-11</sub> -Int-3-S.....   | 61 |
| Figure 4.44: Optimized geometry of EL <sub>-11</sub> -Int-3-L .....  | 61 |
| Figure 4.45: Structure of EL <sub>-11</sub> -Int-4 geometry with SH group at C <sub>1</sub> and C <sub>2</sub> .....   | 62 |
| Figure 4.46: Optimized geometry of EL <sub>-11</sub> -Int-4-S.....   | 63 |
| Figure 4.47: Optimized geometry of EL <sub>-11</sub> -Int-4-L .....  | 63 |
| Figure 4.48: Relative energies (kcal/mol) graph in solvent phase of Educt-Ligand Complex (EL-11L).....   | 65 |
| Figure 4.49: Relative energies (kcal/mol) graph in solvent phase of Educt-Ligand Complex (EL-11S).....   | 65 |
| Figure 4.50: Structure of P <sub>1,2</sub> -5 geometry in which Cu attached with S <sub>1</sub> and S <sub>2</sub> (A) and structure without copper (B)..... | 66 |
| Figure 4.51: Optimized structure of P <sub>1,2</sub> -5-S .....  | 67 |
| Figure 4.52: Optimized structure of P <sub>1,2</sub> -5-L .....  | 67 |
| Figure 5.1: First binding pocket (P <sub>0</sub> ), generated by DoGsite scorer .....  | 69 |

|  |    |
|--|----|
| Figure 5.2: First Binding pocket (1) generated from RaptorX.....   | 70 |
| Figure 5.3: Relative energies (kcal/mol) graph in solvent phase of EL <sub>01</sub> L and EL <sub>01</sub> S .....   | 73 |
| Figure 5.4: Relative energies (kcal/mol) graph in solvent phase of EL <sub>-11</sub> L and EL <sub>-11</sub> S ..... | 74 |
| Figure 5.5: Relative energies (kcal/mol) comparison graph of LANL2DZ.....  | 74 |

## ABSTRACT

Formation of molybdopterin (MPT) from cyclic Pyranopterin Mono-Phosphate (cPMP) is the second step in Molybdenum cofactor (MoCo) biosynthesis, which is important for molybdo-enzymes of all organisms. Deficiency of MoCo can cause serious neurological disorders. The focus of this work was to understand the biosynthesis of MPT from cPMP using MPT synthase with the help of computational techniques. Binding pocket for cPMP in MPT synthase was identified using binding pocket prediction software followed by molecular docking. Quantum Mechanical Calculations were performed to understand the reaction mechanism of MPT formation from cPMP using Density Functional Theory (DFT) methods. For Quantum Mechanical Calculations all geometries involved in the reaction mechanism of MPT formation from cPMP were modeled. Geometry optimization and single point energy calculations were performed on all the modeled geometries using B3LYP (Becke, 3-parameter, Lee–Yang–Parr)/ SDD(Stuttgart/Dresden) and B3LYP/LANL2DZ (Los Alamos Na-La 2-Double-zeta) levels of DFT. Keeping the possibility of phosphate group attached with the cPMP, to be protonated or un-protonated, the overall geometry charge and multiplicity of 0,1 and -1,1 were considered, respectively. After analyzing and comparing the results obtained from computational analysis using different levels of DFT on the reaction mechanism of MPT synthase, it has been observed that LANL2DZ gives more feasible results, low energy pathway on the potential energy surface as compared to SDD basis set. In addition, it has also been revealed that cPMP is present in its unprotonated/ anionic form



when attached with the MPT synthase for the formation of MPT. When analyzing the role of copper in the MPT formation, computational results indicated that copper plays no specific role in the formation of MPT from cPMP, however, it can provide the place of attachment for molybdate which is the next step of biosynthetic pathway of Moco.

# CHAPTER 1

## INTRODUCTION

### 1 Introduction to Molybdenum

Molybdenum is included in group of trace elements, is a micro-nutrient essential for living organisms.<sup>1</sup> (Mo) Molybdenum is 4d transition metal in periodic table that is essential for biological activities.<sup>2</sup> Molybdenum alone in biological systems is catalytically inactive so it needs to be complexed by cofactor, commonly called molybdenum cofactor (Moco), to gain its catalytic activity.<sup>3</sup> Biologically active form of molybdenum is molybdate, and in this form it enters the cell by active transport system<sup>4</sup> (Figure 1.1). Molybdenum along with enzymes plays important role in body. It helps in digestion of food, production of energy and in removal of waste from the body.

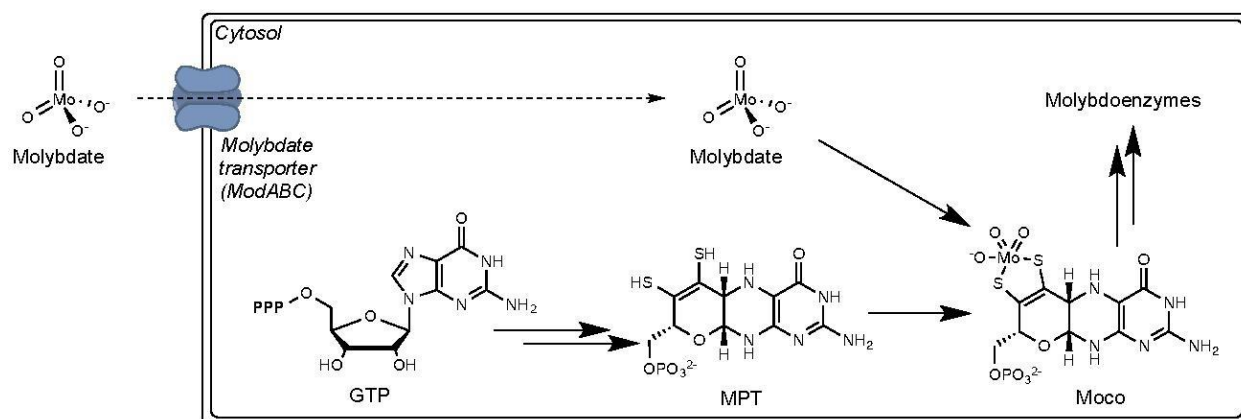


Figure 1.1: Transport of molybdate into cell<sup>4</sup>

## 1.1 Molybdenum Enzymes

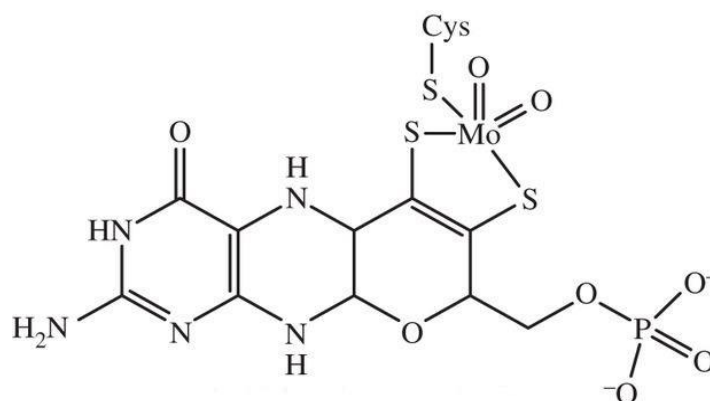
Molybdenum is essential part of several biological enzymes such as sulfite oxidase xanthine oxidase, that plays important processes inside the body. These enzymes are essential constituents for the metabolism of sulfur, carbon and nitrogen, as well as manage chemical reactions. Mo-enzymes catalyze those key reactions that underlay pathways (metabolism of sulfur, carbon and nitrogen) and are characterized by transfer of an oxygen atom.<sup>2, 3</sup>

Molybdo-enzymes, containing the pyranopterin cofactor depends on the coordination chemistry of the Molybdenum (Mo). Moreover, these enzymes share common amino acid sequences and base structure surrounding the pyranopterin cofactor/s in coordination of their central Mo ion.<sup>3,5</sup>

So far more than 40 different Molybdenum enzymes exist in nature, whereas in human body only four (Sulfite Oxidase (SO), Mitochondrial Amidoxime Reducing Component (mARC), XO/xanthine dehydrogenase (XDH), and Aldehyde Oxidase (AO)) are known and are divided into two families.<sup>5,6</sup>

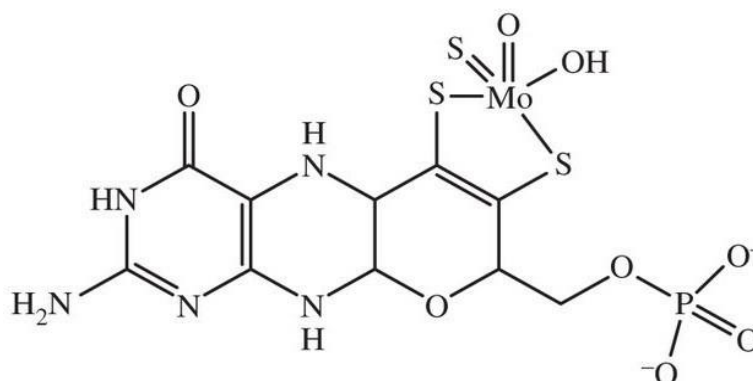
### 1.1.1 Sulfite Oxidase Family

In sulfite oxidase (SO) family, Mo is coordinated with three sulfur and two oxygen atoms (Figure 1.2). Out of three, two sulfur atoms are from Moco, whereas, third equatorial sulfur ligand is from conserved cysteine. The sulfite oxidase family includes members i.e., Sulfite Oxidase and Mitochondrial Amidoxime Reducing Components (mARC). SO catalyzes oxidation of sulfite to sulfate and it is present in inner mitochondrial membrane. Sulfite may be supplied through the diet because it is used as an antimicrobial agent during processing. There are two expressed isoforms of Mitochondrial amidoxime reducing component (mARC), mARC1 and mARC2, and is believed to have functions of detoxification in drug metabolism.<sup>5,6,7</sup>

Figure 1.2: Sulfite Oxidase Family<sup>9</sup>

### 1.1.2 Xanthine Oxidase Family

In xanthine oxidase family, active sites contain a molybdenum atom coordinated with two dithiolene sulfur atoms of the molybdopterin, one sulfur atom of terminal sulfido ligand, an oxygen atom and a hydroxide/water molecule (Figure 1.3). The xanthine oxidase family is formed by two homologous members (xanthine oxidase XO/xanthine dehydrogenase XDH) and Aldehyde Oxidase (AO). XO/XDH is an enzyme involved in purine catabolism and absence or deficiency of XO can lead to kidney stones and renal failure. Whereas AO catalyzes the oxidative hydroxylation of various aldehydes. AO is important for blood circulation, making steroid hormones and also necessary for healthy lungs.<sup>5,7,8</sup>

Figure 1.3: Xanthine Oxidase Family<sup>9,10</sup>

## 1.2 Molybdenum Cofactor

Molybdate, after uptake into the cell, has to be complexed by a unique compound named as tricyclic pteridin or pyranopterin ring to gain catalytic/biological activity.<sup>2</sup> Mo functions as cofactor when coordinated with a pyranopterin ring and form prosthetic group named as molybdenum cofactor (Figure 1.4).<sup>3</sup>

Moco does not occur free in the cell, because it is highly unstable and once liberated from proteins it loses its Mo atom and as a result it undergoes irreversible and rapid loss of function due to oxidation. Therefore, Moco should be transferred soon after its biosynthesis to an enzyme named as apo-target-enzyme, or Moco can also be bound to a carrier protein, this will protect and store it for further use.<sup>8</sup>

In molybdenum cofactor, the function of pterin is to position the catalytic Mo within the active center correctly and also to control its redox behavior.<sup>6</sup> Molybdenum cofactor (MoCo) is important for the molybdo-enzymes of all organisms because it forms the active site of all molybdenum enzymes (Figure 1.4) with exception of the nitrogenases that contains an iron-molybdenum-cofactor<sup>8</sup> (Figure 1.5).

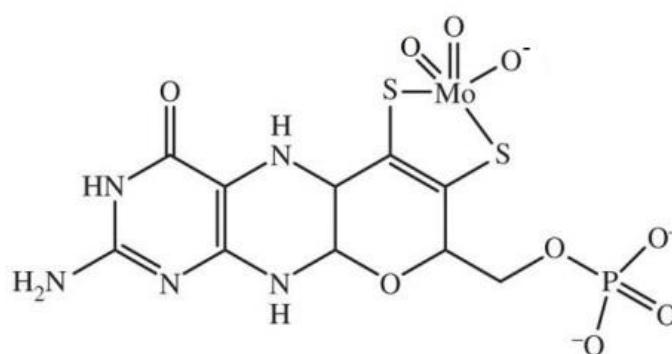


Figure 1.4: Molybdenum Cofactor<sup>9,10</sup>

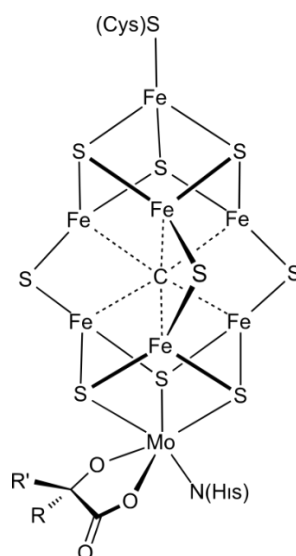


Figure 1.5: Structure of nitrogenases that contains an iron-molybdenum-cofactor<sup>11</sup>

### 1.3 Molybdenum Cofactor Biosynthesis

Molybdenum Cofactor (MoCo) is synthesized by a biosynthetic pathway in all kingdoms of life. It is a process consisted of four steps which works with the help of six proteins (MOCS1(A/B), MOCS2(A/B), MOCS3, GEPH-(G&E) that directly catalyze Moco biosynthesis. This process also requires ATP, iron and copper (Cu).<sup>12</sup>

In first step of synthesis, Guanosine triphosphate is converted to cyclic pyranopterin monophosphate, and two protein MOCS1A and MOCS1B catalyzes this reaction. cPMP, also named as precursor Z.<sup>13</sup>

In the second step of this process two sulfur atom are transferred to cyclic pyranopterin monophosphate, and cPMP by the help of enzyme MPT synthase, is converted into the MPT dithiolate.

In the third step of moco biosynthesis pathway, MPT is adenylated by the protein GEPHYRIN-G, to become receptive for insertion of Molybdenum (Mo). And in final step GEPHYRIN-E inserts Mo in MPT thus forming molybdenum cofactor<sup>12, 13</sup>(Figure 1.6)<sup>14</sup>.

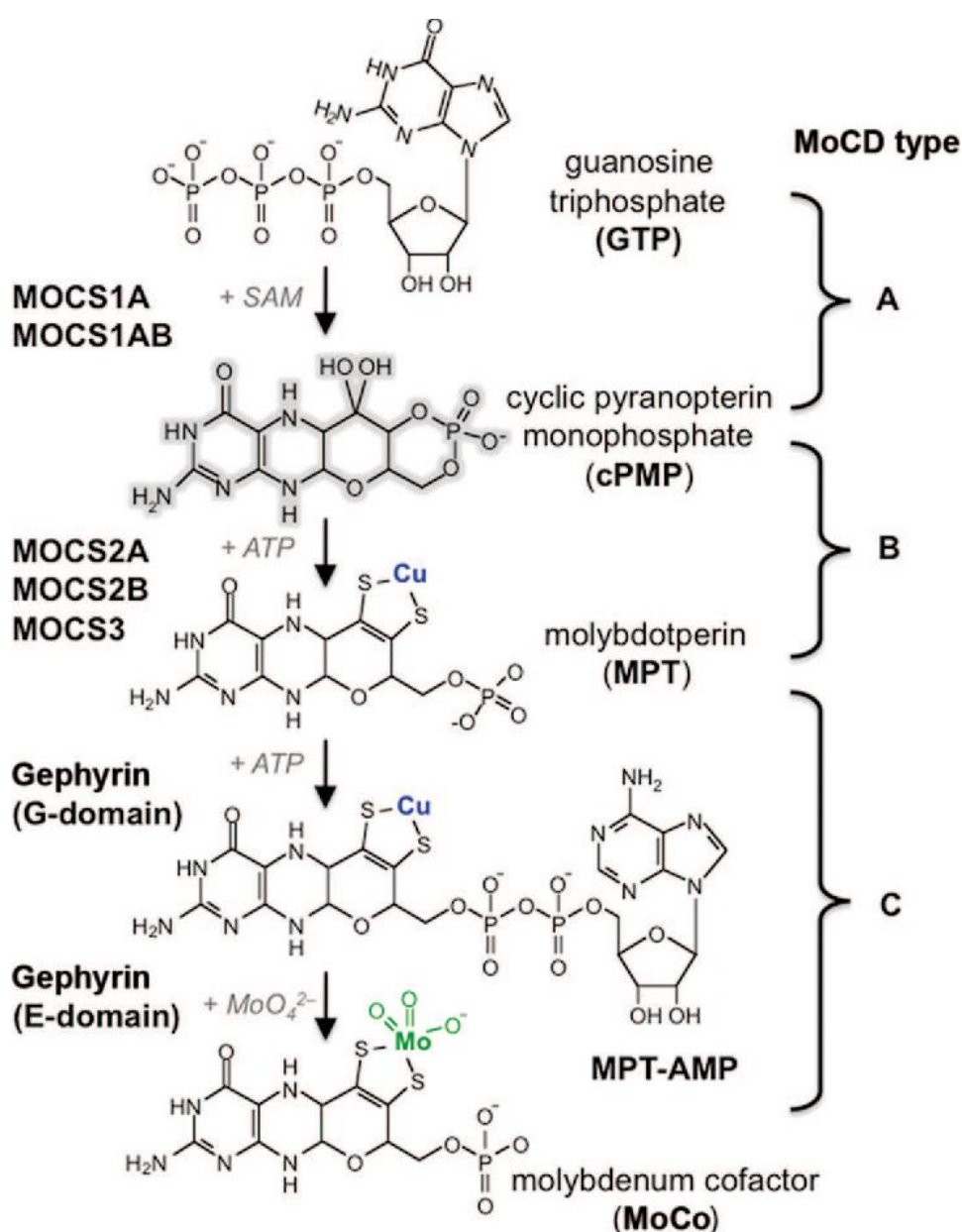


Figure 1.6: Pathway of molybdenum cofactor biosynthesis<sup>14</sup>

In this study, the focus will be on the second step i.e., in formation of MPT from cPMP (precursor Z) by enzyme MPT synthase. In this step, two sulfur atoms are transferred to cPMP by enzyme MPT synthase. It is hetero-tetrameric composed of two subunits encoded by MOCS2A and MOCS2B genes. C-terminal glycine of MOCS2A (smaller subunit) serves as sulfur donor to dithiolene group of MPT. Here another protein MPT-synthase

sulfurase(MOCS3), used for resulfuration of smaller subunit of molybdopterin synthase (MOCS2A) to reactive the enzyme for next cPMP conversion cycle.

#### **1.4 Molybdenum Cofactor Deficiency (MoCD)**

Mutation in any of step of molybdenum cofactor biosynthetic pathway leads to Moco deficiency. MoCo deficiency is rare inborn, an autosomal recessive disorder that is characterized by progressive and severe neurological abnormalities, facial dimorphisms, intractable and untreatable seizure, in most cases dislocated ocular lenses microcephaly, feeding difficulties and death in early childhood.<sup>15,16</sup>

Molybdenum cofactor deficiency (MoCD) can be caused by the deficiency of the molybdoenzymes xanthine dehydrogenase, aldehyde oxidase and sulphite oxidase.<sup>8</sup> MoCD is also caused by defects in the biosynthesis of the Mo-complexed pterin that is the cofactor for enzymes: sulfite oxidase(SO), xanthine dehydrogenase and aldehyde oxidase (AO). And absence of molybdopterin causes in deficient activity of all of these three enzymes.<sup>16,17</sup>

#### **1.5 Computational Analysis**

Computational analysis is performed by using computational chemistry techniques. Computational chemistry is the discipline of science, used to gain chemical information, by using computers. Computational chemistry works at molecules using various models and compares the results of different molecules. In Computational chemistry, geometry of molecules (shape, bond lengths, and angles of molecules), their transition states, chemical reactivity and energies are investigated. Also the interaction of enzymes with substrate, that how a ligand fit into an enzyme's active site is studied using computational chemistry techniques.<sup>19</sup>



Computational methods used for Computational chemistry analysis are divided into two main branches: one is quantum mechanical methods and the other is classical mechanical methods.

### 1.5.1 Classical Mechanical Methods

Classical mechanical methods are further divided into two methods<sup>20</sup>:

#### 1.5.1.1 Molecular Mechanics Method

#### 1.5.1.2 Molecular Dynamic Method

#### 1.5.1.1 Molecular Mechanics(MM) Method

Molecular mechanics are based on classical mechanical methods and are used to compute the geometries for large molecules and their energies in less time period. Molecular mechanics methods are computationally fast with good quality results.<sup>20</sup> MM uses classical physics, and is restricted to the description of only equilibrium structure and conformations. MM is computationally fast, less intensive and can work within limited computer resources compared to Molecular Dynamics method. Structures of large molecules like DNA or protein can be best studied by using Molecular Mechanics.<sup>18,20</sup> Molecular Mechanics used molecular models of molecules. These models of molecules are held together as a collection of balls (atoms) with springs (bonds) (Figure 1.6).

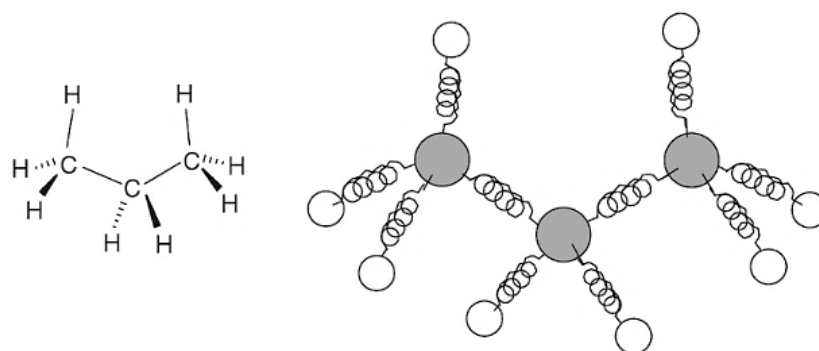


Figure 1.7: Molecular Mechanics Model<sup>20</sup>

### 1.5.1.2 Molecular Dynamic Method

The MD method is used for calculation of the molecular motion of several particles that are interacting in classical manner. In Molecular dynamics, motions of molecules are within some specific time-frame. i.e. inside molecules, atoms have some velocities and are subjected to some forces. The MD simulation is used to study the conformational changes in molecules by integrating Newton 's second law of motion ( $F=ma$ ). Motions of single molecule and also large amounts of molecules can be studied by this method.<sup>18,20</sup>

### 1.5.2 Quantum Mechanical Method

The QM Methods are really based on finding the solutions to Schrödinger wave equation<sup>21</sup>, which describes the complete electronic structure of a molecule. Schrödinger wave equation cannot be solved for any molecule with more than one electron.<sup>20</sup> QM deals with the arrangements of electrons and interactions of those electrons with the electrons and nuclei of other molecules. Quantum methods are not based on experimental information; it can be applied to those areas where there is no prior information and experience. The quantum mechanical methods are further divided into three methods:

1.5.1.1 Ab initio Method

1.5.1.2 Semi-empirical method

1.5.1.3 Density Functional Theory Method

#### 1.5.2.1 Ab Initio Method

Ab initio method of quantum chemistry has become an important tool in the study of atoms and molecules in past three decades. This method is based on Schrödinger wave equation, where the energies of molecules are based on wave function. It is increasingly used in modeling complex systems, especially those arising in biology and materials sciences.<sup>19</sup> Study of large molecules by this method is computationally very expensive, so a small region

of a molecules or system can be modeled and studied by using this method and large molecules can be studied by means of other methods e.g. (molecular mechanics (MM) or continuum solvation models).<sup>19,20</sup>

### 1.5.2.2 Semi-empirical Method

This method is also based on the Schrödinger wave equation, but derived with experimental value or empirical parameters. This method is the mixture of theory and experiment that makes the method semi empirical. Moreover, semi empirical calculations are little slower than molecular mechanics(MM) calculations, but are faster than ab initio.<sup>19,20</sup>

### 1.5.2.3 Density Functional Theory (DFT) Method

DFT methods or calculations are also are based on the Schrödinger equation. However, DFT does not calculate wave function, rather it derives electron density function.

Density functional theory is new approach and its limitations are not so clear yet.<sup>20</sup> DFT calculations are usually slower than Semi-empirical (SE), and advantage of DFT is that it is faster than ab initio but suitable method such as basis set for a particular problem, is difficult in DFT method.<sup>19</sup> DFT investigates the bonding, spectroscopic properties such as absorption and emission spectra, NMR chemical shifts and interactions, and also about structure of many compounds.<sup>20,43</sup>

DFT is capable of performing<sup>43</sup>:

- Geometry optimization (GO)
- Single-point energy (SPE) calculation
- Spectroscopic analysis
- Frequency calculations
- Calculation of dipole moments, atomic charges, electrostatic

- Potentials, polarizabilities etc.

Some of the Most Popular Density Functions<sup>20,43</sup> are:

- **BLYP** is one of the earliest functions and it predicts long bonds.
- **BP86** predicts vibrational frequencies and geometries.
- **OLYP** is a DFT method that performs well for molecules.
- **B3LYP** is the most widely used hybrid functional in quantum studies and showed excellence performance in calculations of energies properties and structures of molecules.
- **PBE0** is in competition with B3LYP and is widely used. It provides excellent accuracy.
- **B2PLYP** is excellent for energies and geometries.

### 1.5.3 Molecular Docking:

It is also the application of computational chemistry, in which a molecule is docked into the active site of enzymes. Usually docking is done with Molecular Mechanics, because of the involvement of large molecules, although selected portions of the biomolecules can also be studied by any of quantum mechanical methods. Docking results are used as a guide for drug designing in better way.<sup>20,21</sup>

# CHAPTER 2

## LITERATURE REVIEW

### 2 Literature Review

#### 2.1 Molybdenum Cofactor

Molybdenum cofactor is present in active site of all molybdoenzymes. The existence of molybdenum cofactor (Moco) was first obtained from genetic studies that characterized mutation in genes; these mutations were affecting the activity of nitrate reductase in *Aspergillus nidulans* (filamentous fungus).<sup>5</sup> Further studies on the expression of nitrate reductase (NR) and xanthine dehydrogenase (XDH) in *Aspergillus nidulans* discovered the pleiotropic mutants that produced to the inactivity of both forms of enzyme.<sup>2,3,5</sup>

Discovery of Moco in humans is associated with sulfite oxidase, one of the Moco containing enzymes. In 1967, deficiency of sulfite oxidase in human (infant) was studied. In year 1978, it was found that sulfite oxidase is a molybdenum containing protein, and molybdenum is essential for normal human development. After these findings, next step was to study the nature of moco (metal containing prosthetic group, common to all molybdoenzymes) and its synthesis.<sup>40</sup>

#### 2.2 Biosynthesis of Molybdenum Cofactor

In the year 1982, Rajagopalan and his coworkers determined the chemical nature of Molybdenum cofactor(Moco).<sup>11</sup> Molybdenum cofactor is synthesized by a biosynthetic pathway, divided into four steps, contains intermediates that are stable and can be isolated and studied intensively in *Escherichia coli*.<sup>15,41</sup>

These steps of biosynthesis are: cyclic pyranopterin monophosphate (cPMP, also known as precursor Z) synthesis from GTP, conversion of cPMP into (MPT) by insertion of two sulfur, and introduction of molybdenum to form Moco. Moreover, six proteins are identified that are involved in catalyzing Molybdenum cofactor biosynthesis pathway in humans, bacteria, plants and also in fungi.<sup>2, 5, 11, 31</sup>

In humans these proteins are MOCS1 (A&B), MOCS2 (A&B), MOCS3, and gephyrinG/E. Synthesis of cPMP takes place inside mitochondria by the help of protein MOCS1. Cyclic Pyranopterin monophosphate (cPMP) is first stable intermediate in moco biosynthesis. After its synthesis in mitochondria it is transported out to cytosol. Subsequent steps of biosynthesis that is, synthesis of MPT and finally of Moco, takes place in cytosol.<sup>12,13,14</sup>

### 2.2.1 Synthesis of MPT Dithiolate

Formation of MPT from cPMP is the second step in Moco biosynthesis, where two sulfur atoms are introduced into the cPMP. An enzyme, MPT synthase, catalyzes this reaction. MPT synthase is a hetero-tetramer composed of two small and two large subunits of MOCS2A and MOCS2B, respectively.<sup>5, 11</sup> The sulphur is attached to MOCS2A at its C-terminus as thiocarboxylate (Figure 2.1).

From uniprot (<https://www.uniprot.org/uniprot/O96007>) binding pocket residues (143-144 and 166-168) were found. These residues are located in larger subunit (4AP8)<sup>26</sup> also called as catalytic subunit of MPT synthase as substrate binding residues, but were not found specifically for cPMP ligand. It also found that the smaller subunit (5MPO)<sup>27</sup> of MPT synthase is important as sulfur carrier subunit and terminal glycine of this unit provides sulfur to cPMP for reaction process. This study was limited only for bacteria and it was taken as reference to find complete reaction mechanism in humans.<sup>4,7</sup>

In a separate reaction, sulphur is transferred to the smaller subunit (MOCS2A) to reactivate it for next cycle (transfer of sulfur). This process of re-activation and transfer of sulphur is performed by another enzyme MOCS3 (MPT synthase sulfurase), and its amino acid, cysteine is the donor of reactive mobile Sulphur.<sup>12,13,25</sup>

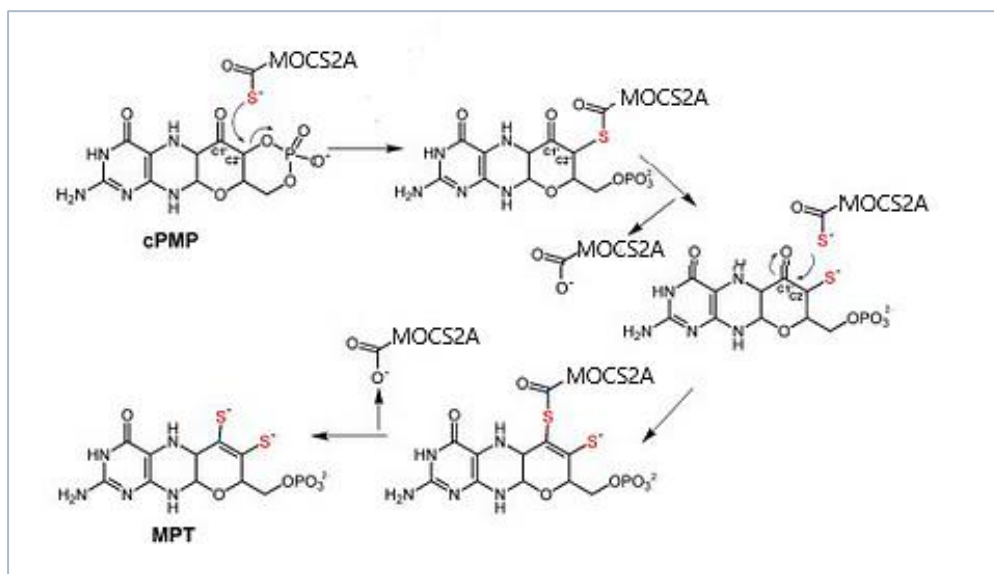


Figure 2.1: Pathway for synthesis of MPT from cPMP<sup>31</sup>

### 2.3 Moco Deficiency

Molybdenum cofactor deficiency in humans is a recessive disorder. It is rare inborn disease and first reported in 1978. Deficiency can be caused by mutations in any of the moco biosynthesis step. Genetic deficiency in any of enzyme, such as MPT synthase, involved in MoCo biosynthesis lead to MoCo deficiency, a rare disease that results in severe neurological abnormalities slow brain growth, feeding difficulties, untreatable seizures and and in some cases also caused death in early childhood.<sup>40,41</sup>

It was identified by biochemical studies that there are two types of patients that were classified as Molybdenum cofactor deficiency (MoCD) ‘Type A’ and ‘Type B’. Defect in the formation of cPMP due the the enzymes (MOCS1A/B) that are involved in this step, is

named as type A deficiency, and lack of the enzymes (MOCS2A/B) caused type B deficiency as a result MPT formation is stopped.<sup>16,17</sup>(Figure 2.2).<sup>39</sup>

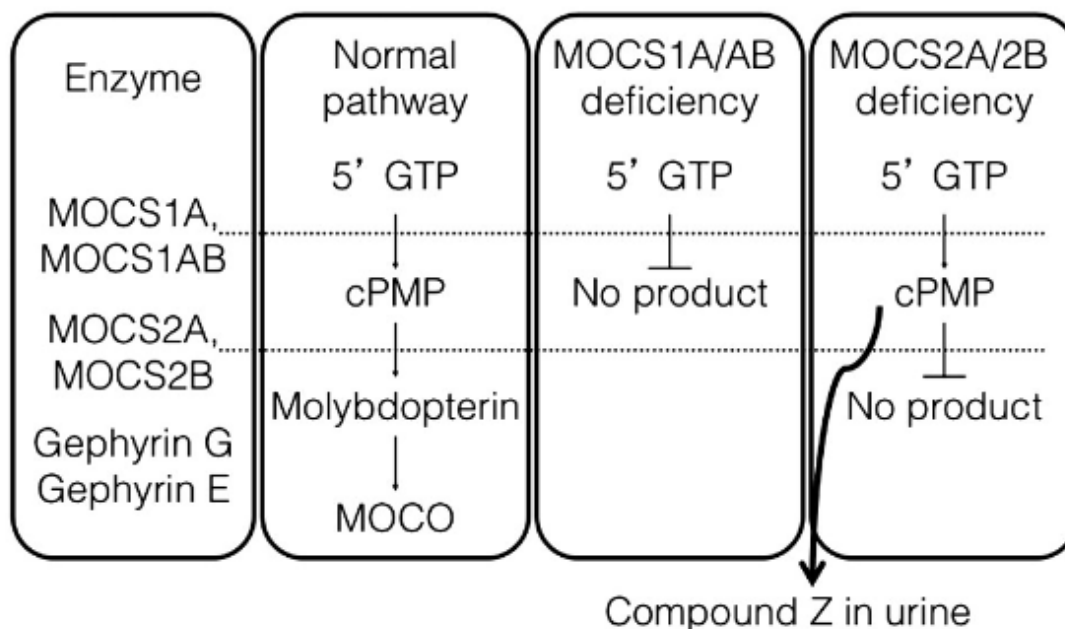


Figure 2.2: Normal/deficient Pathways for Moco Biosynthesis<sup>39</sup>

Symptoms of this disease usually developed after birth, when the neonate started feeding and its process of metabolism in body started to work. It is found that toxic metabolites (mainly sulfite), accumulated within the body. Later, it is also found that two third of the patients have a defect in the formation of first step of Moco biosynthesis (formation of cPMP-precursor Z) and classified with type A deficiency. First treatment of a patient with cPMP has also been performed recently, that had been diagnosed with the deficiency on 6<sup>th</sup> day of life and treatment was started on 36<sup>th</sup> day of life by providing cPMP in the form of injection.<sup>40, 41</sup>

It is noticed that some patients could not survive by providing cPMP supplementation, because they had mutation in MOCS2 genes- this was named as type 2 deficiency, but no mutation was found in MOCS3. It is also seen that GPHN gene mutations can cause severe forms of Moco deficiency.<sup>41, 42</sup>



## **Problem Statement**

Molybdenum cofactor (MOCO) deficiency is a rare, metabolic disorder and characterized by severe neurological abnormalities caused by the functional loss of sulfite oxidase-molybdenum dependent enzyme in affected patient. Until recently, no effective therapy is available. So it is important to know the detailed reaction mechanism for synthesis of MPT (that is actually second step in row) in MoCo biosynthesis, and also the role of copper using in this step. This work will become a step towards understanding/ finding the reason for MoCo deficiency and propose its treatment, by understanding its detailed reaction mechanism using Density Functional Theory (DFT) approach.

## **Objectives**

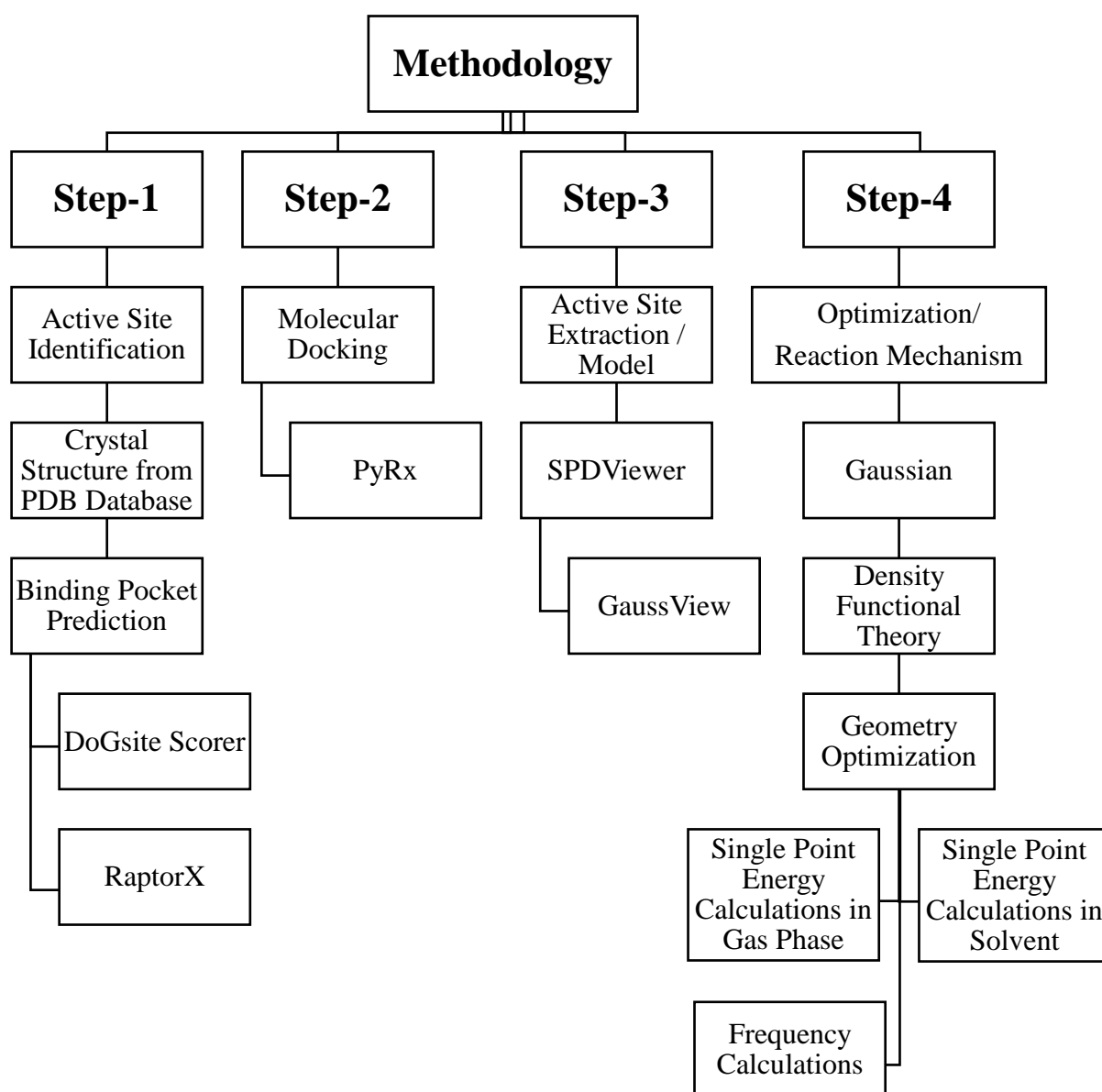
- Identification of active site for cPMP in MPT synthase (MOCS2).
- Understand the reaction mechanism for synthesis of MPT from cPMP.
- Understand the role of copper found in MPT.

# CHAPTER 3

## METHODOLOGY

### 3 Methodology

This research work was carried out to understand the reaction mechanism of MPT synthesis from cPMP with the help of enzyme MPT synthase, which contains two subunits MOCS2A (small) and MOCS2B (large)<sup>12,13</sup>. The project is completed in 04 steps mentioned below:



### 3.1 Active Site Search

MPT synthase contains different binding pockets for different ligands in different species e.g., in Bacterial enzyme (*E.coli*) (PDB ID: 3BII)<sup>23</sup> ligands present are CL (CHLORIDE ION), GL3 (THIOGLYCIN) and (PDB ID: 1NVI)<sup>24</sup> ligands present are SO4 (SULFATE ION), GOL (GLYCEROL). In *Staphylococcus aureus* (PDB ID: 2Q5W)<sup>25</sup> ligand present is GOL (GLYCEROL) and (PDB ID: 2QIE)<sup>25</sup> ligand present is 8CS3 (C<sub>10</sub> H<sub>12</sub> N<sub>5</sub> O<sub>7</sub> P). Whereas in human EDO (1,2-ETHANEDIOL), and GOL (GLYCEROL) are the ligands attached with larger subunit of MPT synthase (PDB ID: 4AP8)<sup>26</sup> but there was no suggested binding pocket for cPMP.

Uniprot data indicated some residues present at the binding site but did not mentioned any ligand. On the other hand, there were no active site or binding site residues spotted for any ligand, shown in smaller subunit of MPT synthase (PDB ID 5MPO).<sup>27</sup> Moreover, literature only indicates that glycine from smaller subunit of MT synthase provides sulfur in the conversion of cPMP to MPT.<sup>3, 4, 6</sup> Keeping in view the above mentioned points, first task was designed to find the active site in MPT Synthase for cPMP through following steps.

#### 3.1.1 Crystal Structure from PDB (Protein Data Bank)

#### 3.1.2 Binding Pocket Prediction

### 3.1.1 Crystal Structure from PDB (Protein Data Bank)

To find the active site of MPT synthase, this is involved in the formation of MPT from cPMP. The X-ray crystal data of protein, MOCS2-MPT Synthase (humans) was obtained from Protein Data Bank with PDB ID: 5MPO.<sup>27</sup> Ligand (cPMP) was extracted from a crystal structure of bacterial Enzyme-MoaC (PDB ID: 4PYD)<sup>28</sup> complexed with cPMP.

### 3.1.2 Binding Pocket Prediction

After obtaining protein and ligand from PDB, next step was the prediction of protein binding pocket for cPMP. For this purpose, following two well-known software for binding pocket prediction were used. Two software were used for comparison and validation of results.

#### 3.1.2.1 DoGsite Scorer<sup>29, 30</sup>

#### 3.1.2.2 RaptorX<sup>31</sup>

#### 3.1.2.1 DoGsite Scorer<sup>29</sup>

DoGsite Scorer is a web-server for automatic binding pocket prediction and analysis. It is newly developed tool and available freely for academics. Files of Protein and Ligand were uploaded in the provided box to run the process and obtain best binding pocket for cPMP<sup>29, 30, 32</sup>(Figure 3.1).

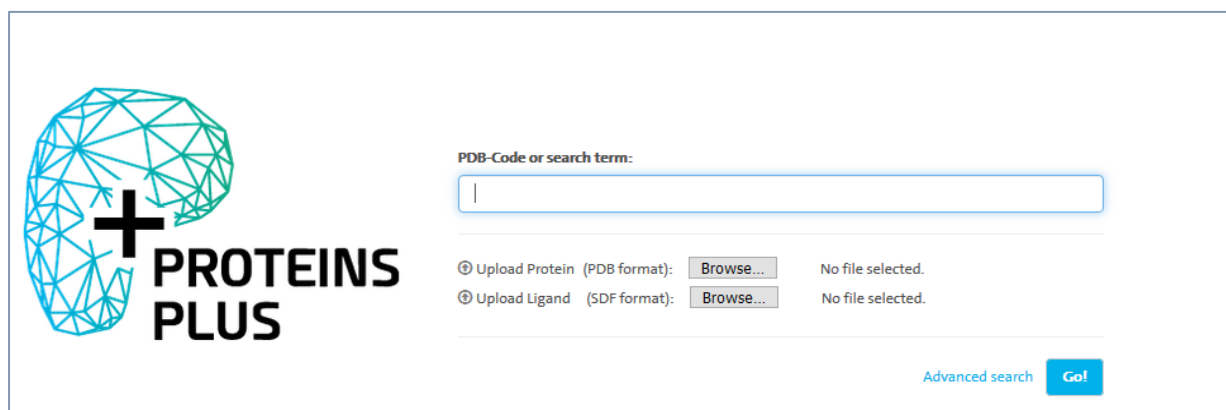


Figure 3.1: Web-Server-GUI of DoGsite Scorer

Results generated from DoGsite Scorer were arranged on the basis of drug score, more the drug score is , more the binding pocket is good for relative ligand. The results are downloaded and viewed in PyMOL<sup>50</sup>.

#### 3.1.2.2 RaptorX<sup>31</sup>

RaptorX is also a web portal for protein structure prediction and binding site prediction, disordered region and protein function. Data of protein and ligand was provided to software.

As a result seven binding pockets were generated for different ligands that can bind with this MPT synthase in some other processes, but at different binding sites. Each pocket contained different multiplicity values. Multiplicity is used to measure the quality of a predicted binding pocket; the higher value is considered better. From results, pockets those were for cPMP ligand was viewed in PyMOL and residues were labeled (THR-77, ARG-79, LYS-159, LYS-166).

## 3.2 Molecular Docking

Molecular Docking is an important tool used in structural molecular biology and drug discovery. Ligand-protein docking is used to predict the binding mode(s) of a ligand with a protein that has a known three-dimensional structure.<sup>20,21</sup> By taking the suggested residues of protein with high drug scores, from binding pocket prediction software, they are forwarded for docking. For docking purpose following step was carried out by using software (PyRx).

### 3.2.1 PyRx

PyRx is virtual screening software used in computational drug discovery to screen compounds against their potential drug targets. PyRx includes Dock vina wizard with easy-to-use user interface, which makes it a good, valuable tool for computer-aided drug design.<sup>33</sup>  
<sup>34</sup> In this study, it is used for docking of protein with ligand (MPT synthase with cPMP).

In PyRx window, first of all protein and ligand is loaded. Next, residues that were from predicted from softwares were selected from protein chains and processed the docking (Figure 3.2). After completion of docking different binding pockets were generated. Results of those pockets along with their binding affinity values were saved and viewed.

These binding affinity values depend upon the interactions between the active site of protein and ligand. Moreover, based on high binding affinity value the relative binding pocket is selected as best docking site for cPMP ligand.

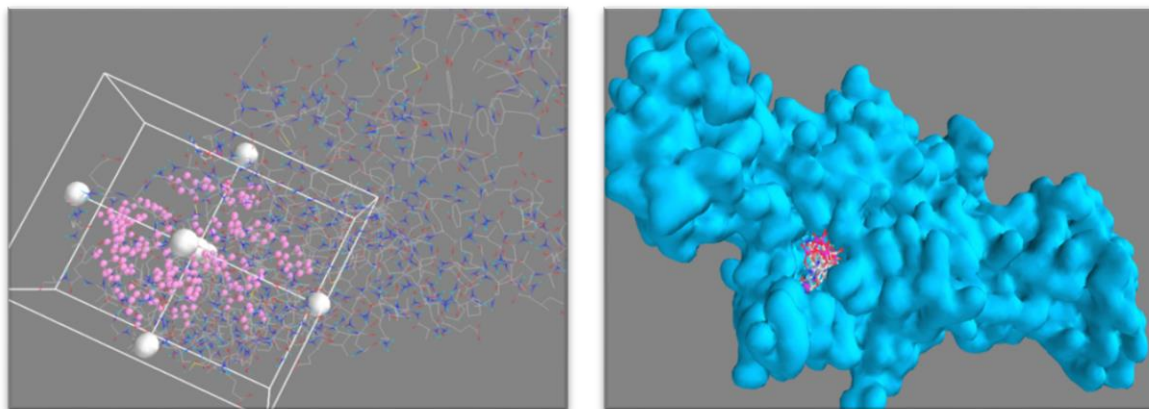


Figure 3.2: Toggled residues in PyRx(A), different confirmation generated by PyRx(B)

### 3.3 Active Site Extraction / Model

After docking process, it was necessary to extract the active site of enzyme from Xray crystal structure of MPT Synthase. Extraction of active site was necessary for fast and accurate optimization as it is not possible to do DFT studies on the whole protein molecule. This process was done with the help of Swiss-Pdb Viewer.

#### 3.3.1 SPDB Viewer<sup>36</sup>

Swiss-Pdb viewer was developed by Nicolas Guex in 1993. It is an application with user friendly interface that allows to analyze different proteins at the same time.<sup>36</sup> Amino acid mutations, distances, angles and H-bonds between atoms are easy to understand and obtain.

The geometry was extracted and then utilized to proceed for next step, generation of model complex geometry by using Gauss view.

#### 3.3.2 Gauss view

Gauss view<sup>37,38</sup> is a graphical user interface for Gaussian. It is used for Gaussian input file generation and also for the visualization of Gaussian output files like optimized molecular geometries, their molecular orbitals (HOMO and LUMO) and UV/IR/NMR graphs etc.<sup>37</sup> By using Gauss View one can also build and edit model structure for Gaussian calculations.

In this study Gauss View is used to edit the amino acids present at the active site of MPT synthase and to add hydrogen atoms wherever required for valency satisfaction.

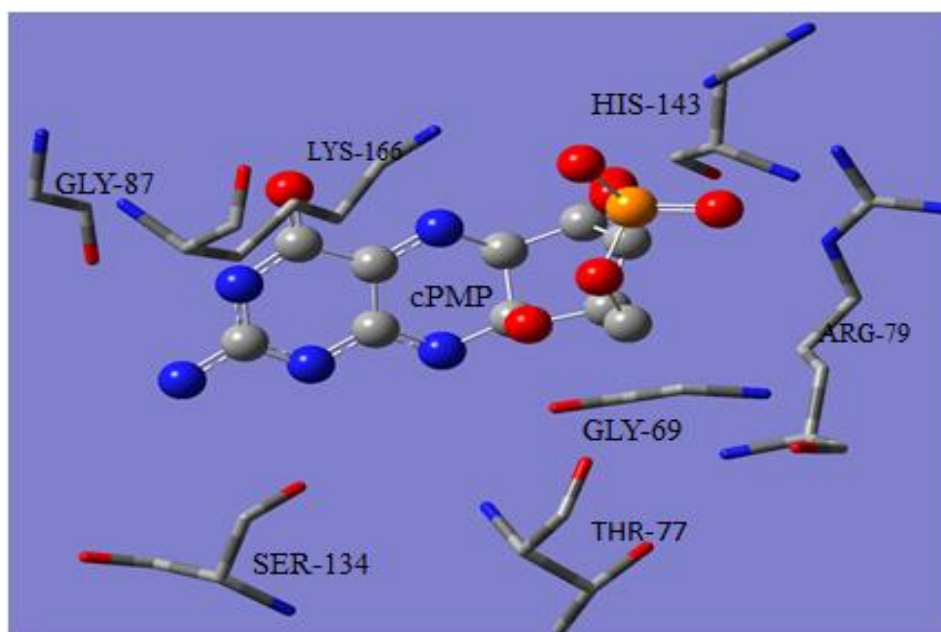


Figure 3.3: Gauss View built and edited model structure

### 3.4 Quantum Studies/Calculations- Reaction Mechanism

Molecular dynamics (MD) simulations are an attractive process for structural refinement and quantum studies of the final docked complex structures. In structure optimization, a Molecular dynamics simulation allows flexibility for both the protein and ligand receptor and also facilitates the relaxation of the complete system. Geometry/Structure optimization is a process/method used to predict the arrangement of the atoms in a molecule in three-dimensions by means of minimizing the energy of a model.<sup>19,20</sup>

Molybdenum cofactor plays an important role in body. By knowing reaction mechanism of Moco biosynthesis it would be possible to understand the changes occurred which cause Moco deficiency and related disease in body. Formation of MPT is the second step in Moco biosynthesis and its reaction mechanism could be understood through following processes.



### 3.4.1 Gaussian-09

Gaussian software was introduced in 1970 by John Pople.<sup>38</sup> The Gaussian package is used for finding energies, molecular geometries, molecular orbitals (LUMO and HOMO), spectroscopy as well as the transition states involved in reaction mechanisms.

In this study, Gaussian 09 is used for DFT studies on the geometries involved in the reaction mechanism of MPT formation from cPMP.

### 3.4.2 Density functional theory (DFT) method

Density Functional Theory (DFT) method is good for studying structure, electronic state, different spectra and for mechanistic studies. In DFT method the properties of a molecule are calculated based on electron density. The energy of a molecule is a function of electron density and electron density is function of electron position in x, y and z coordinates.<sup>20,43</sup>

### 3.4.3 Geometry Optimization/Energy Optimization

Energy Minimization/Geometry Optimization is a process used to find an atomic arrangement in space where the molecule is more stable. The model complex geometry of MPT synthase containing amino acids of Chain-A, Chain-C and Chain-D was considered for geometry optimization and computational analysis.

All the educt, educt-substrate, intermediate and product complex model geometries were optimized using hybrid functional B3LYP<sup>20</sup> and basis set used were LANL2DZ<sup>47</sup>, SDD<sup>47</sup>.

#### 3.4.3.1 Frequency Calculation

Frequency calculations were performed on all the optimized geometries to calculate the nuclear vibrational motion modes in a molecule and to confirm that the minimized geometry has no imaginary frequency. Purpose of frequency calculations was also to get the zero-point correction energies.

### 3.4.3.2 Single point Energy and Self-Consistent Reaction Field Calculation<sup>46</sup>

Single point energies and Self-Consistent Reaction Field calculations<sup>46</sup> are used to calculate electronic energy of a specific arrangement of atoms in the gas phase as well as in the solvent phase. For all optimized geometries, the single point energy calculations were performed using Stuttgart Dresden effective core<sup>48</sup> potential in gas phase and in solvent phase.

### 3.5 Molden<sup>49</sup>

In quantum chemical calculations Molden software is used to visualize the output file of Gamess, Molpro (ab initio calculation), Mopac (semi empirical calculation) and Gaussian softwares.<sup>44</sup>In this research Molden package is used for visualizing the Gaussian output files involved in reaction mechanism of MPT Synthase.

# CHAPTER 4

## RESULTS

### 4 Results

Molybdenum cofactor biosynthesis is a stepwise process that consists of four steps.<sup>2</sup> This study is focused on the computational analysis of synthesis of its second step i.e., the formation of metal binding pterin (MPT) with the help of enzymes MPT synthase (MOCS2).

This study was completed in following four stages:

4.1 Active Sites/Binding Pocket Search

4.2 Molecular Docking

4.3 Active site Extraction / Model

4.4 Quantum Studies/Reaction Mechanism

#### 4.1 Active Sites/Binding Pocket Search

There is no identified binding pocket in enzyme (MPT synthase) for cPMP ligand found in literature. Uniprot data (<http://www.uniprot.org/uniprot/O96007>) only suggests the active site residues located around 143-144 and 166-168 number residues of MPT synthase as substrate binding residues but these binding pocket residues are not for cPMP ligand, therefore it was required to find the binding pocket (active site) for cPMP (Figure4.1).







| Region              |             |  |         |   |        |
|---------------------|-------------|--|---------|---|--------|
| Feature key         | Position(s) | Description  | Actions | Graphical view  | Length |
| Region <sup>1</sup> | 143 - 144   | Substrate binding  UniRule annotation  |         |  | 2      |
| Region <sup>2</sup> | 166 - 168   | Substrate binding  UniRule annotation  |         |  | 3      |

Figure 4.1: Suggested positions for ligand binding, given in uniprot

To obtain binding pocket for cPMP, crystal structure of human MPT synthase is obtained from PDB data base, (PDB ID: 5MPO)<sup>27</sup> and ligand is extracted from a bacterial enzyme (PDB ID: 4PYD)<sup>28</sup>. For active site search, two open source binding pocket prediction software, DoGsite Scoer and RaptorX were used. Purpose of using two software was to ensure the authenticity of our active site binding pocket prediction results.

#### 4.1.1 Results obtained from DoGsite Scorer

First software used for active site search was DoGsite Scorer<sup>29-30</sup>, results obtained show different binding pockets at different site of enzyme or at different chains of enzyme with different residues (Figure 4.2). Different poses and orientations of ligand were found and each binding pose (mode) has different druggability score. In Figure 4.2, these binding pockets (modes) were labeled with different colors as well as arranged on the basis of their druggability scores from high to low. First binding pocket that is of yellow color (P\_0) shows highest drug score of 0.67 and is considered as best binding pocket for cPMP ligand in MPT synthase (Figure 4.2). The selected binding pocket (P\_0) contains, Chain A (GLY-87), Chain C (GLY-69, ALA-70, HIS-143, ARG-144) and Chain D (GLY-76, THR-77, THR-78, ARG-79, ASN-81, PHE-82, GLU-168).

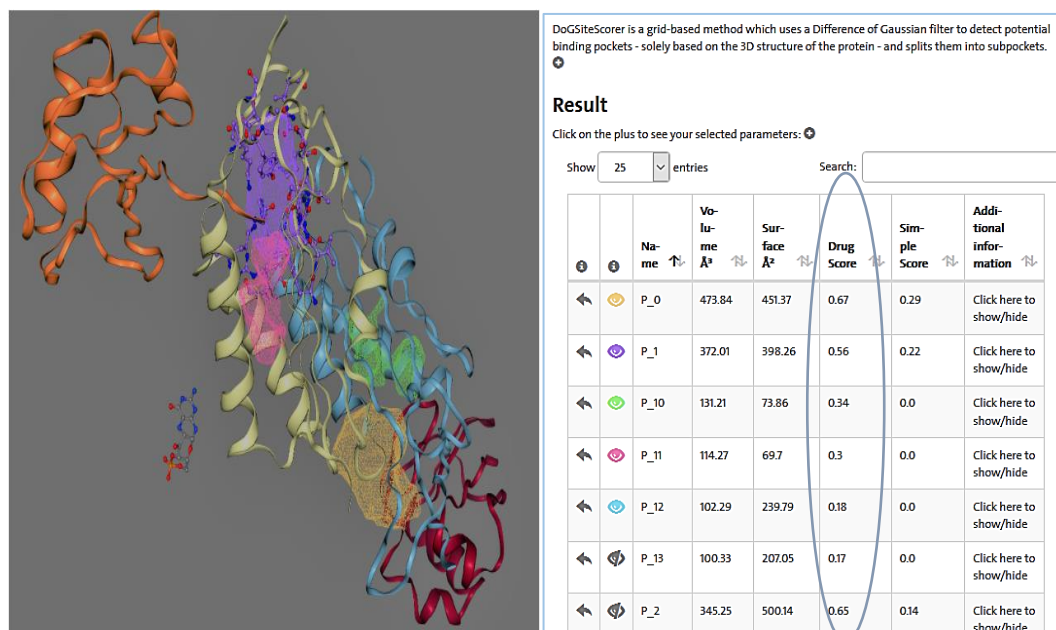


Figure 4.2: Results generated from DoGsite Scorer

Subsequently, the data generated by DoGsite Scorer, chains present in binding pockets along with residues are arranged in table (Table 4.1). For clear understanding these interactions between their residues were also studied using PyMOL (Figure 4.3- Figure 4.6). The binding pockets with their chains and residues present at active site are labeled in Figure 4.3, 4.4. 4.5 & 4.6)

Tabel 4.1: DoGsite Scorer generated binding pocket residues along with chains and Drug Scores

|                    | Chains/Residues  | Drug Score |
|--------------------|--|------------|
| Pocket-1<br>(P_0)  | Chain A (GLY-87),<br>Chain C (GLY-69, ALA-70, HIS-143, ARG-144)<br>Chain D (GLY-76, THR-77, THR-78, ARG-79, ASN-81,<br>PHE-82, GLU-168)  | 0.67       |
| Pocket-2<br>(P_1)  | Chain B (GLY-88)<br>Chain C (GLY-69, GLY-76, HIS-143, ARG-144)<br>Chain D (THR-77, THR-78, ARG-79, ASN-81, PHE-81,<br>PHE-82, GLU-168)   | 0.56       |
| Pocket-3<br>(P_10) | Chain C (LEU-54, SER-55, VAL-56, GLY-68, GLY-69,<br>THR-77, THR-78, ARG-79, SER-131, GLU-132, ALA-133)<br>Chain D (ASP-57, SER-60, GLN-61, VAL-63, ILE-64, SER-<br>65, PRO-66, CYS-68, GLY-69, ALA-70) | 0.34       |
| Pocket-4<br>(P_11) | Chain C (SER-60, GLN-61, ILE-64, SER-65, PRO-66, CSY-<br>68, GLY-69, ALA-70, and SER-140)<br>Chain D (LYS-53, LEU-54, SER-55, ARG-79, SER-131,<br>GLU-132, and ALA-133)                                | 0.3        |

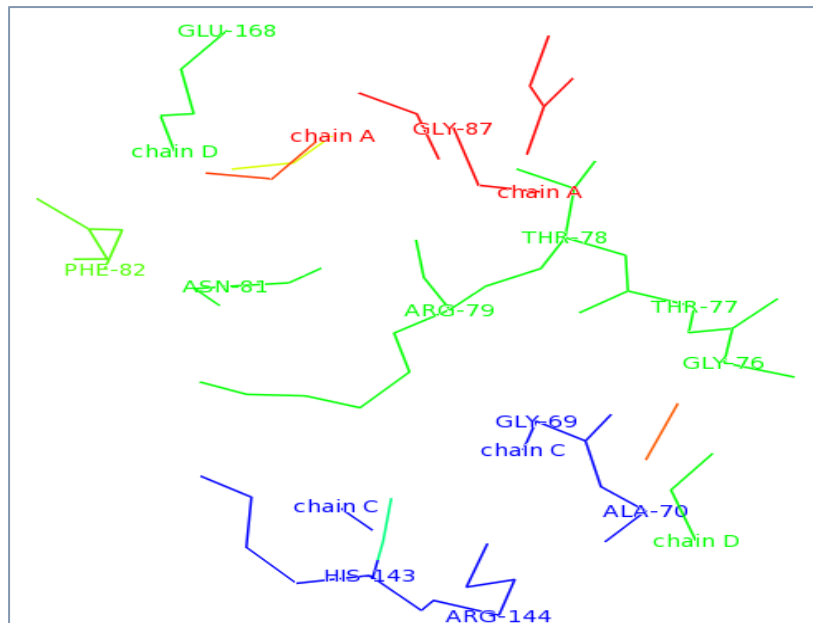


Figure 4.3: PyMOL view of first binding pocket (P\_0)

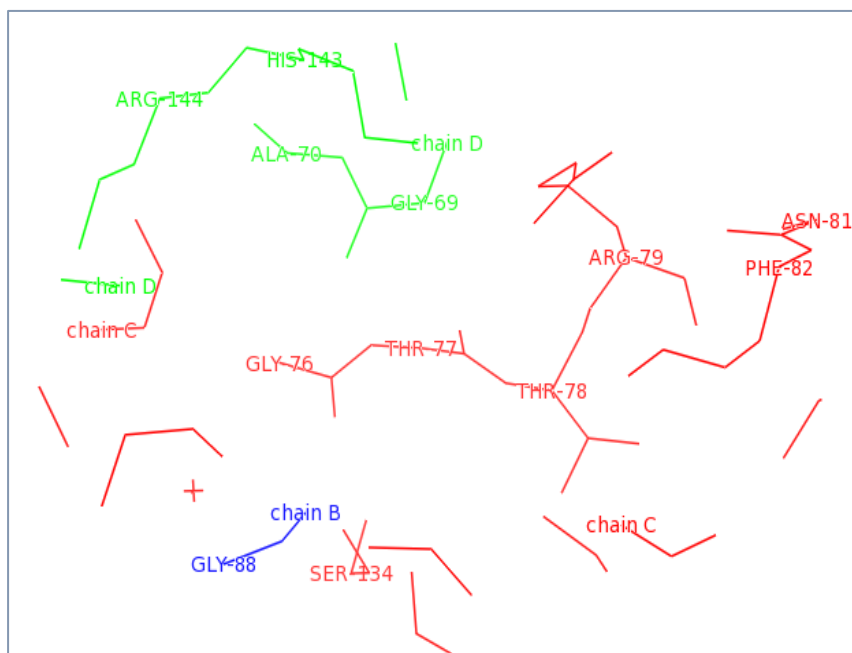


Figure 4.4: PyMOL view of second binding pocket (P\_1)

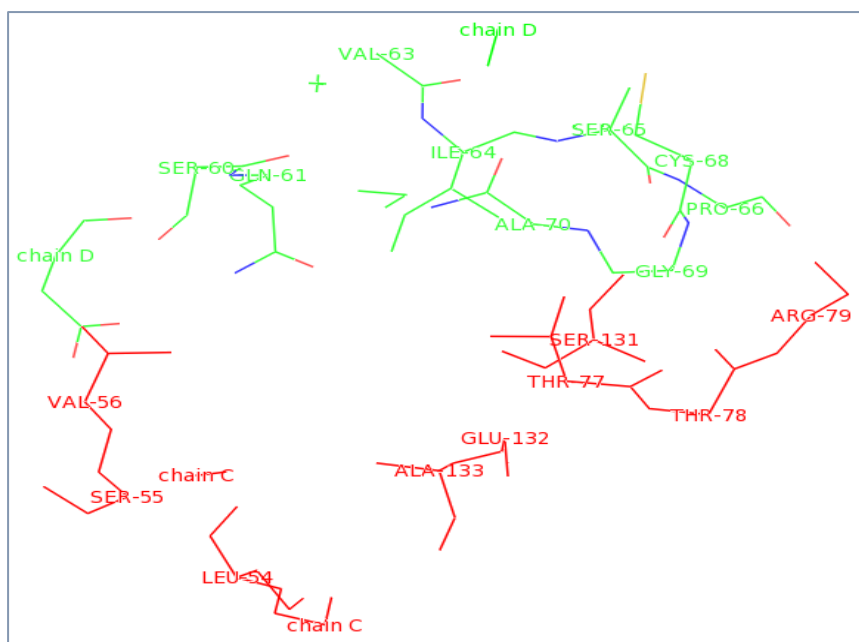


Figure 4.5: PyMOL view of third binding pocket (P\_10)

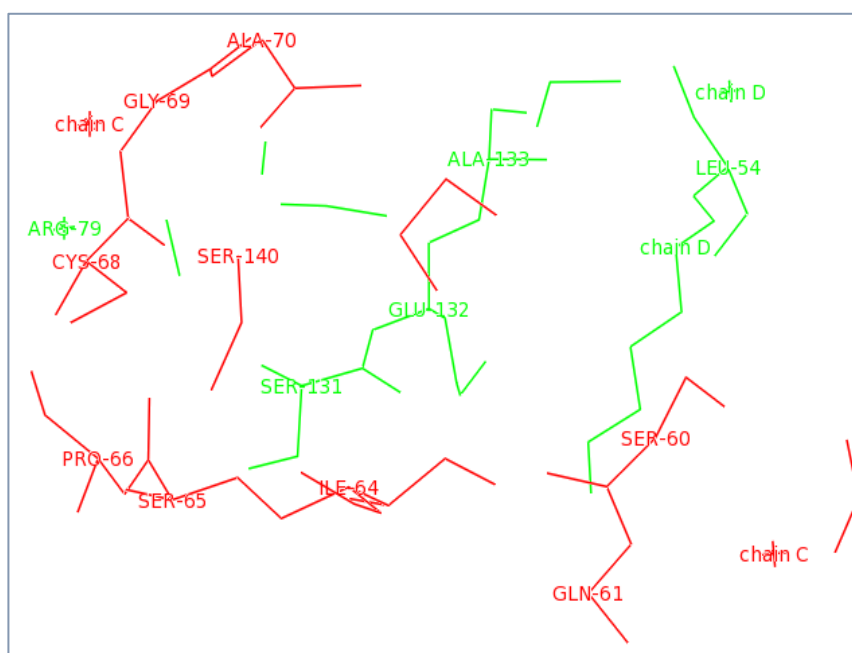


Figure 4.6: PyMOL view of fourth binding pocket (P\_11)

After analyzing all these binding pockets based on their drug score values in DoGsite Scorer, the first binding pocket i.e., P\_0 with highest drug score (0.67) was suggested as the best binding pocket for substrate/ ligand cPMP.



### 4.1.2 Results obtained from RaptorX

To validate our selected binding pocket for MPT synthase, RaptorX<sup>31</sup> software was selected for binding pocket prediction. Results generated from RaptorX shows seven different binding pockets of MPT synthase for different ligands. Each pocket contains different residues for the desired ligand (cPMP). Seven binding pockets, their multiplicity, top ligands and binding residues are presented in table (Table 4.2).

It has been observed from the results of raptorX that only first two binding pockets i.e., 1 & 2 with multiplicity (binding residues) 17(H143, R144) and 14(F74, G76, T77, R79, K159, K166), respectively were suitable for our desired ligand i.e., cPMP, rest of the pockets were not showed pocket for cPMP. Other binding pockets suggested were for binding of ligands other than cPMP. Therefore, Pocket-1, based on high multiplicity (17), was taken as best suggested binding pocket from the results obtained from RaptorX. Results generated from Raptor X were viewed in PyMOL and presented in Figure 4.7 & 4.8.

Tabel 4.2: Binding pockets residues from RaptorX of MPT synthase for cPMP ligand

| Pocket | Multiplicity | Top Ligand | Binding Residues                      |
|--------|--------------|------------|---------------------------------------|
| 1      | 17           | 8CS-501    | H143, R144                            |
| 2      | 14           | 8CS-501    | F74 ,G76, T77, R79, K159, K166        |
| 3      | 8            | DC         | N48, F49, L96, E100, N101, R104, C107 |
| 4      | 6            | NA         | S131, E132, A133                      |
| 5      | 6            | EDO        | A94, Y95, L96, P97, M98               |
| 6      | 5            | GOL        | L54, E58, V59, F122                   |
| 7      | 4            | NA         | V56                                   |

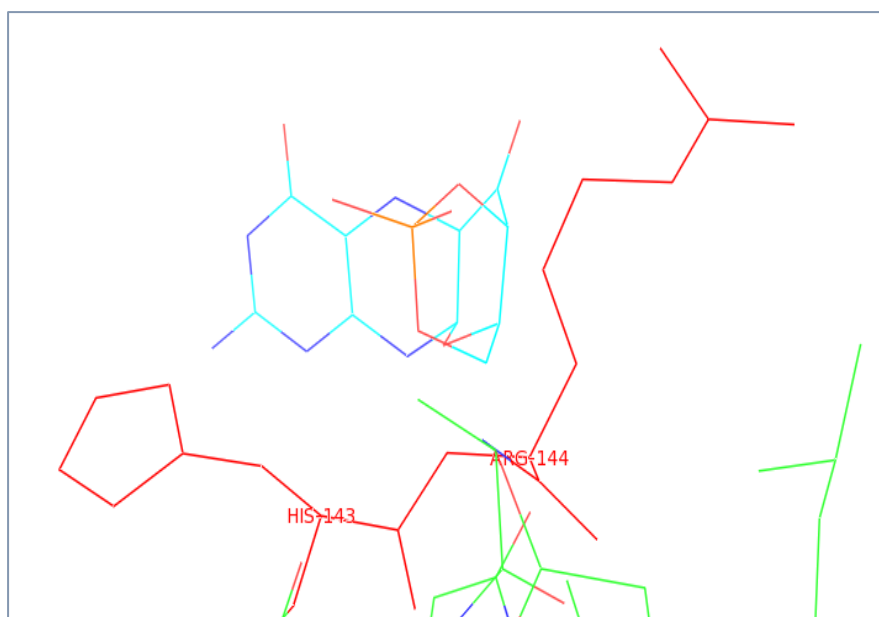


Figure 4.7: First Binding pocket (1) for cPMP

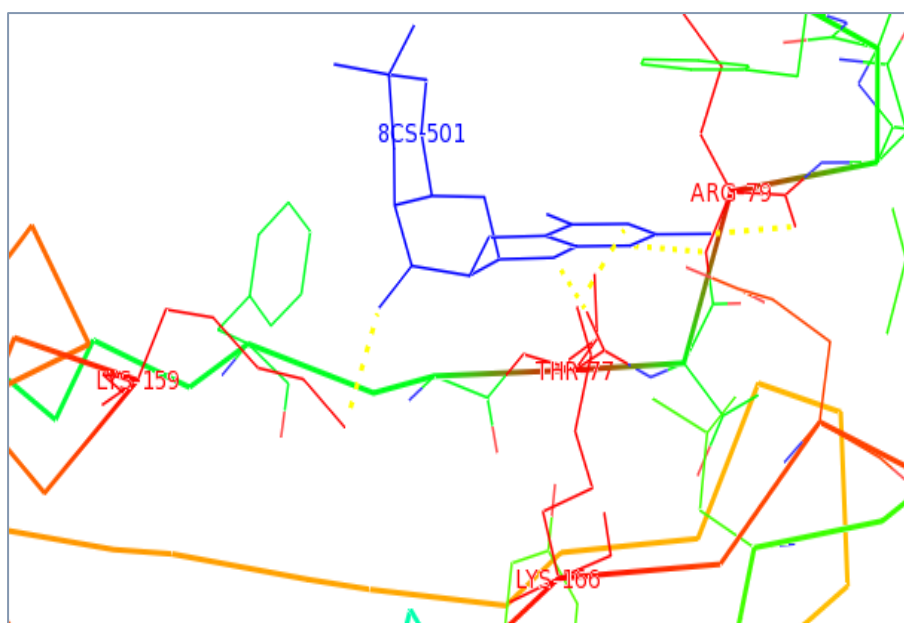


Figure 4.8: Second Binding pocket (2) for cPMP

After comparing the results obtained from DoGsite Scorer and RaptorX, it has been observed that there are some residues common in active site binding pockets for cPMP suggested by both softwares used for the identification of Active site/binding pocket

search, this validated the pocket as being ideal for cPMP. As a result, following chains (residues) were selected for molecular docking study, which is the next step of this research.

- Chain A (GLY-87)
- Chain C (GLY-69, ALA-70, HIS-143, ARG-144, LYS-159, LYS-166)
- Chain D (GLY-76, THR-77, THR-78, ARG-79, ASN-81, PHE-82, GLU-168)

## 4.2 Molecular Docking

For docking purpose, GLY-87, GLY-69, ALA-70, HIS-143, ARG-144, LYS-159, LYS-166, GLY-76, THR-77, THR-78, ARG-79, ASN-81, PHE-82, GLU-168 residues present in the active site binding pocket of MPT synthase were considered.

Molecular docking was done by PyRx and it shows 09 different binding pockets with different binding affinity scores which are populated in the form of table (Table 4.3) with the binding affinity (kcal/mol) values. Every pocket show different binding affinity scores, like Protein\_Mligand-1 shows binding affinity score of -9.4, Protein\_Mligand-2 shows -9.3, Protein\_Mligand-3 shows -8.6, and so on. The software to binding pockets (Table 4.3) gave this name Protein\_Mligand.

According to the literature, more negative the binding affinity value better is the orientation of ligand in that binding site. Therefore, Protein\_Mligand-1(-9.4) is considered as best binding pocket for cPMP ligand as its binding affinity for the orientation of ligand is high in negative value (Table 4.3).

Tabel 4.3: Docking Results, generated by PyRx

| S.No | Pocket/Ligand     | Binding Affinity<br>Kcal/mol |
|------|-------------------|------------------------------|
| 1    | Protein_Mligand-1 | -9.4                         |
| 2    | Protein_Mligand-2 | -9.3                         |
| 3    | Protein_Mligand-3 | -8.6                         |
| 4    | Protein_Mligand-4 | -8.5                         |
| 5    | Protein_Mligand-5 | -8.3                         |
| 6    | Protein_Mligand-6 | -8.2                         |
| 7    | Protein_Mligand-7 | -8.0                         |
| 8    | Protein_Mligand-8 | -7.4                         |
| 9    | Protein_Mligand-9 | -7.4                         |

Complete interaction of residues between protein and ligand was also visualized online in PLIP<sup>35</sup> (Protein-Ligand Interaction Profiler) by uploading docked protein (MPT synthase-cPMP). PLIP showed that interactions between ligand and protein were present through Chain-A (Gly-87), Chain-C (Gly-69, His-143) and Chain-D (77THR, ARG-79, SER-134, LYS-166) (Figure 4.9).

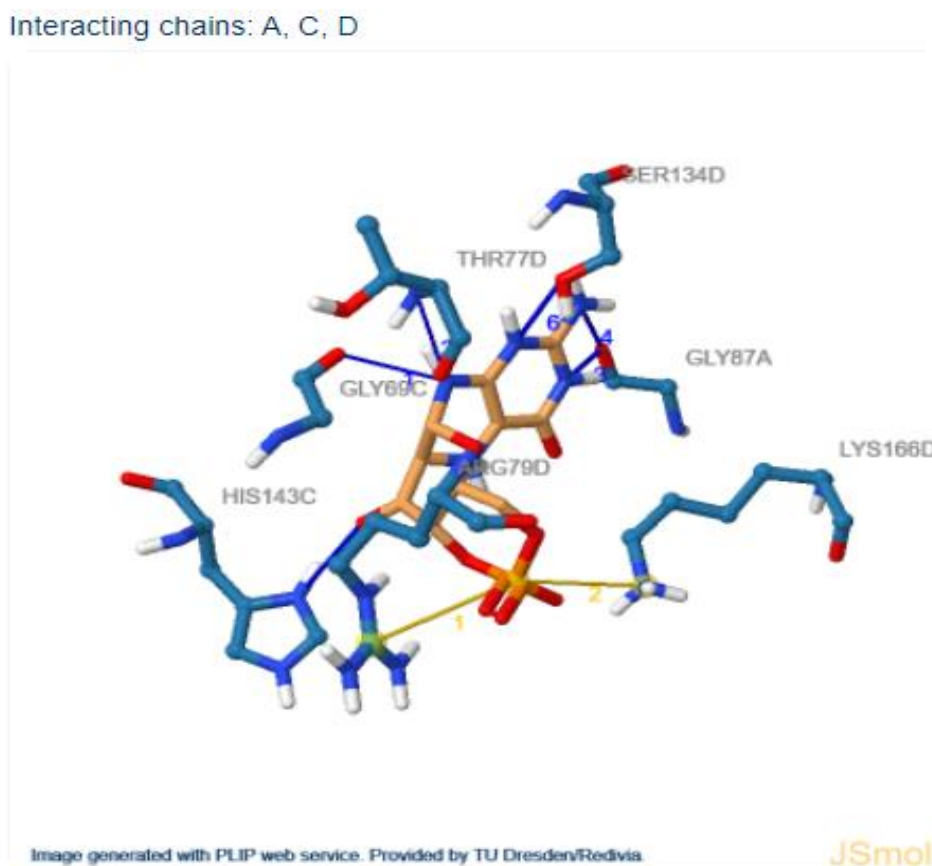


Figure 4.9: Protein-Ligand Interactions after docking (PLIP view)<sup>35</sup>

### 4.3 Active Site Extraction/Model

After docking process, it has been confirmed that residues of Chain-A (Gly-87), Chain-C (Gly-69, His-143) and Chain-D (77THR, ARG-79, SER-134, LYS-166) present at the active site of MPT synthase are involved in MPT synthesis of from cPMP. These residues were then extracted from crystal structure of MPT Synthase (PDB ID: 5MPO)<sup>27</sup> using Swiss-Pdb Viewer<sup>36</sup>. Extraction of desired amino acids from active site binding pocket of protein is necessary for fast and accurate optimization as due to computational cost and time it is not possible to do DFT studies on the whole protein molecule. Therefore, amino acid residues present within 8Å region of the active site and are in close proximity to cPMP-ligand were selected and extracted for quantum studies i.e., DFT analysis (Figure4.10).

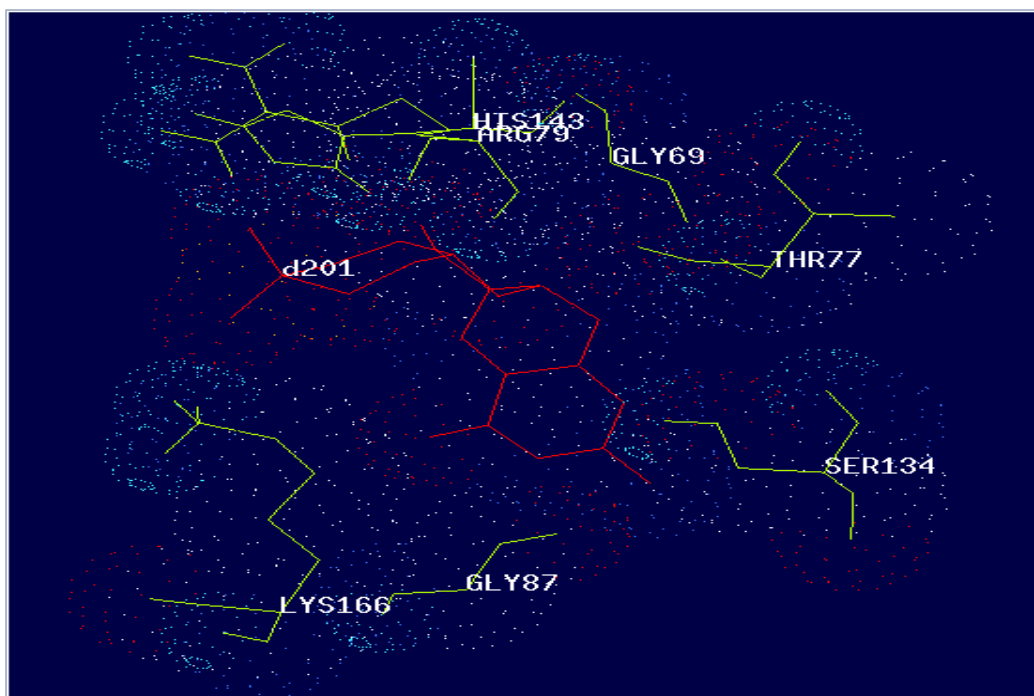


Figure 4.10: Active site of MPT synthase extracted by Swiss-PDB Viewer

#### 4.4 Quantum Studies/Calculations- Reaction Mechanism

To study the reaction mechanism, amino acid residues extracted from MPT synthase (PDB ID:5MPO)<sup>44</sup> using SPDB Viewer (Figure 4.11) were utilized to generate model complex geometries involved in the formation of MPT from cPMP for Quantum calculations. The proposed reaction mechanism to be studied using DFT is presented in Figure 4.13.

##### 4.4.1 Educt Complex Geometry (E)

To model active site model geometry (educt) amino acid residues extracted were truncated and modified wherever required. Gly-87, Gly-69, Ser-134 and HIS-143 were truncated at their  $\alpha$  carbons, Arg-79 was truncated at nitrogen atom attached with  $\gamma$  carbon and Lys-166 at nitrogen atom attached to the  $\alpha$  carbon. Hydrogen atoms were included wherever required to satisfy the valency.

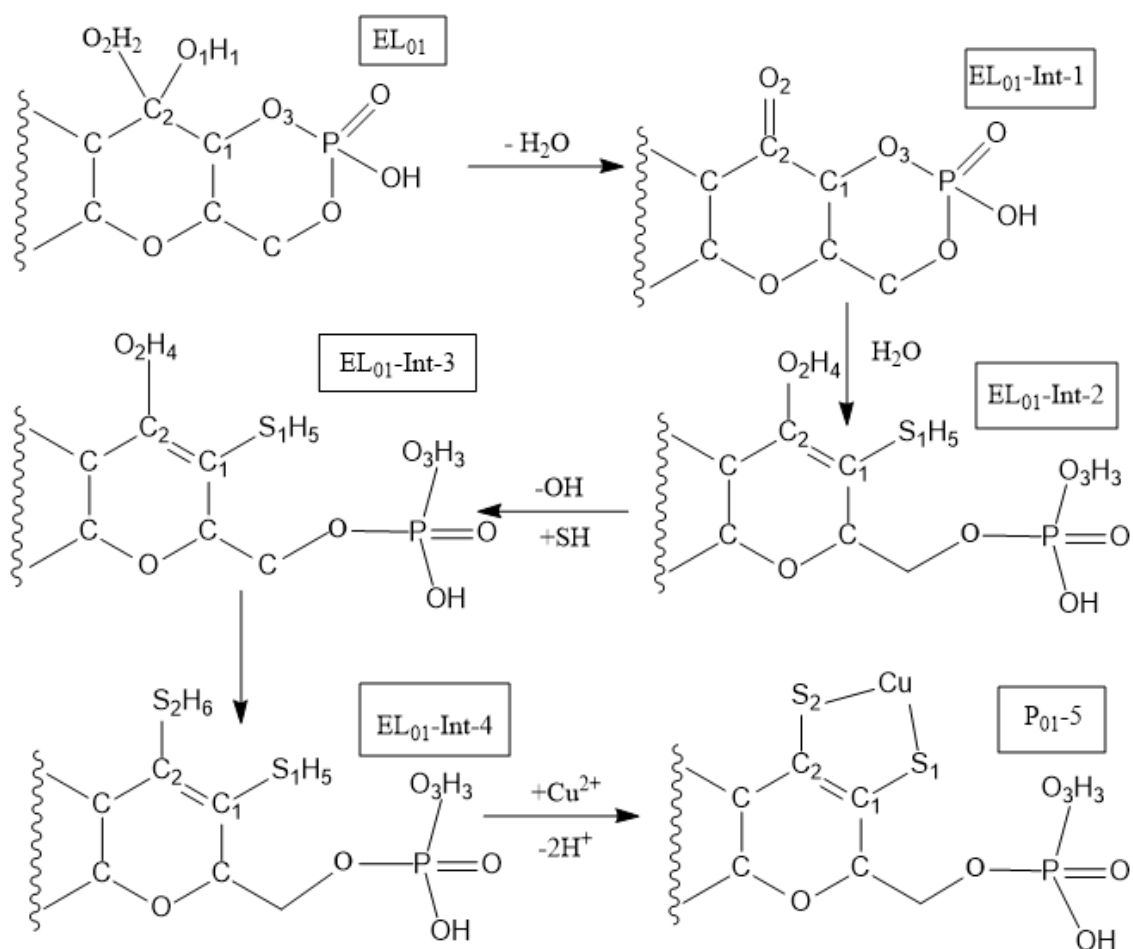


Figure 4.11: Proposed reaction mechanism

In the first step, only hydrogen atoms were optimized while all atoms other than hydrogen were kept fixed. In the second step geometry optimization was carried out on the hydrogen optimized geometry while fixing few atoms of the amino acid residues to keep the residues at their positions of X-ray Crystal structure and also to keep the effect of protein moiety.  $\alpha$  carbon atom of Gly-87, Gly-69, Ser-134 and HIS-143, nitrogen atom attached with  $\gamma$  carbon of Arg-79, and nitrogen attached to  $\alpha$  carbon of Lys-166 were kept fixed for the whole quantum study.

Geometry optimizations were done using B3LYP hybrid density functional and LANL2DZ and SDD<sup>47</sup> basis sets. Single point energies were calculated on the optimized geometries in the gas using B3LYP/LANL2DZ and B3LYP/SDD level of Density Functional Theory as well as in the solvent phase using conductor like polarizable continuum models (CPCM). Frequency calculations were also carried out on all the optimized geometries. All the results obtained from single point energy calculations as well as geometry optimization are presented in this thesis.

#### 4.4.2 Ligand Geometry (L)

Ligand (cPMP) geometry was extracted from X-ray crystal structure of bacterial Enzyme-MoaC (PDB ID: 4PYD)<sup>45</sup>. Two configurations of cPMP ligand considered on the basis of its overall charge are:

4.4.2.1 L<sub>01</sub>: where overall charge is zero, and multiplicity is 1 (Figure 4.12A)

4.4.2.2 L<sub>11</sub>: where overall charge is -1 and multiplicity is 1 (Figure 4.12B).

Purpose of considering two ligand structures is to identify in which configuration cPMP is more reactive.

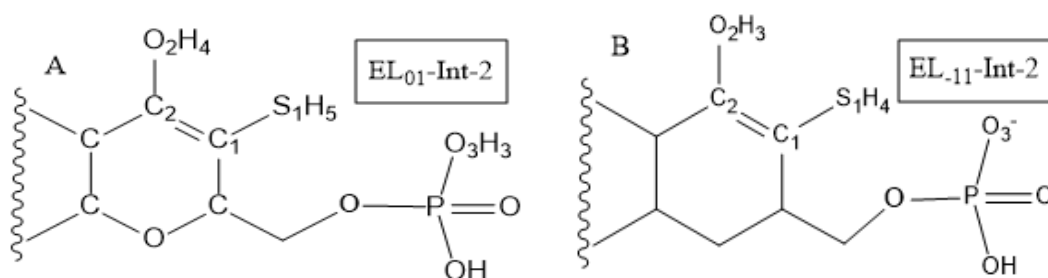


Figure 4.12: Protonated and unprotonated cPMP

#### 4.4.3 Educt-Ligand Complex Geometry (EL)

Reaction mechanism of MPT formation from cPMP involves 04 steps (five structures in each mechanism) starting from the Educt-Ligand Complex, EL. Based on the two



configurations of ligands  $L_{01}$  and  $L_{-11}$ , following two configurations of Educt-Ligand Complex geometry are generated for further study:

#### 4.4.3.1 Educt-Ligand Complex ( $EL_{01}$ )

#### 4.4.3.2 Educt-Ligand Complex ( $EL_{-11}$ )

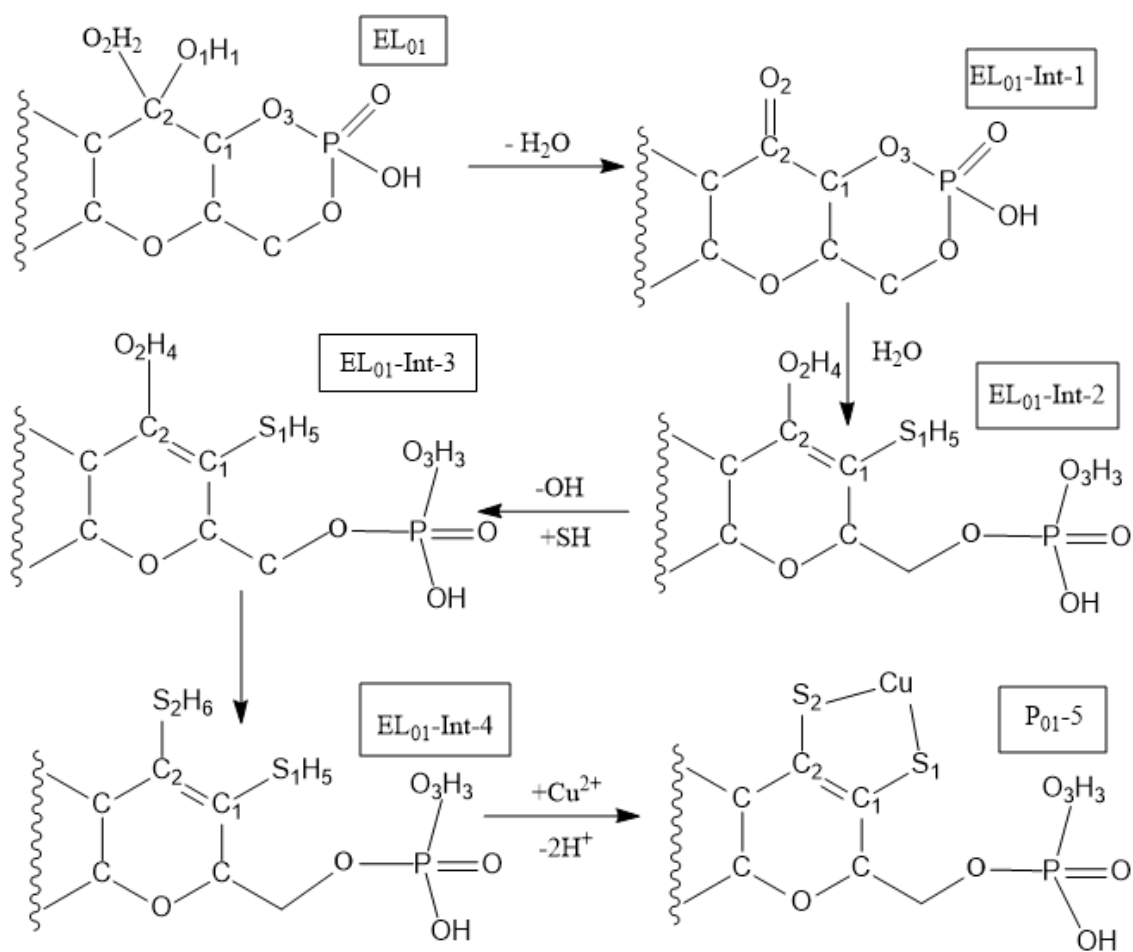
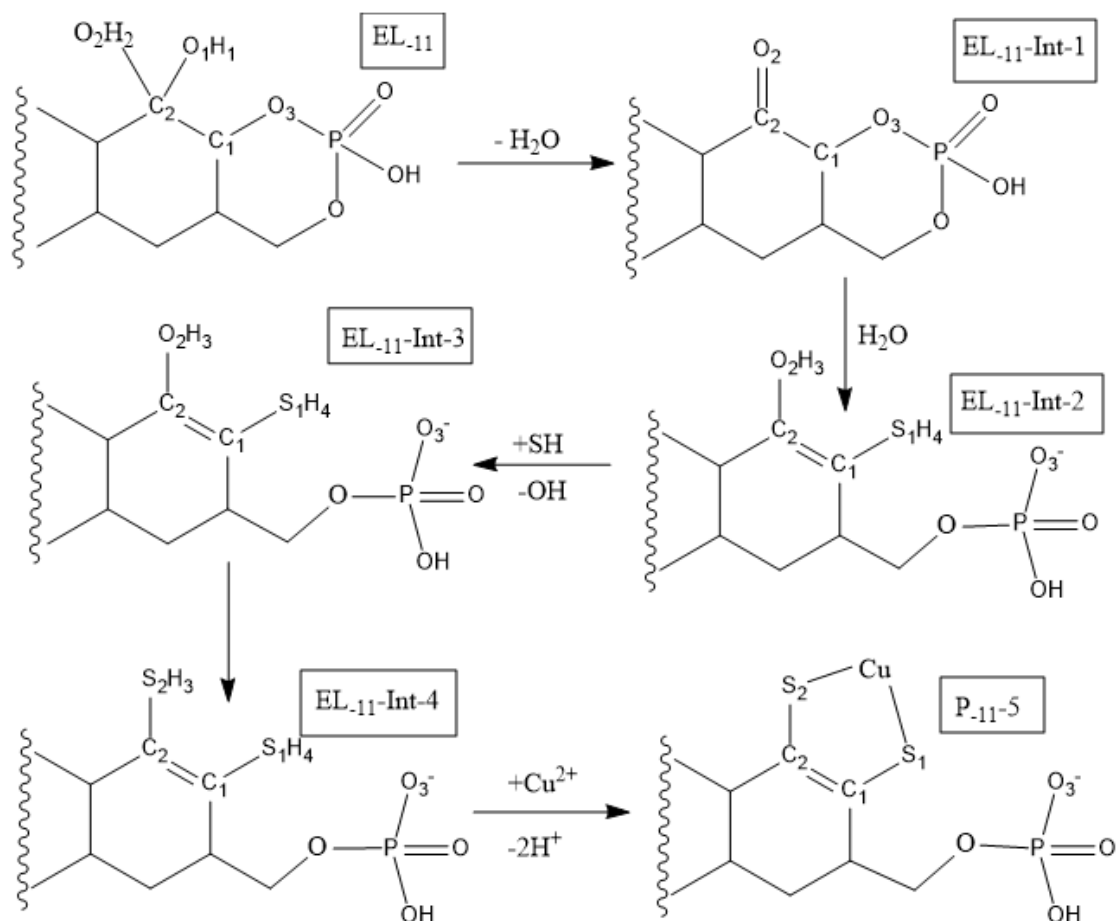


Figure 4.13: Stepwise formation of MTP from  $EL_{01}$  to  $P_{01-5}$

Figure 4.14: Stepwise formation of MTP from EL<sub>11</sub> to P<sub>11-5</sub>

#### 4.4.3.1 Educt-Ligand Complex Geometry (EL<sub>01</sub>)

In EL<sub>01</sub> complex geometry, the overall charge on the active site of protein and ligand complex was kept zero and multiplicity one (0, 1). Results obtained from the computational analysis of EL<sub>01</sub> educt-ligand complex shows no imaginary frequency which indicates that the geometry is fully optimized and at minimum energy on the potential energy surface. Imaginary frequencies arise when one or more of our value is in negative number, after completion of optimization of structure. By getting no imaginary frequencies, we can get the true local minimum at (PES) potential energy surface.

It is observed that the formation of complex EL<sub>01</sub> is exothermic, release energy of 28 kcal/mol with B3LYP/LANL2DZ level of DFT (-28 kcal/mol with B3LYP/SDD),

relative to separate educt and ligand complex geometries (Table 4.4). Optimized geometry indicates two hydroxyl groups  $O_1H_1$  and  $O_2H_2$  present at  $C_2$  position (Figure 4.15) with bond distance of about 1.430 and 1.421 Å respectively. These hydroxyl groups are directly involved in the reaction mechanism. (Table 4.5). Optimized geometries of  $EL_{01}S$  (Educt-Ligand Complex optimized using SDD basis set) and  $EL_{01}L$  (Educt-Ligand Complex optimized using LANL2DZ basis set) are shown in Figure 4.16 and 4.17 respectively.

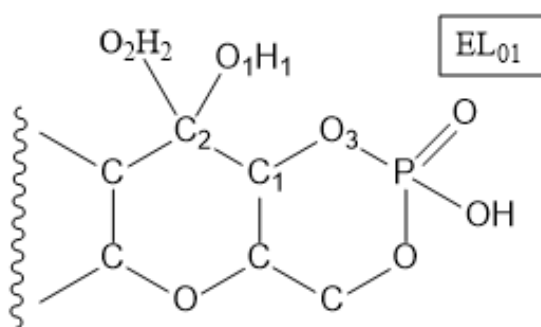
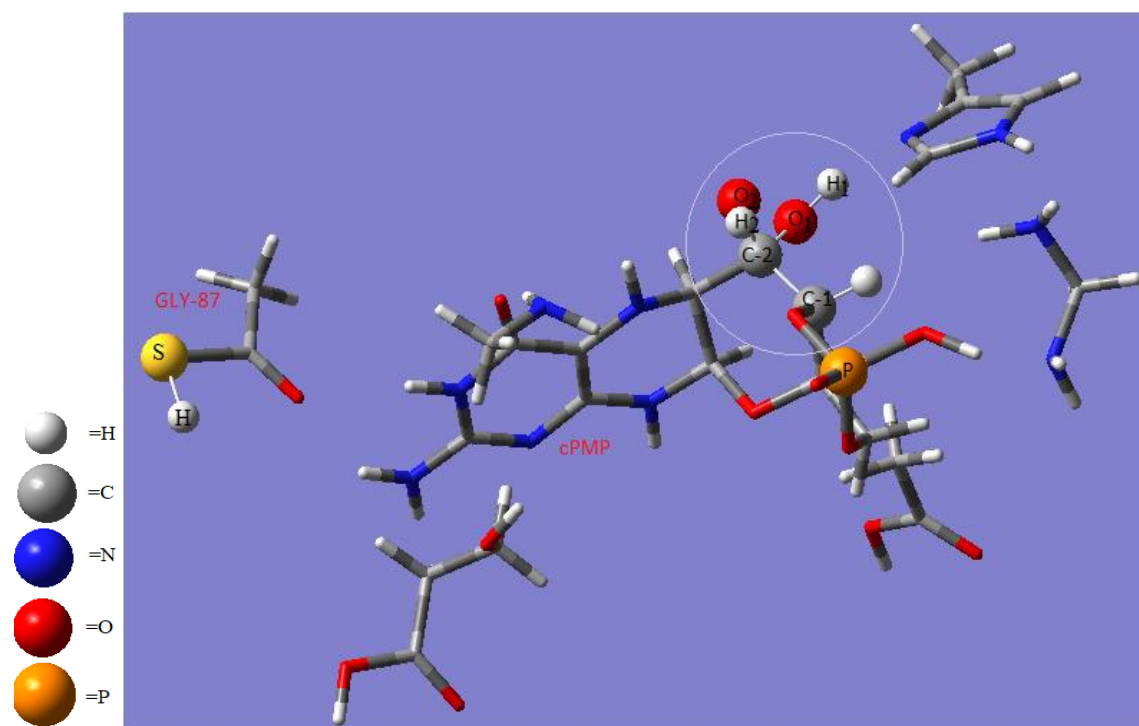
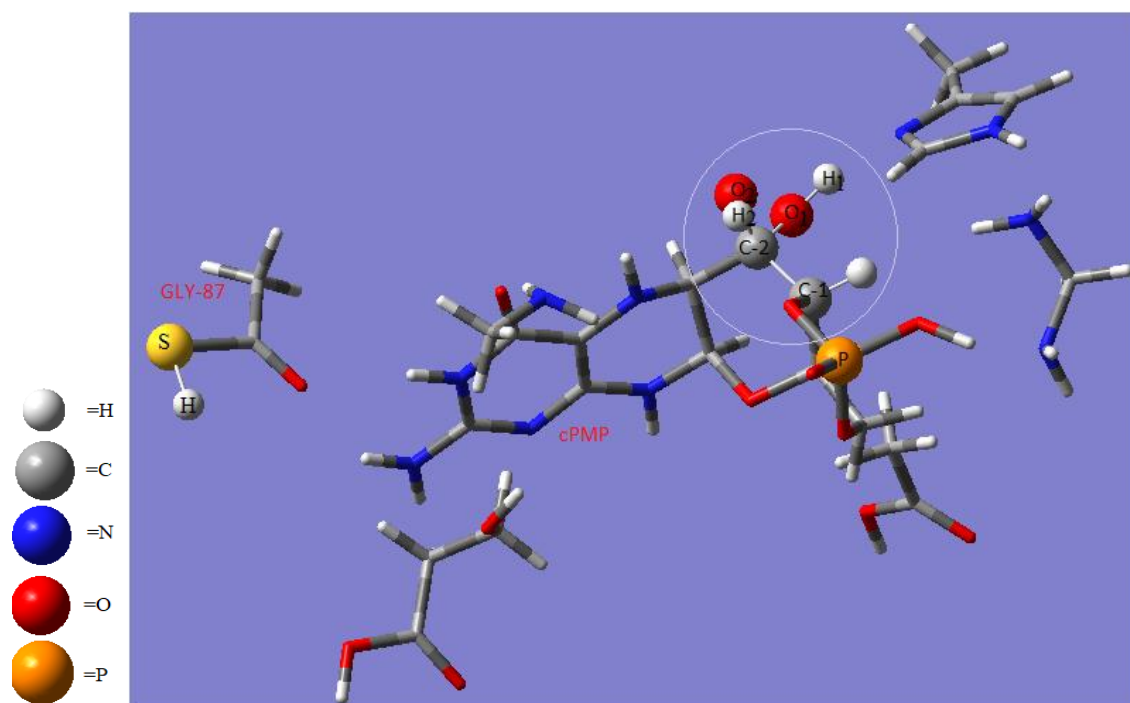


Figure 4.15: Structure of  $EL_{01}$  geometry with two hydroxyl groups at  $C_2$

Figure 4.16: Optimized structure of EL<sub>01</sub>SFigure 4.17: Optimized structure of EL<sub>01</sub>L

#### 4.4.3.1.1 Intermediate Complex Geometry (EL<sub>01</sub>Int-1)

Dehydration of Educt-Ligand Complex EL<sub>01</sub> gives the first intermediate geometry, EL<sub>01</sub>-Int-1, where carbon oxygen (C<sub>2</sub>=O<sub>2</sub>) double bond arises at position C<sub>2</sub> by removal of water (one O<sub>1</sub>H<sub>1</sub> and H<sub>2</sub> is removed as H<sub>2</sub>O) from two hydroxyl groups (O<sub>1</sub>H<sub>1</sub> & O<sub>2</sub>H<sub>2</sub>) present at C<sub>2</sub>, (Figure 4.18). This reaction is neither exothermic nor endothermic relative to separate educt ligand complex geometries as the relative energy obtained on the potential energy surface is -0.3 with B3LYP/SDD level of DFT. However, it is endothermic with B3LYP/LANL2DZ level of DFT as the relative energy obtained is 11 kcal/mol (Table 4.4, Figure 4.19 & 4.20).

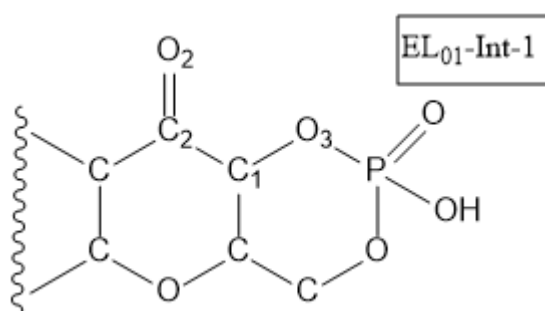
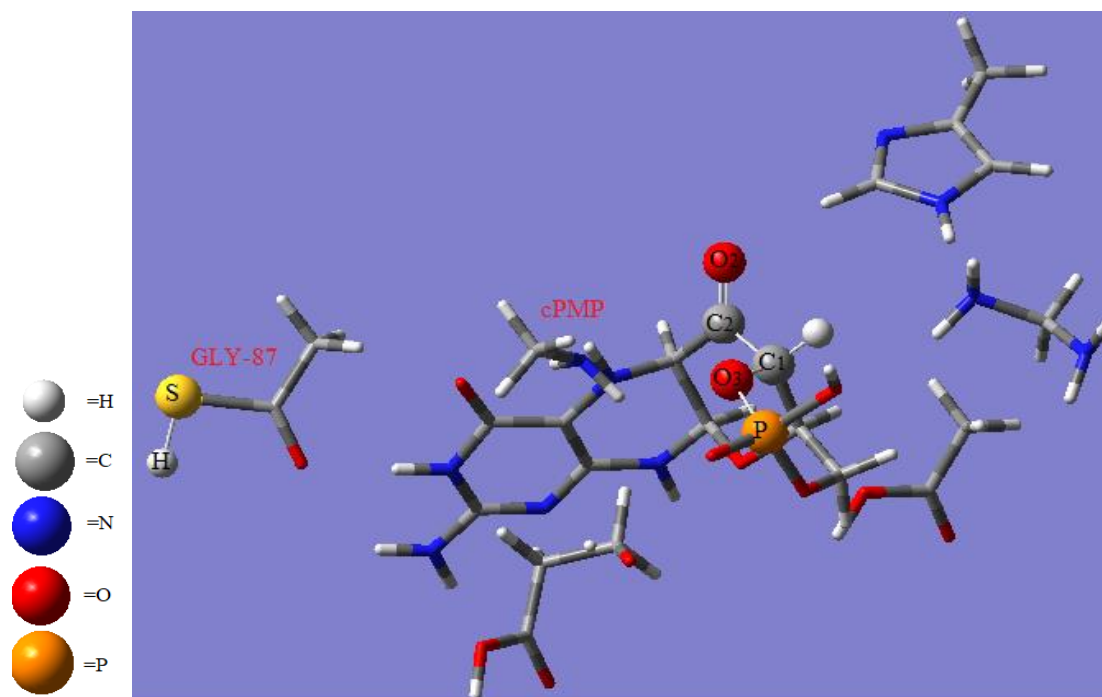
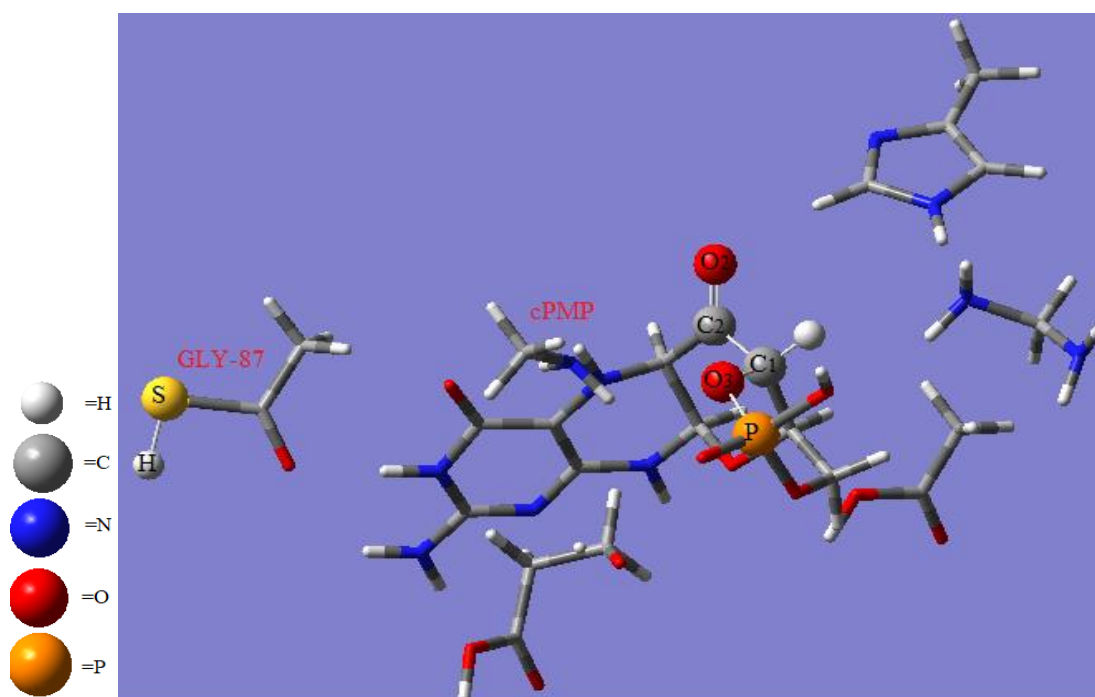


Figure 4.18: Structure of EL<sub>01</sub>-Int-1 geometry with carbon (C<sub>2</sub>) oxygen (O<sub>2</sub>) double bond

Figure 4.19: Optimized structure of EL<sub>01</sub>-Int-1-SFigure 4.20: Optimized structure of EL<sub>01</sub>-Int-1-L

#### 4.4.3.1.2 Intermediate Complex Geometry (EL<sub>01</sub>-Int-2)

At the next step (C<sub>1</sub>-O<sub>3</sub>) bond of intermediate, EL<sub>01</sub>-Int-1 is broken resulting in the breakdown of ring, it gives second intermediate geometry EL<sub>01</sub>-Int-2. The oxygen atom

(O<sub>3</sub>) attached with phosphorus in the ring, becomes protonated by gaining hydrogen atom (O<sub>3</sub>H<sub>3</sub>) (Figure 4.21). Second change that occurs at this step after the breakdown of ring is, the shifting of double bond from (C<sub>2</sub>=O<sub>2</sub>) to (C<sub>2</sub>-C<sub>1</sub>), and sulfide group (SH) of glycine (GLY-87) of smaller subunit of protein (MOCS2A) is attached with the carbon, C<sub>1</sub> as S<sub>1</sub>H<sub>5</sub>. This reaction is exothermic, i.e., 26.3 kcal/mol energy is released with both B3LYP/LANL2DZ and B3LYP/SDD levels of DFT (Table 4.4, Figure 4.22 & 4.23).

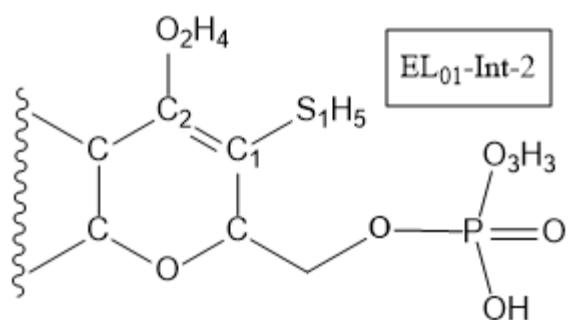


Figure 4.21: Structure of EL<sub>01</sub>-Int-2 geometry with S<sub>1</sub>H<sub>5</sub> group at C<sub>1</sub>

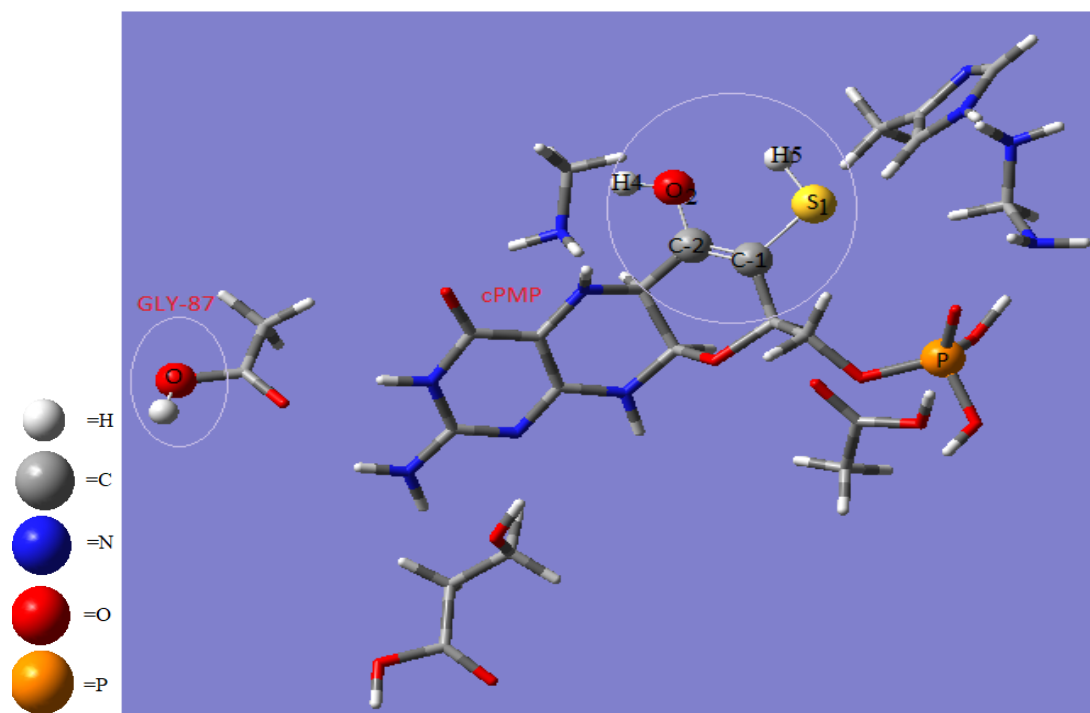


Figure 4.22: Optimized structure of EL<sub>01</sub>-Int-2-S

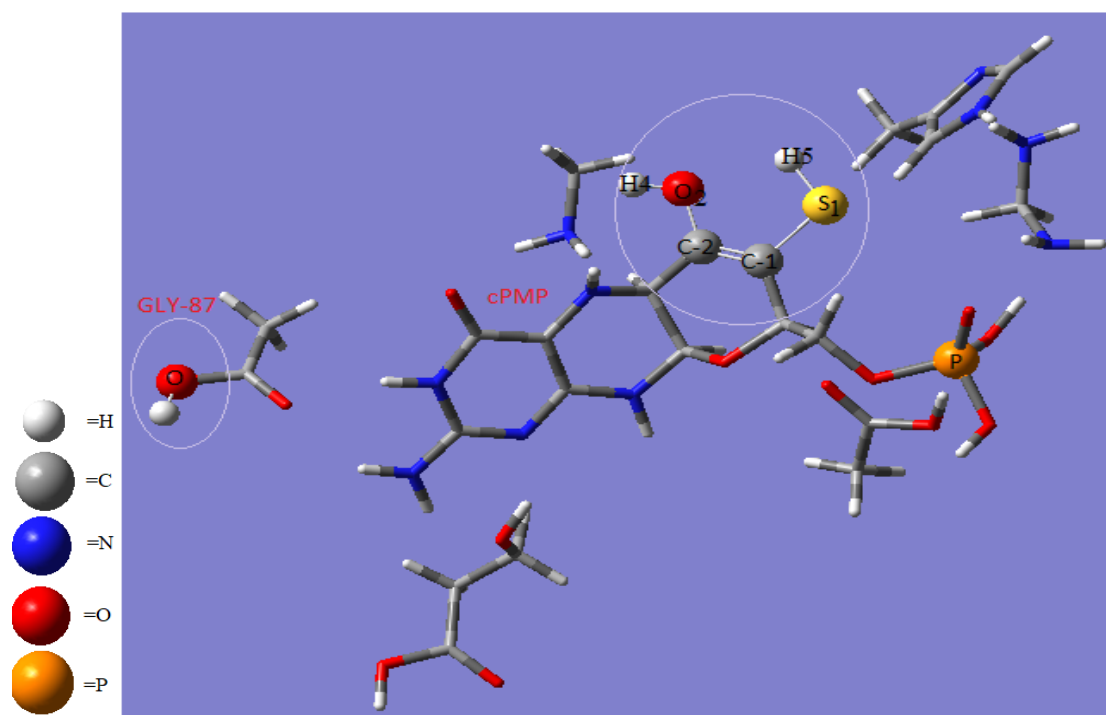


Figure 4.23: Optimized structure of  $EL_{01}$ -Int-2-L

#### 4.4.3.1.3 Intermediate Complex Geometry ( $EL_{01}$ Int-3)

Upon conversion of  $EL_{01}$ -Int-2 to third intermediate geometry  $EL_{01}$ -Int-3, active site remained same, only changes occurred at GLY-87 at Chain A of smaller subunit of MOCS2 protein as a result one SH group is added with help of enzyme MOCS3 (MPT-synthase sulfurase) at by replacing OH of GLY-87 (Figure 4.25 & 4.26)

$EL_{01}$ -Int-3 could not be optimized by charge 0 and multiplicity 1, because of change in structure (addition of SH in GLY-87 of protein chain-1 instead of OH) charge on whole molecule is changed so it is optimized with charge -1 and multiplicity 1. Therefore, instead of using structure with zero charge, another structure with charge -1 and multiplicity 1 was taken for optimization (Figure 4.23). This reaction is endothermic, and absorbed 21.7 kcal/mol energy with B3LYP/LANL2DZ level of DFT and 27.6 kcal/mol with B3LYP/SDD level of DFT relative to the separate protein and ligand (Table 4.4, Figure 4.26 & 4.27).



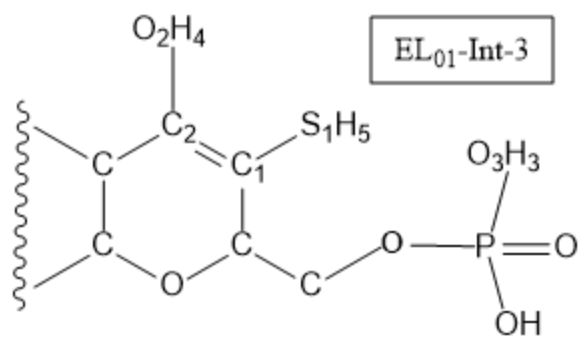


Figure 4.24: Structure of EL<sub>01</sub>-Int-3 geometry with SH group at C<sub>1</sub>

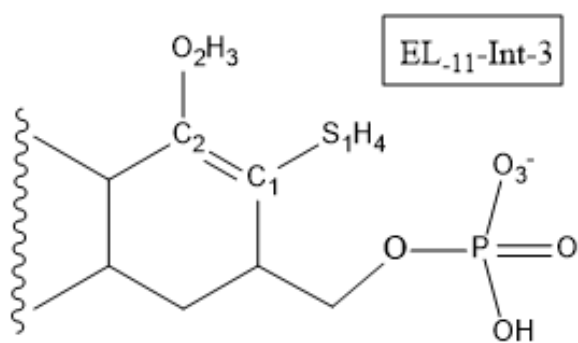


Figure 4.25: Structure of EL<sub>11</sub>-Int-3 geometry with S<sub>1</sub>H<sub>1</sub> group at C<sub>1</sub>

Figure 4.26: Optimized structure of EL-<sub>11</sub>-Int-3-SFigure 4.27: Optimized structure of EL-<sub>11</sub>-Int-3-L

#### 4.4.3.1.4 Intermediate Complex Geometry (EL<sub>01</sub>-Int-4)

From EL<sub>01</sub>-Int-3, EL<sub>01</sub>-Int-4 was formed by replacement of O<sub>2</sub>H<sub>4</sub> with S<sub>2</sub>H<sub>6</sub> group from GLY-87 (MPT synthase) at carbon (C<sub>2</sub>) of cPMP (Figure 4.28). This reaction is endothermic, relative to the separate protein and ligand, with absorption of 9.7 kcal/mol energy with B3LYP/LANL2DZ and 13.6 with B3LYP/SDD level of DFT (Table 4.4, Figure 4.29 & 4.30).

In this step as a side reaction an OH group is added to GLY-87 by removal of hydrogen sulfide (SH group is added as second S<sub>2</sub>H<sub>2</sub> at C<sub>2</sub> of cPMP) (Figure 4.29), and it will move to next step for addition of sulfur with the help of MOCS3 and the cycle goes on for the biosynthesis of MPT from cPMP.

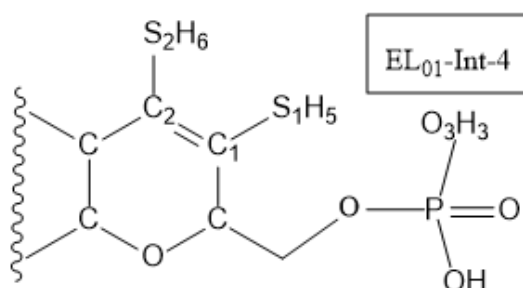
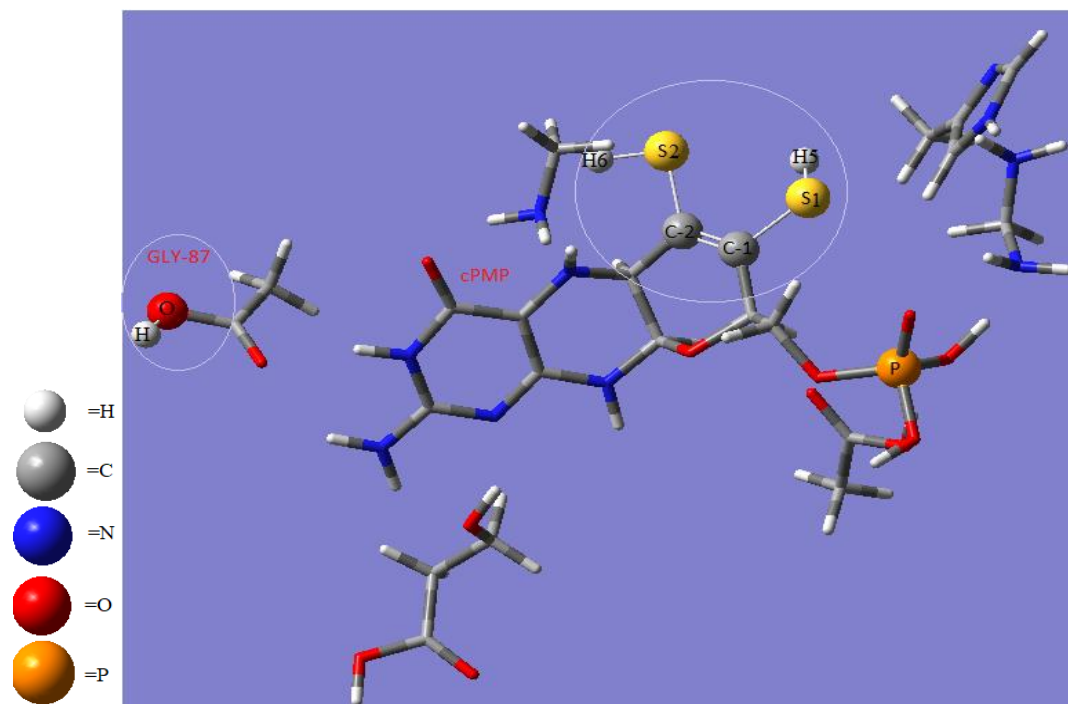
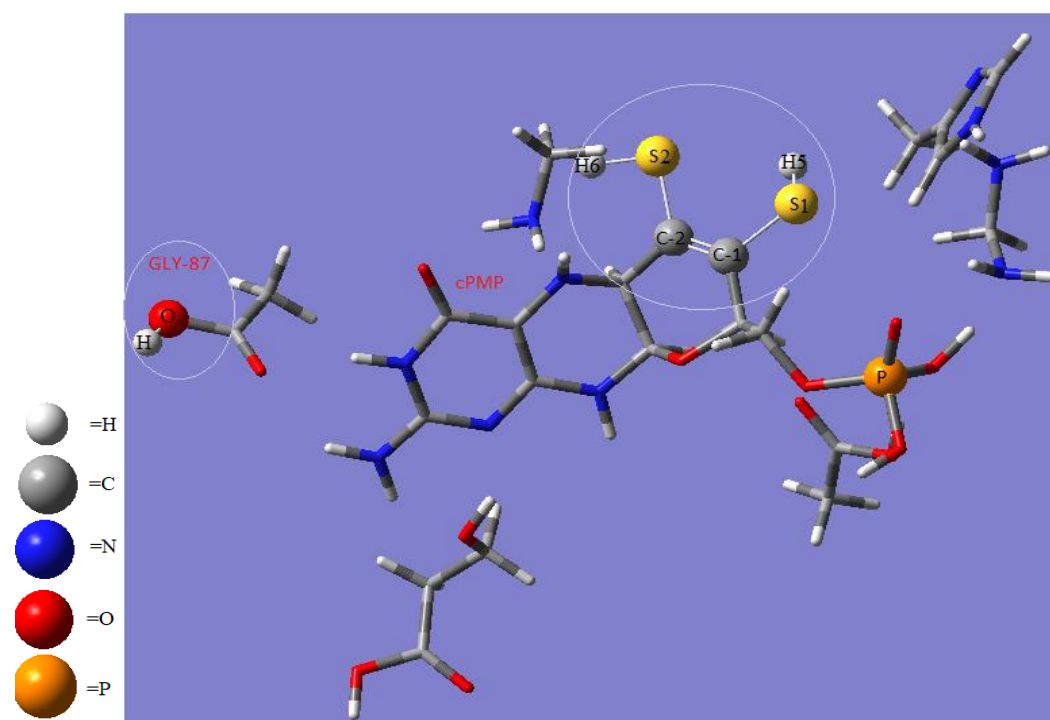


Figure 4.28: Structure of EL<sub>01</sub>-Int-4 geometry with SH groups at C<sub>1</sub> and C<sub>2</sub>

Figure 4.29: Optimized structure of EL<sub>01</sub>-Int-4-SFigure 4.30: Optimized structure EL<sub>01</sub>-Int-4 -L

Tabel 4.4: Relative energies (kcal/mol ) graph in solvent phase of Educt-Ligand Complex (EL<sub>01</sub>L) and Educt-Ligand Complex (EL<sub>01</sub>S)

|                                | (EL <sub>01</sub> L)<br>(Plot 4.1) | (EL <sub>01</sub> S)<br>(Plot 4.2) |
|--------------------------------|------------------------------------|------------------------------------|
| Protein+ Ligand                | 0                                  | 0                                  |
| EL <sub>01</sub>               | -28.7                              | -28                                |
| EL <sub>01</sub> -Int-1        | 11                                 | -0.3                               |
| EL <sub>01</sub> -Int-2        | -26.3                              | -26.3                              |
| EL <sub>01</sub> -Int-3 (-1 1) | 21.7                               | 27.6                               |
| EL <sub>01</sub> -Int-4        | 9.7                                | 13.6                               |

Obtained results of all of these structures are presented separately in Fig 4.31 and Fig 4.32.

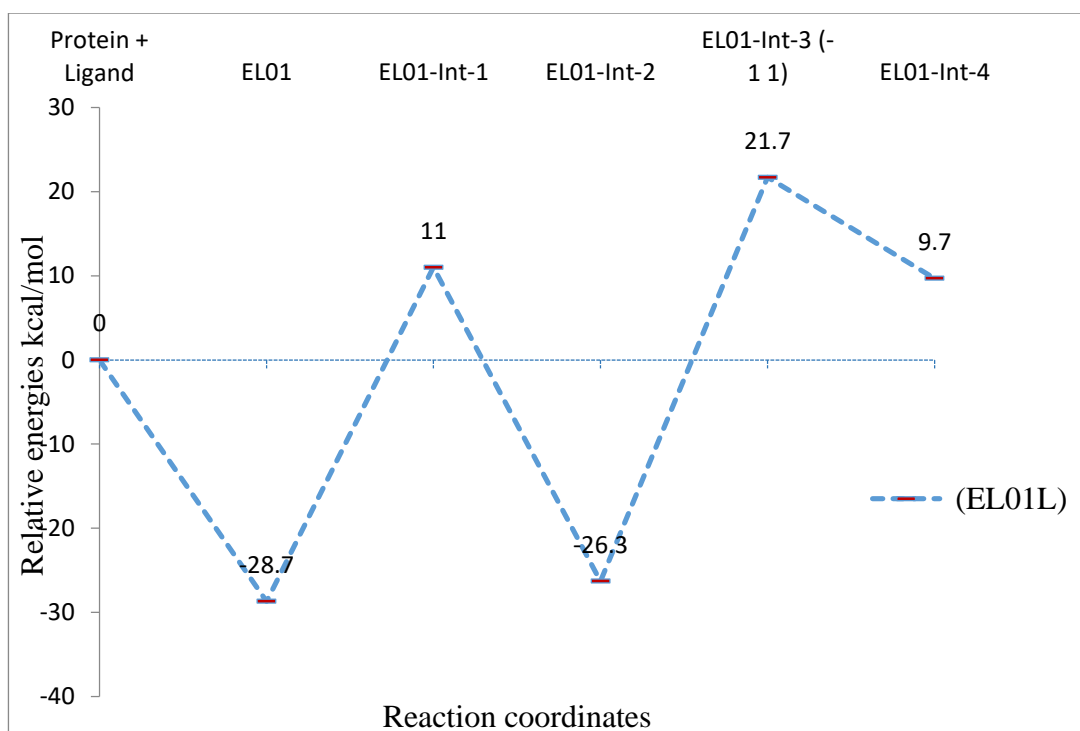


Figure 4.31: Relative energies(kcal/mol) graph in solvent phase of Educt-Ligand Complex ( $EL_{01}L$ )

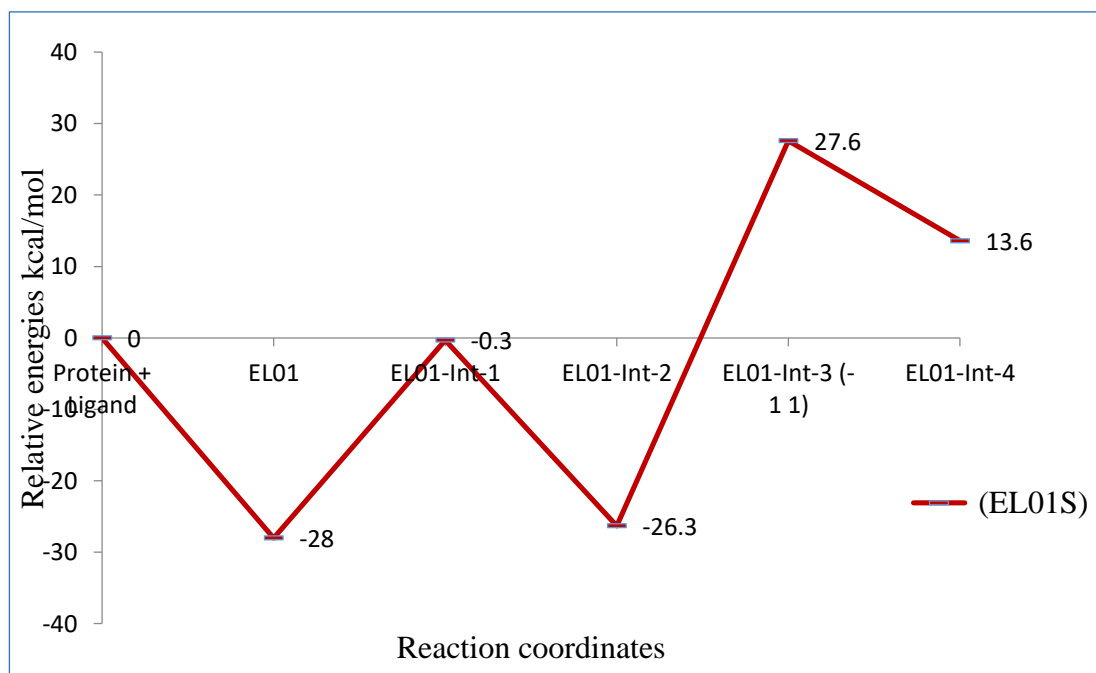


Figure 4.32: Relative energies(kcal/mol) graph in kcal/mol in solvent phase of Educt-Ligand Complex ( $EL_{01}S$ )

#### 4.4.3.2 Educt-Ligand Complex (EL<sub>-11</sub>)

In this Educt-Ligand, complex geometry (EL<sub>-11</sub>) the overall charge on the molecule was kept -1 and multiplicity 1 on the active site of protein and ligand complex. There was no imaginary frequency seen in results obtained from the computational analysis of EL<sub>-11</sub>, which indicates that the geometry is fully optimized and it is at minimum energy on the potential energy surface.

Formation of EL<sub>-11</sub> geometry was an exothermic process with the release of energy 14 kcal/mol with B3LYP/LANL2DZ level of DFT and 18 kcal/mol with B3LYP/SDD level of DFT (Table 4.5). There are two hydroxyl groups O<sub>1</sub>H<sub>1</sub> and O<sub>2</sub>H<sub>2</sub> present at C<sub>2</sub> position (Figure 4.33) with bond distance of about 1.445 and 1.438Å.

Optimized structures of EL<sub>-11</sub>S (Educt-Ligand Complex with B3LYP/SDD level of DFT) and EL<sub>-11</sub>L (Educt-Ligand Complex with B3LYP/LANL2DZ level of DFT) are shown in Figure 4.34 and 4.35 respectively.

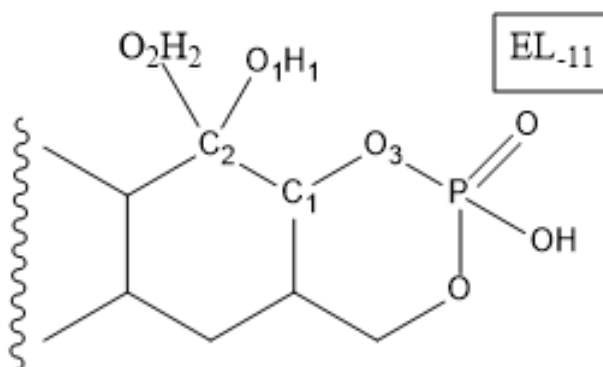
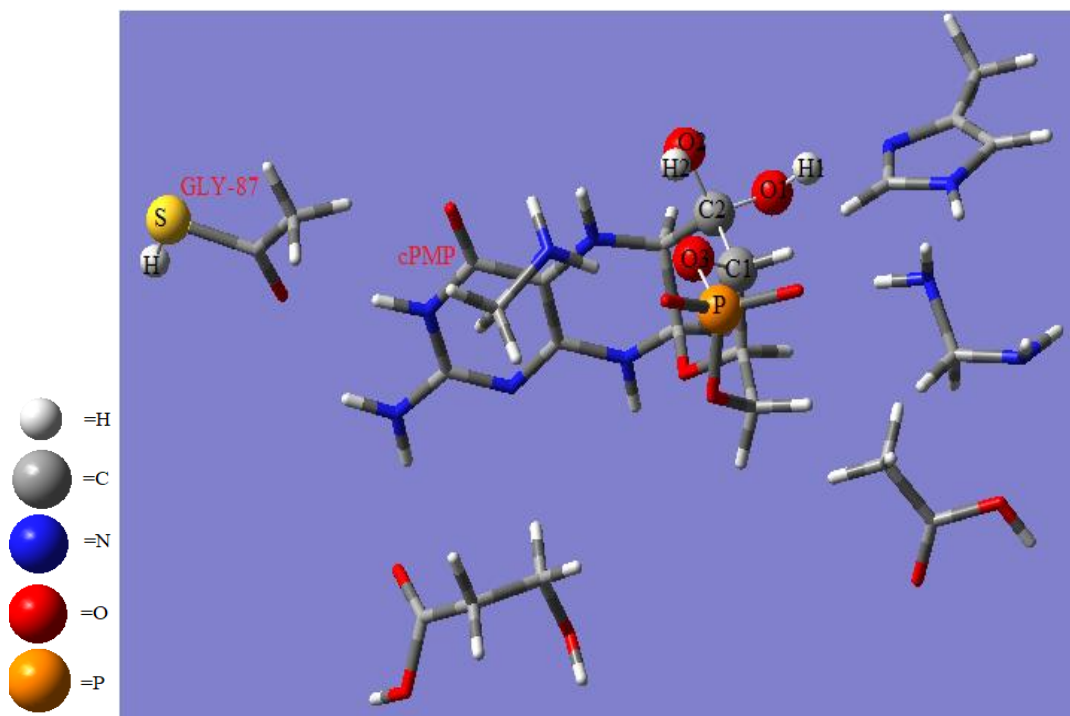
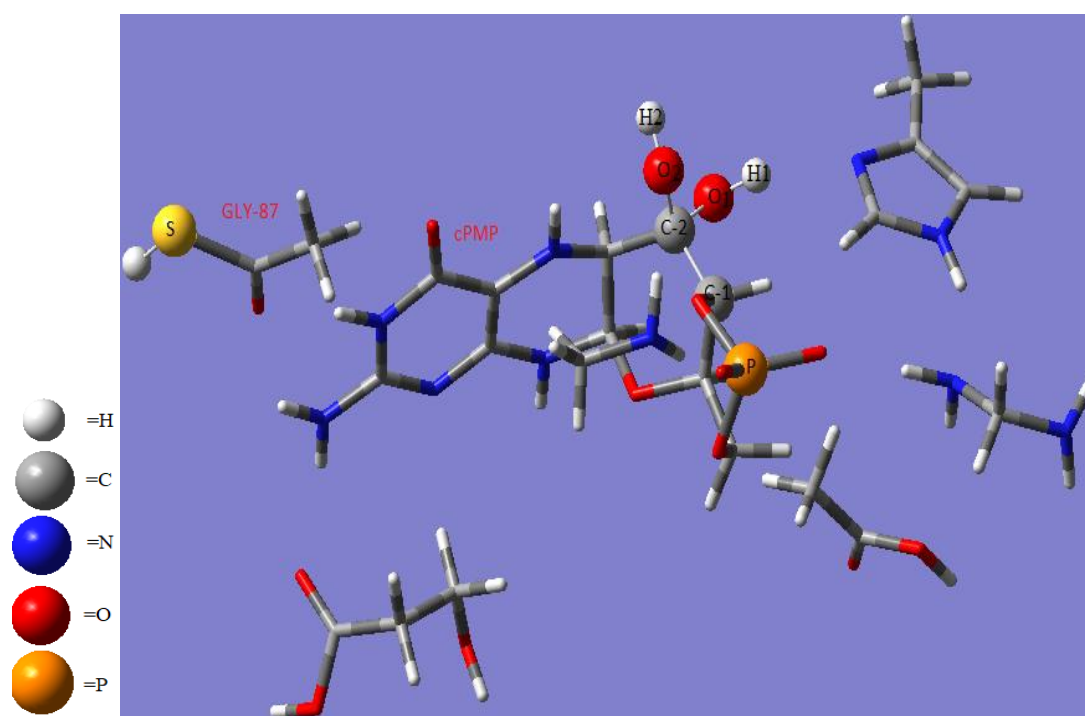


Figure 4.33: Structure of EL<sub>-11</sub> geometry with two hydroxyl groups

Figure 4.34: Optimized geometry of EL<sub>11</sub>-SFigure 4.35: Optimized geometries EL<sub>11</sub>-L



#### 4.4.3.2.1 Intermediate Complex Geometry (EL<sub>11</sub>Int-1)

At next step, dehydration of EL<sub>11</sub> took place at C<sub>2</sub> and gives the first Intermediate geometry, EL<sub>11</sub>-Int-1, where double bond is formed between C<sub>2</sub> and Oxygen (O<sub>2</sub>) upon removal of one O<sub>1</sub>H<sub>1</sub> and H<sub>2</sub> (Figure 4.36). This reaction is exothermic with release of 11.8 kcal/mol energy on potential energy surface with B3LYP/SDD level of DFT and 1.5kcal/mol energy with B3LYP/LANL2DZ level of DFT (Table 4.5 Figure 4.37 & 38).

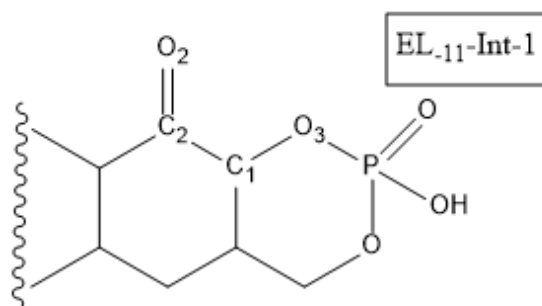


Figure 4.36: Structure of EL<sub>11</sub>-Int-1 geometry with carbon oxygen double bond at C<sub>2</sub>

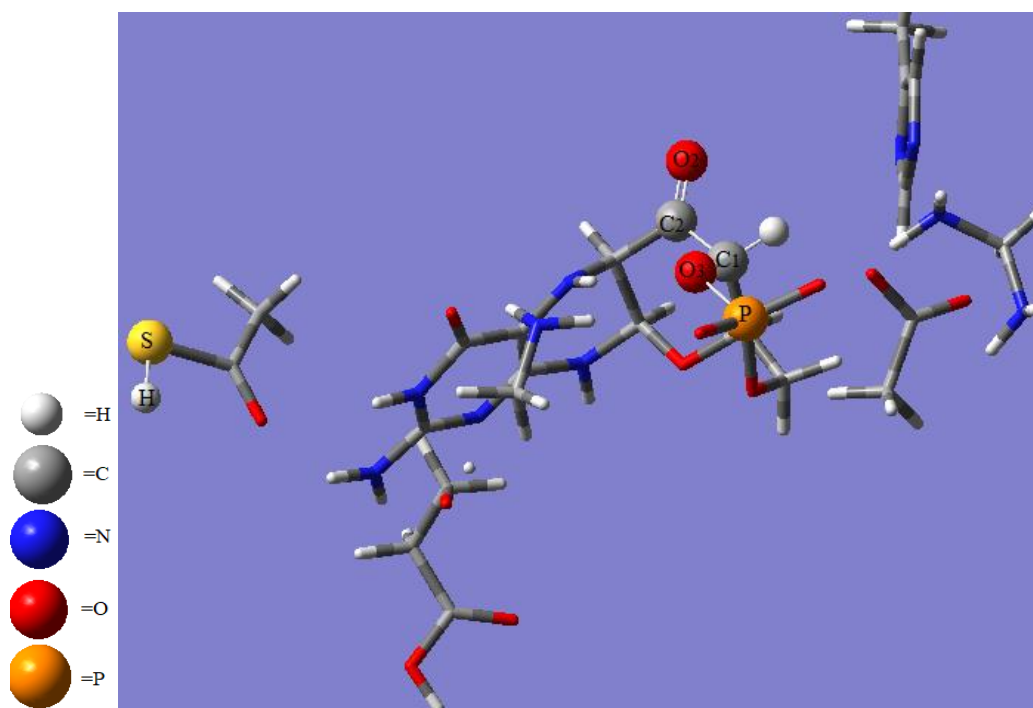


Figure 4.37: Optimized geometry of EL<sub>11</sub>-Int-1-S

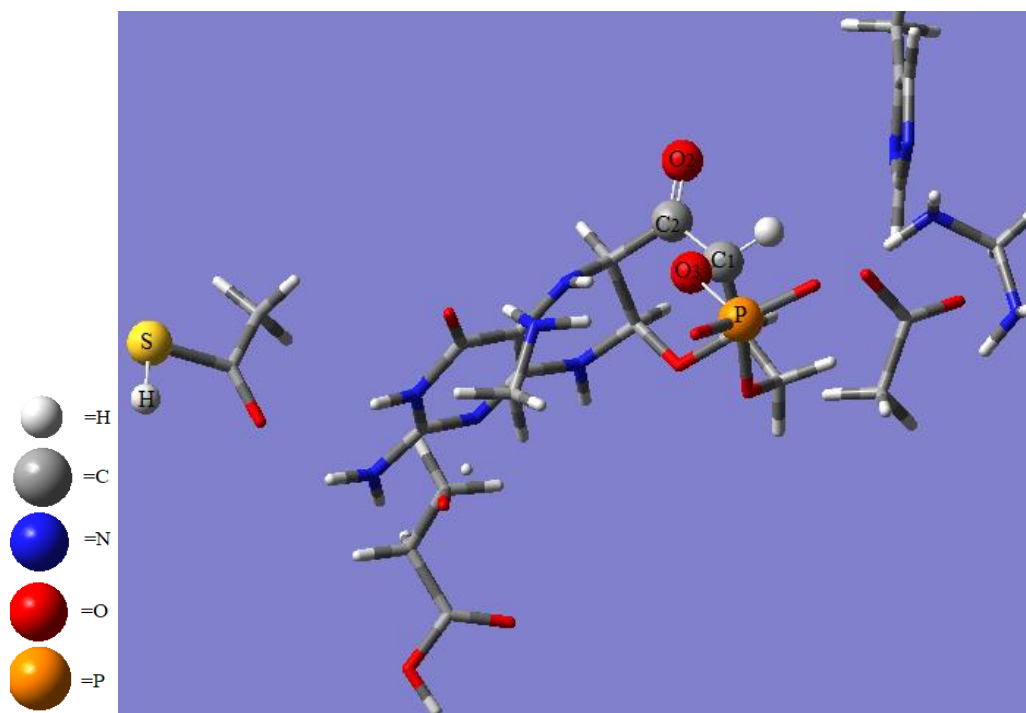


Figure 4.38: Optimized geometry of EL<sub>11</sub>-Int-1-L

#### 4.4.3.2.2 Intermediate Complex Geometry (EL<sub>11</sub>Int-2)

Upon conversion of EL<sub>11</sub>-Int-1 to second intermediate geometry EL<sub>11</sub>-Int-2, there are two changes took place, first is the breakdown of ring, this breakdown occurs at the bond between oxygen (O<sub>3</sub>) and carbon(C<sub>1</sub>). After breakdown of this bond, O<sub>3</sub> bonded with phosphorus at one end become negatively charged. Second change is that sulfide group (S<sub>1</sub>H<sub>4</sub>) is added in this geometry from glycine (GLY-87) of smaller subunit of protein (MOCS2A) to carbon atom present at position one (C<sub>1</sub>) (Figure 4.39). This reaction is exothermic with release of 25.2 kcal/mol energy with B3LYP/LANL2DZ level of DFT and 25.1 kcal/mol energy with B3LYP/SDD level of DFT (Table 4.5, Figure 4.40&4.41).

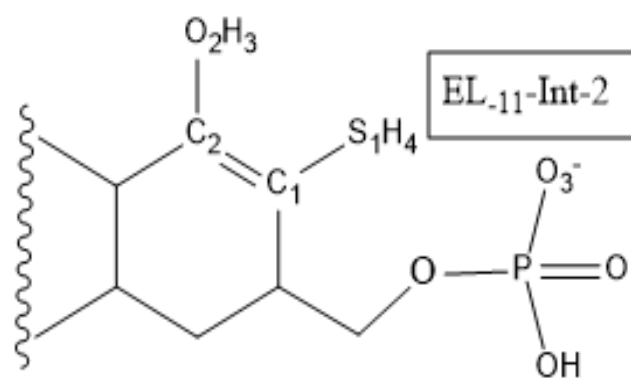


Figure 4.39: Structure of EL<sub>.11</sub>-Int-2 geometry with SH group at C<sub>1</sub>

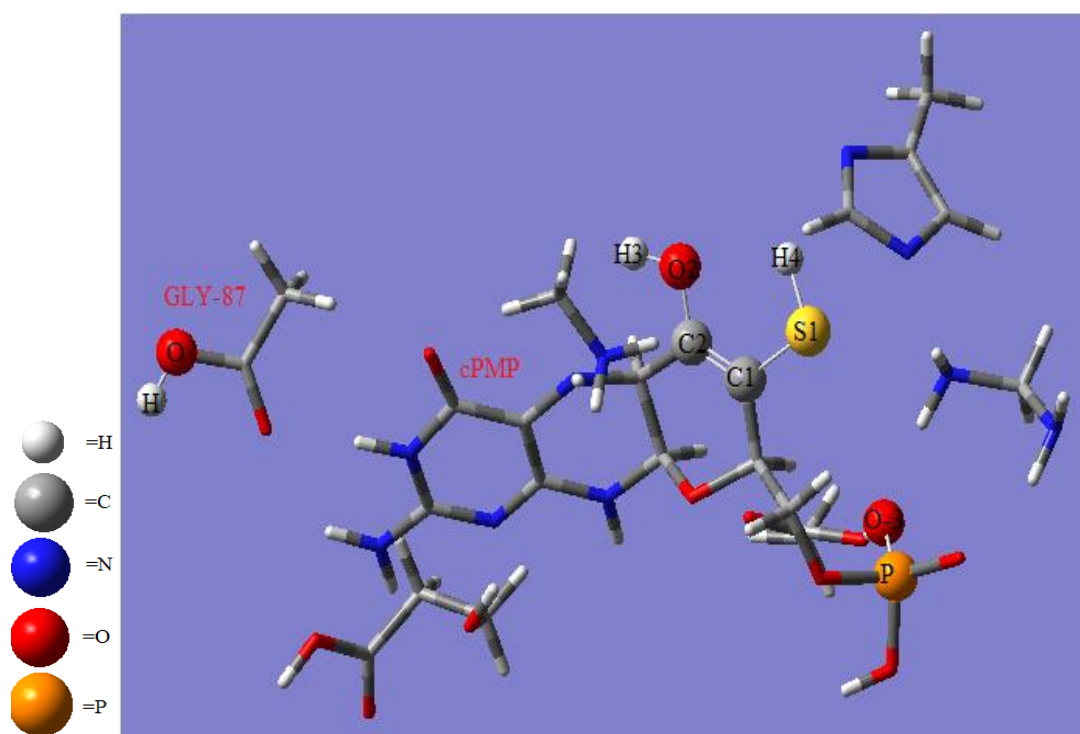


Figure 4.40: Optimized geometry of EL<sub>.11</sub>-Int-2-S

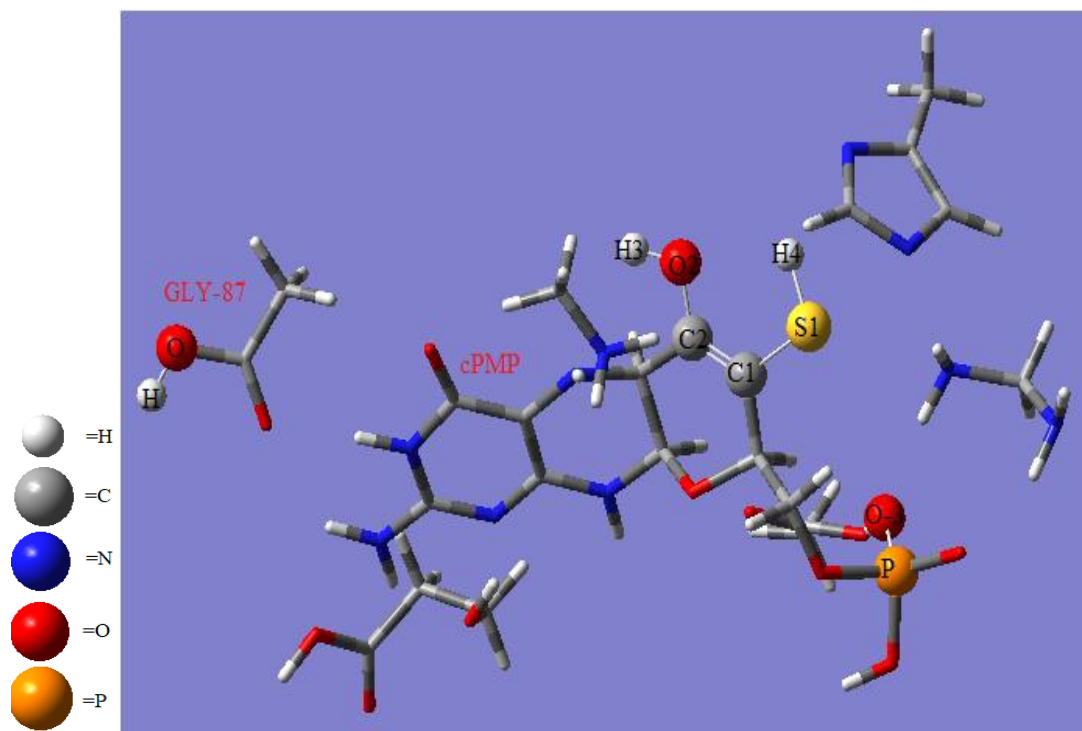


Figure 4.41: Optimized geometry of EL<sub>11</sub>-Int-2-L

#### 4.4.3.2.3 Intermediate Complex Geometry (EL<sub>11</sub>Int-3)

From EL<sub>11</sub>-Int-2 geometry, third intermediate geometry EL<sub>11</sub>-Int-3 is formed. At this step only changes occurred at GLY-87 in Chain A of MPT synthase, by addition of SH group, with the help of MOCS3 (MPT-synthase sulfurase), with the replacement of OH of GLY-87. In this step, no change occurred at the active site of MPT synthase (Figure 4.43-4.44).

This reaction is endothermic with absorption of 21.7 kcal/mol energy with B3LYP/LANL2DZ level of DFT and 27.6 kcal/mol with B3LYP/SDD level of DFT (Table 4.5).

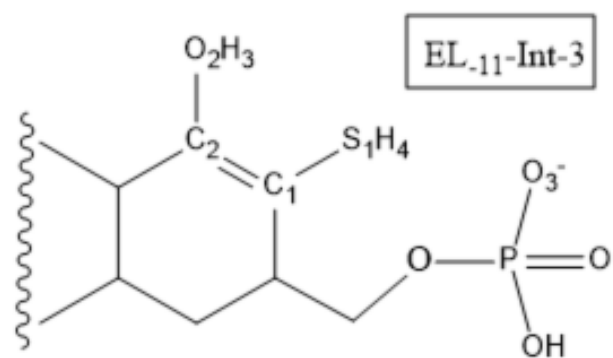


Figure 4.42: Structure of EL<sub>.11</sub>-Int-3 geometry with SH group at C1

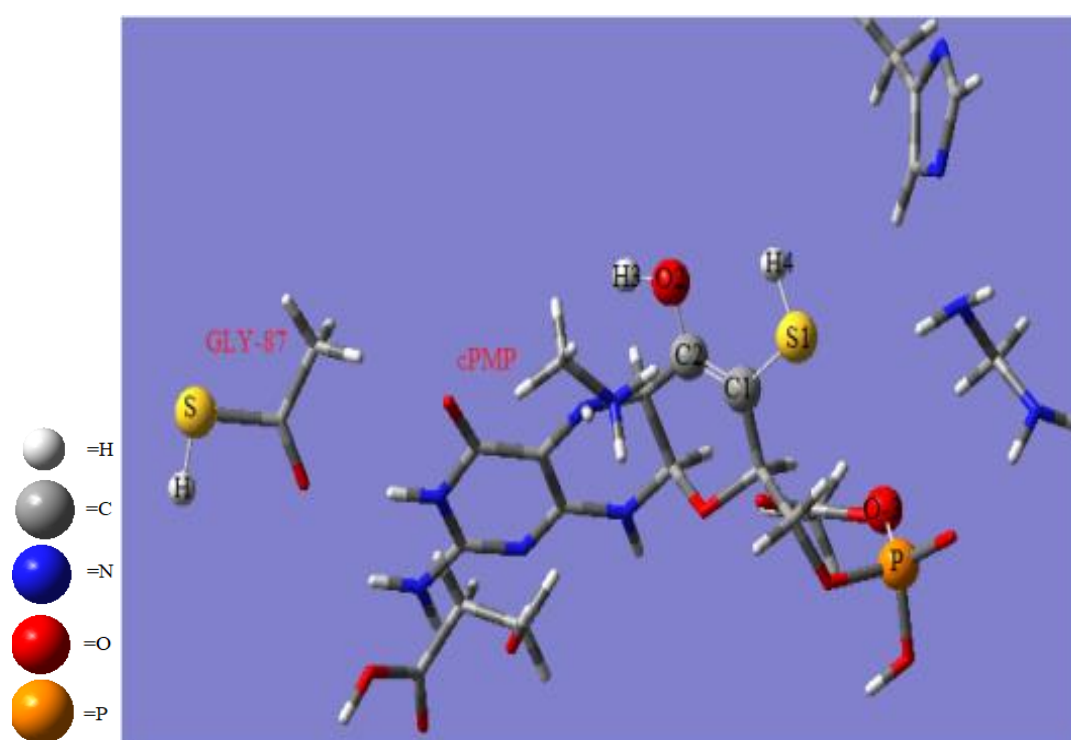


Figure 4.43: Optimized geometry of EL<sub>.11</sub>-Int-3-S

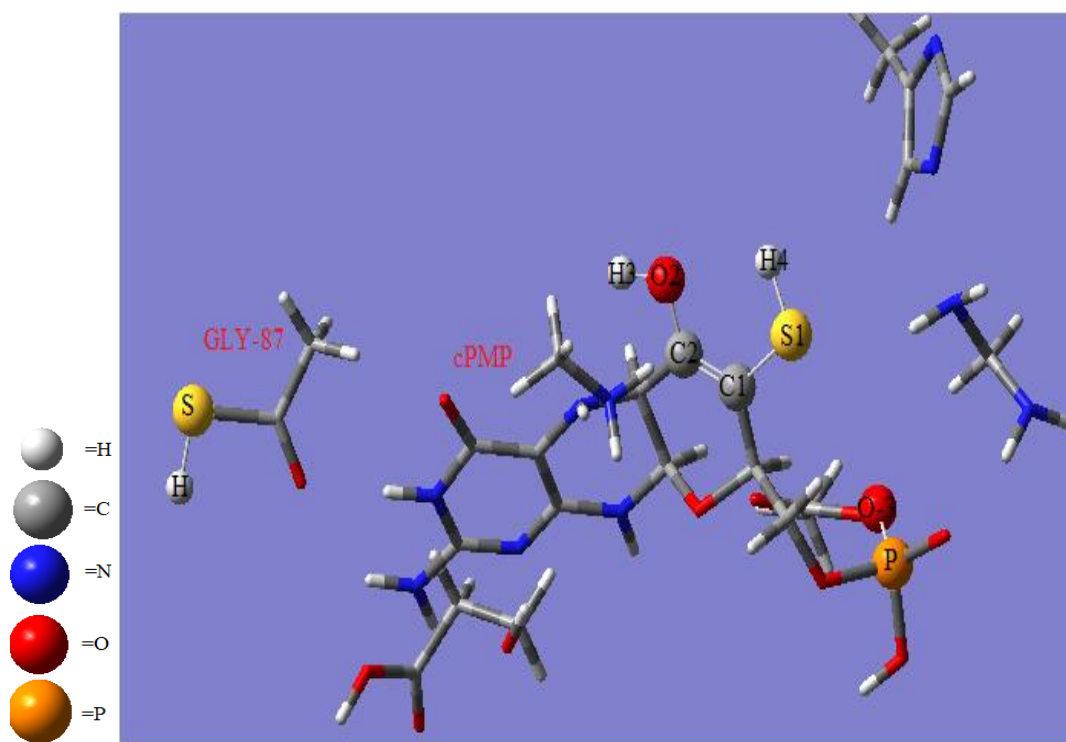


Figure 4.44: Optimized geometry of EL<sub>11</sub>-Int-3-L

#### 4.4.3.2.4 Intermediate Complex Geometry (EL<sub>11</sub>-Int-4)

At next step, EL<sub>11</sub>-Int-3 is converted into Intermediate Complex Geometry EL<sub>11</sub>-Int-4 by addition of second S<sub>2</sub>H<sub>3</sub> group, which was added at place of O<sub>2</sub>H<sub>3</sub> at C<sub>2</sub> of cPMP. Again, this SH is added from GLY-87 of MPT synthase (Figure 4.45). This step is again endothermic with consumption of 17.7 kcal/mol energy with B3LYP/LANAL2DZ level of DFT and 18.6 kcal/mol with B3LYP/SDD level of DFT relative to the separate protein and ligand (Table 4.5, Figure 4.46&4.47).

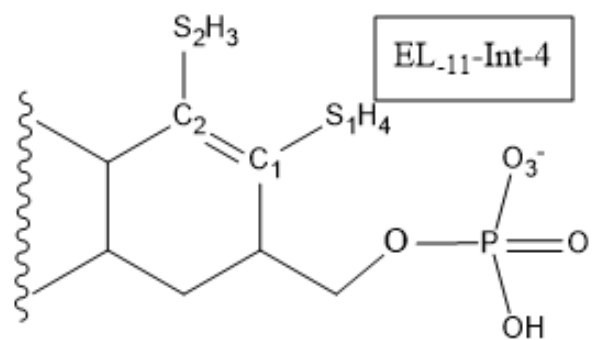


Figure 4.4513: Structure of EL<sub>11</sub>-Int-4 geometry with SH group at C<sub>1</sub> and C<sub>2</sub>

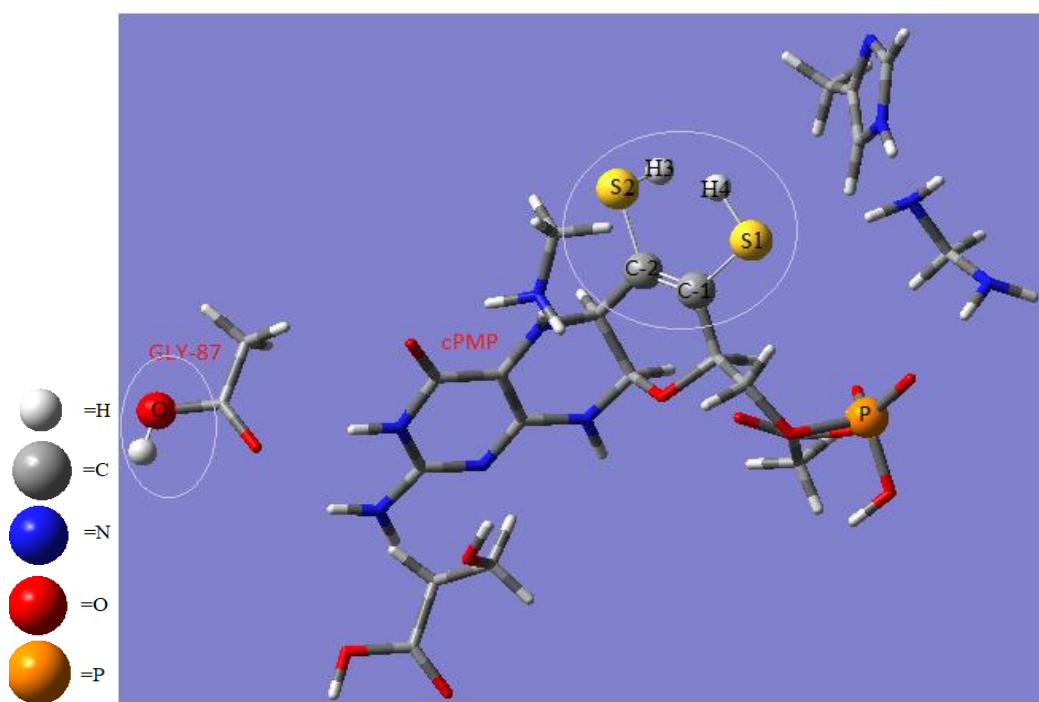
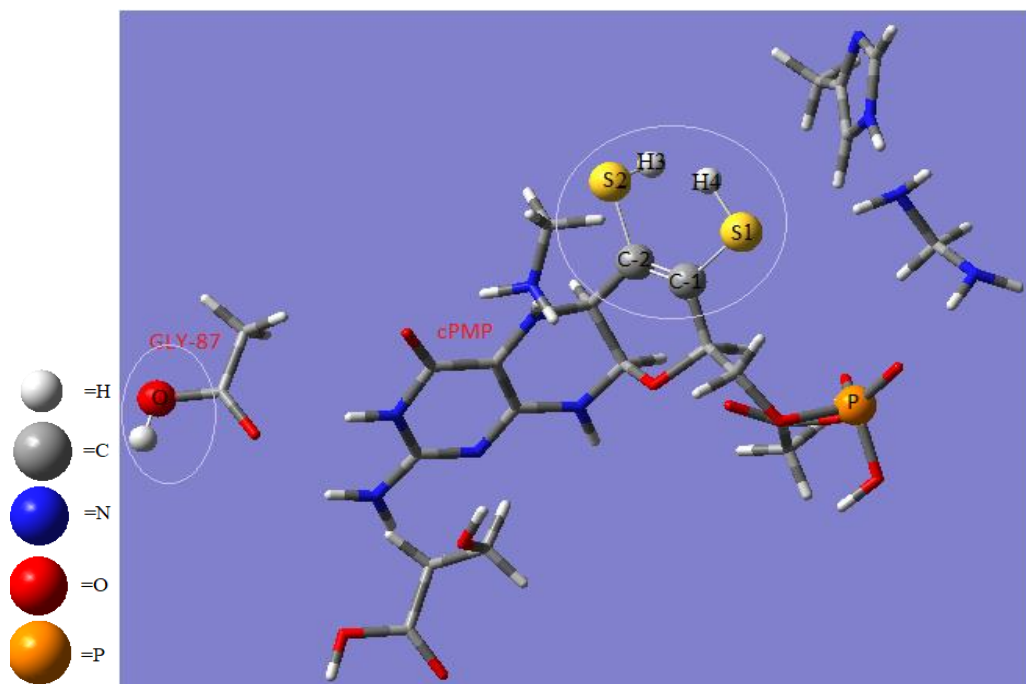


Figure 4.46: Optimized geometry of EL<sub>11</sub>-Int-4-S

Figure 4.47: Optimized geometry of EL<sub>11</sub>-Int-4-L

Tabel 4.5: Relative energies in kcal/mol of optimized geometries in solvent phase of  
 Educt-Ligand Complex (EL<sub>11</sub>L) and Educt-Ligand Complex (EL<sub>11</sub>S)

|                         | EL <sub>11</sub> L<br>(Plot 4.3) | EL <sub>11</sub> S<br>(Plot 4.4) |
|-------------------------|----------------------------------|----------------------------------|
| Protein+Ligand          | 0                                | 0                                |
| EL <sub>11</sub>        | -14                              | -18                              |
| EL <sub>11</sub> -Int-1 | -1.5                             | -11.8                            |
| EL <sub>11</sub> -Int-2 | -25.2                            | -25.1                            |
| EL <sub>11</sub> -Int-3 | 21.7                             | 27.6                             |
| EL <sub>11</sub> -Int-4 | 17.7                             | 18.6                             |
| P <sub>1,2</sub> -5     | -0.01                            | -0.02                            |



Moreover, all these obtained results are also presented separately in Figure 4.48 and Figure 4.49.

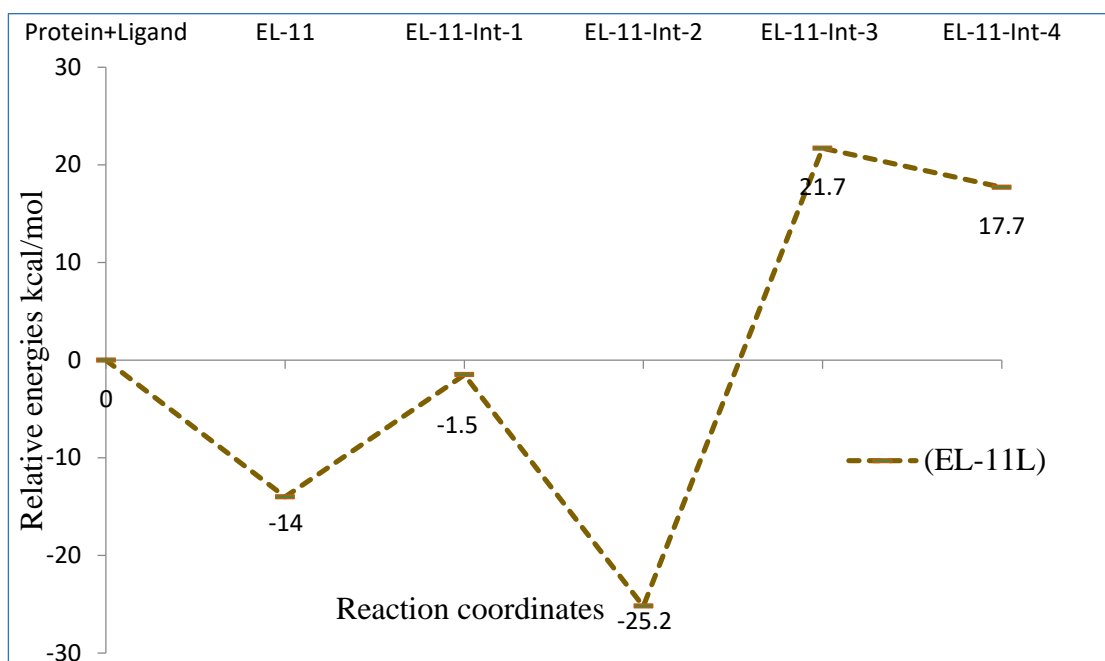


Figure 4.48: Relative energies (kcal/mol) graph in solvent phase of Educt-Ligand Complex (EL<sub>11</sub>L)

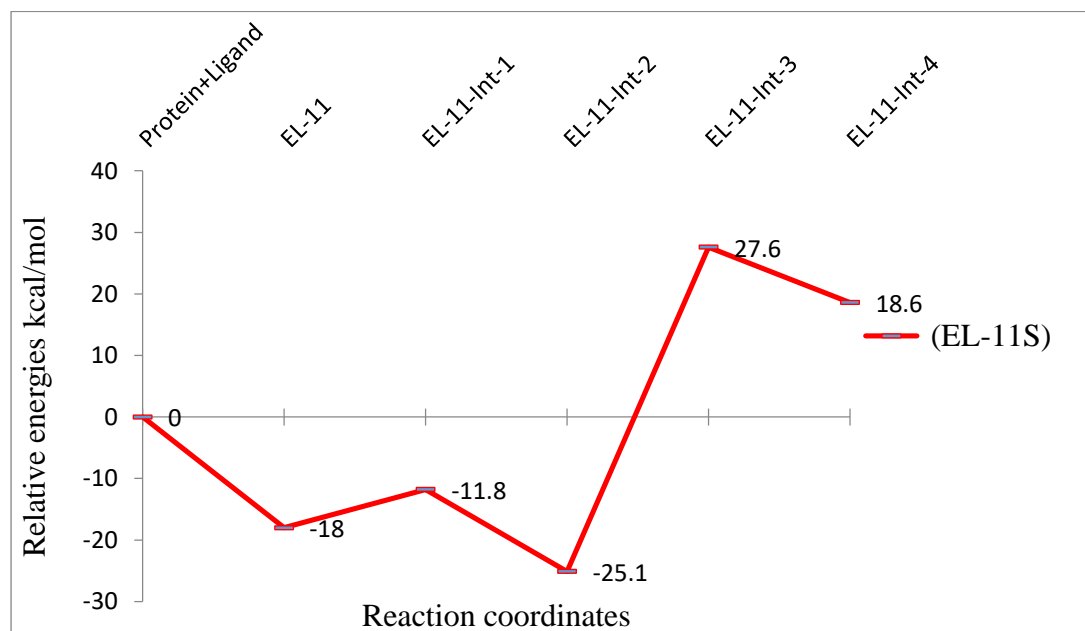


Figure 4.49: Relative energies (kcal/mol) graph in solvent phase of Educt-Ligand Complex (EL<sub>11</sub>S)

### 4.5 Role of copper

According to literature, copper is attached at the Metal Binding pterin (MPT), however, its role is not clear. Therefore, to understand the role of copper, product-complex P-5 was generated by the replacement of a Cu ion with the two hydrogen atoms in the Intermediate complex EL<sub>-11</sub>-Int-4. (Figure 4.50).

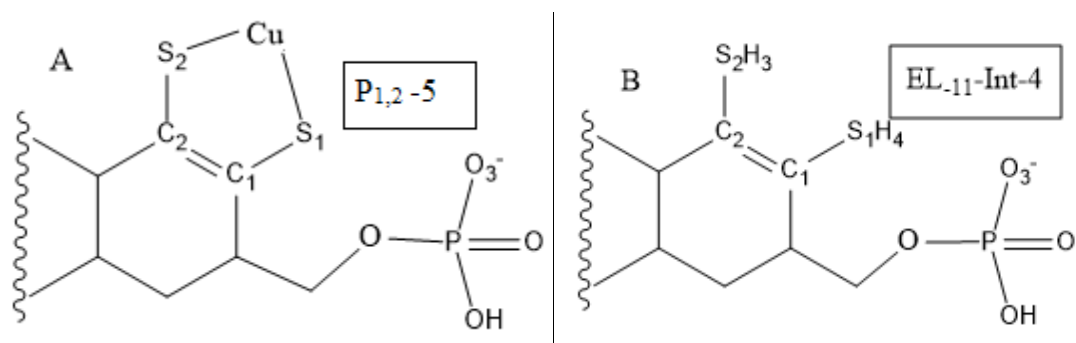
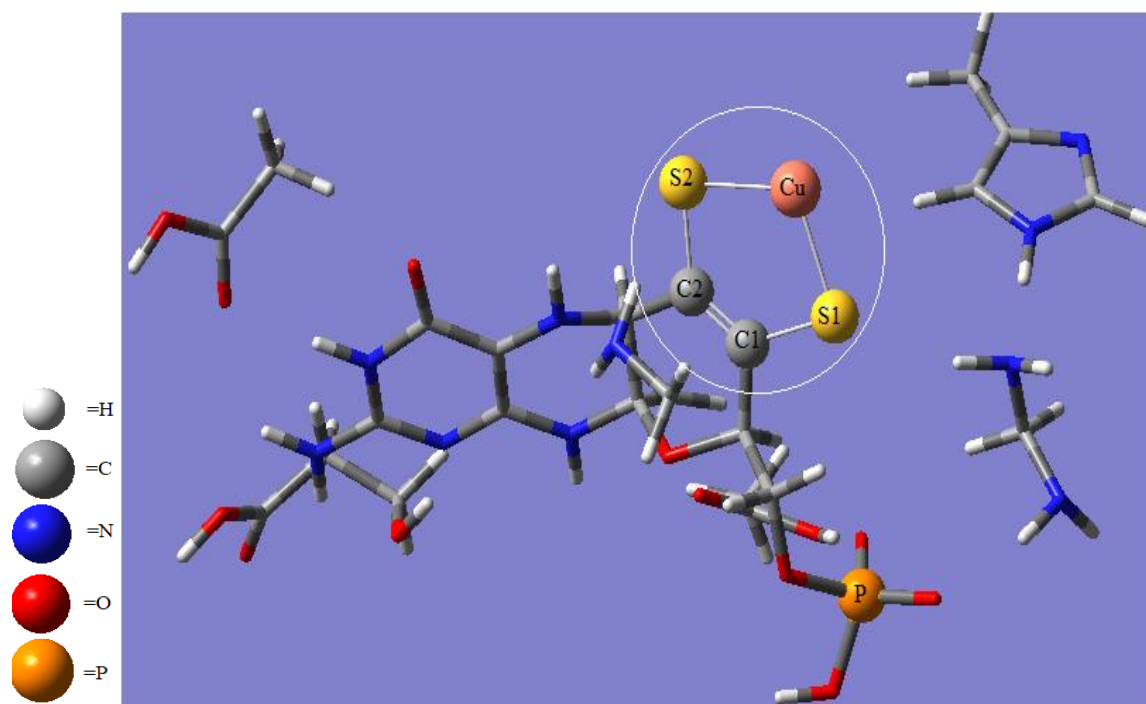
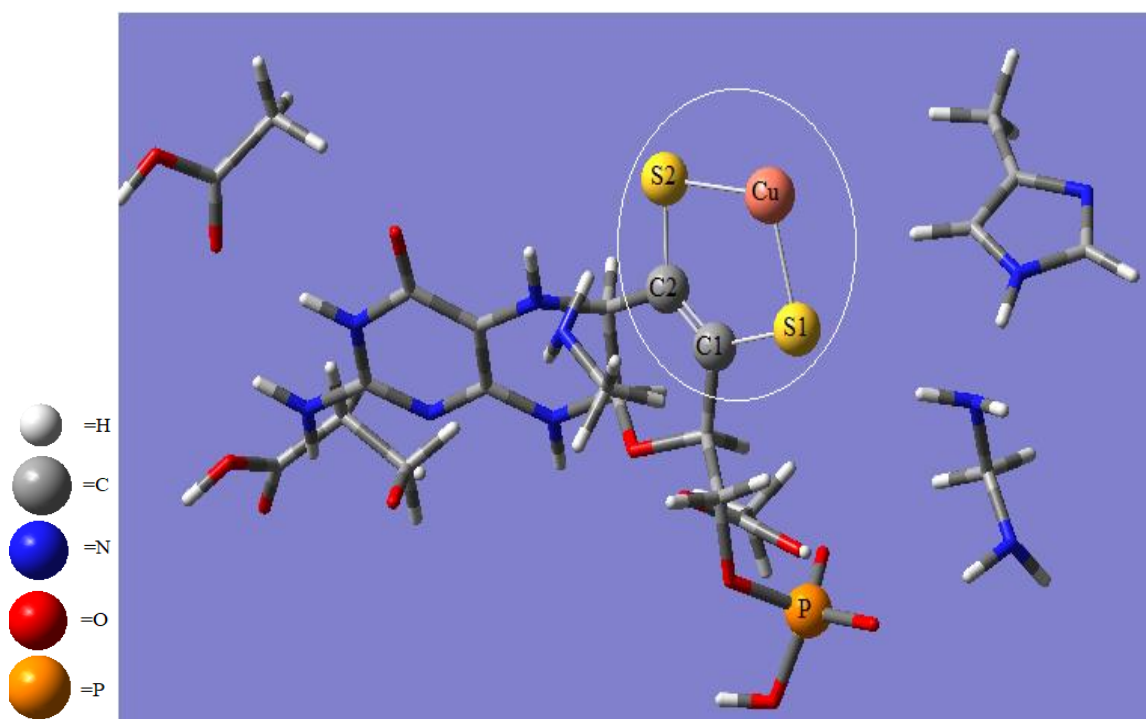


Figure 4.50: Structure of P<sub>1,2</sub>-5 geometry in which Cu attached with S<sub>1</sub> and S<sub>2</sub> (A) and structure without copper (B)

Because of replacement of two hydrogen atoms with a copper ion, charge of molecule was changed to 1 and multiplicity 2. Computational results indicate that this reaction is neither exothermic nor endothermic as the relative energy on the potential energy surface is almost zero (-0.012kcal/mol energy) with B3LYP/SDD level of DFT as well as with B3LYP/LANL2DZ level of DFT (Table 4.5 Figure 4.51 & 4.52). Therefore, we can say that copper plays no role in the formation of MPT but it can play major role in the addition of molybdate which is the third step in the biosynthesis of MoCo.

Figure 4.51: Optimized structure of P<sub>1,2</sub>-5-SFigure 4.52: Optimized structure P<sub>1,2</sub>-5-L

# CHAPTER 5

## DISCUSSION

Molybdenum along with its cofactor (molybdenum cofactor) is present at the active site of all molybdenum-enzymes. According to literature, importance of Moco was studied first time in humans in the year 1967 when the deficiency of sulfite oxidase was observed in an infant. It has been observed that it is a rare unattended disorder which results in severe neurological abnormalities, in some cases caused death in early childhood.<sup>16</sup>

Importance of molybdenum cofactor was studied and it was identified that the molybdenum cofactor is very important for the proper functioning of many enzymes in the human body as well as other eukaryotes and prokaryotes.<sup>17</sup> To find the reason behind the dysfunctioning of molybdenum dependent enzymes or deficiency of molybdenum cofactor in the human body there was a need to understand the biosynthesis of normal molybdenum cofactor.

Molybdenum cofactor biosynthesis is a 4-step process in which almost 06 proteins are involved. The four steps are (1) formation of cPMP from GTP with the help of MOCS1 A and B enzyme, (2) MPT formation from cPMP with the help of MPT synthase (MOCS2 A and B) and MPT synthase sulfurase (MOCS3) (3) Adenylation of MPT with the help of protein GEPHYRIN-G and last is (4) Insertion of Molybdenum into MPT with the help of GEPHYRIN-E to form molybdenum cofactor<sup>12,13</sup>. In this dissertation, our research mainly

focuses on the reaction mechanism of its second step that is MPT formation with the help of MPT synthase (MOCS1A & MOCS1B) from cPMP using Computational methods.

To study its reaction mechanism, the knowledge of active site was very important. literature review revealed that although different binding pockets were reported but no binding pocket was reported for cPMP. Therefore, the first step was to identify the binding pocket/ active site for cPMP which was found by using binding pocket finder softwares (DoGsite Scorer and RaptorX). Results indicated that the binding pocket/ active site for cPMP contains GLY-87, GLY-69, ALA-70, HIS-143, ARG-144, GLY-76, THR-77, THR-78, ARG-79, ASN-81, PHE-82, GLU-168 residues from all four chains of amino acid residues (Table 4.1 & 4.2) (Figure 5.1 & 5.2).

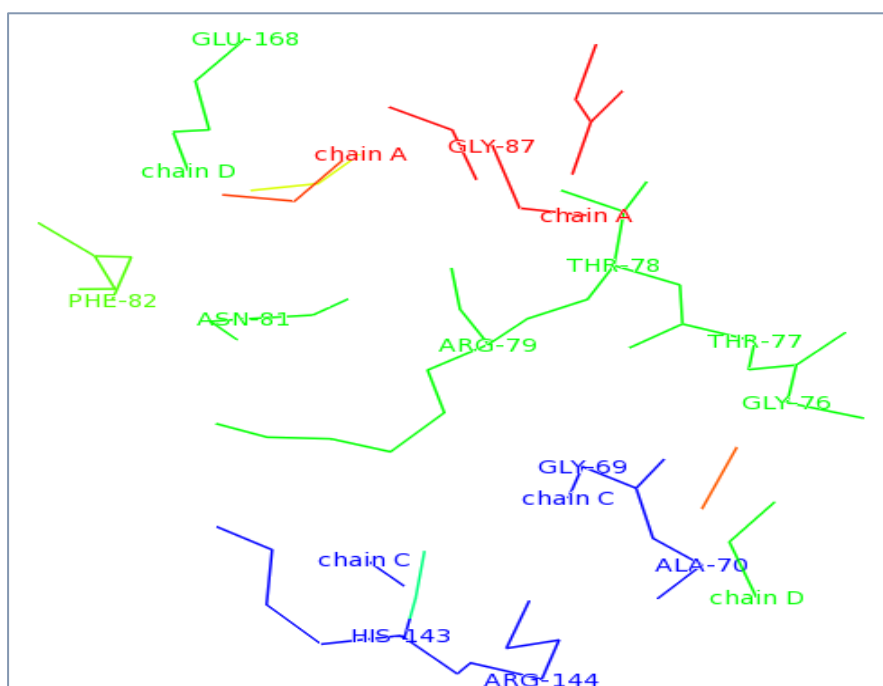


Figure 5.1: First binding pocket (P\_0), generated by DoGsite scorer

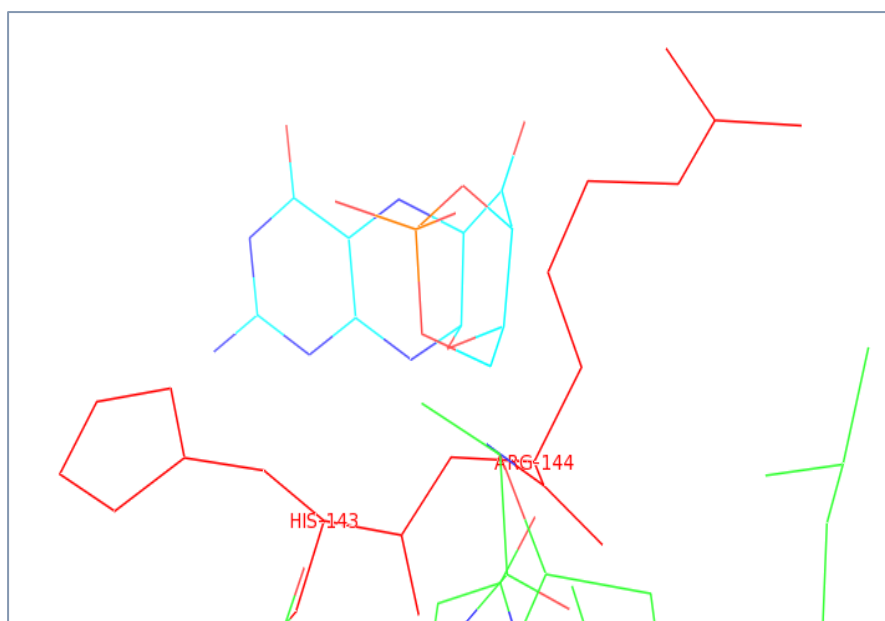


Figure 5.2: First Binding pocket (1) generated from RaptorX

This binding pocket/ amino acid residues/ active site was then selected for the molecular docking. Molecular Docking results showed 09 pockets from which Protein\_Mligand-1 with residues of Chain-A (Gly-87), Chain-C (Gly-69, His-143) and of Chain-D (77THR, ARG-79, SER-134, LYS-166) was selected on the basis of its high binding affinity value (-9.4) (Table 4.3).

Next step was to understand the reaction mechanism of MPT formation from cPMP using quantum mechanical calculations. Proteins/enzymes are the macromolecules containing thousands of atoms and it is not possible to do quantum mechanical studies on the whole protein/ enzyme molecule due to time constraints as well as computational cost. Therefore, need was to identify the amino acid residues which may directly be involved in the conversion of cPMP to MPT. Accordingly, active site region of 8Å from protein (cPMP docked MPT Synthase) was extracted using Swiss-Pdb Viewer which included Gly-87, Gly-69, Ser-134, HIS-143 Arg-79 and Lys-166-amino acid residues of MPT Synthase along with cPMP ligand.

To generate model complex geometries for computational studies, the selected residues were truncated using GUI Gauss view, hydrogen atoms were added wherever required to manage the molecular charge. To keep the effect of protein moiety, atoms from where the chains were truncated, were kept fixed for whole quantum mechanical studies.

For quantum mechanical studies, DFT method B3LYP was used along with basis sets, LANL2DZ<sup>47</sup> and SDD<sup>47</sup>. All geometry optimizations and single point energy calculation in gas phase on the optimized geometries were performed using B3LYP/SDD//B3LYP/LANL2DZ and B3LYP/SDD//B3LYP/SDD levels of DFT followed by frequency calculations and Self-Consistent Reaction Field (SCRF) calculations.

First, hydrogen atoms were optimized applying overall charges (0, assuming phosphate of cPMP is neutral (EL<sub>01</sub>) and one negative, assuming phosphate of cPMP is not neutral (EL<sub>-11</sub>)) and multiplicity 1, keeping all heavy-atoms fixed at their positions. The resulting geometries were then utilized as starting geometries for the generation of input geometries for structures involved in the mechanism for formation of MPT from cPMP (Figure 4.14-A & 4.14-B).

Based on the nature of charge on the cPMP, two mechanisms were proposed each containing almost 05-06 main structures after the formation of educt-substrate complex (Figure 4.13). Computational results indicate that the formation of educt-substrate complex i.e., EL<sub>01</sub> is same with the both basis sets, an exothermic reaction releasing ~28kcal/mol energy relative to separate educt and ligand (Plot 5.1). However, the dehydration step i.e., the first intermediate formation shows an endothermic reaction with LANL2DZ basis set and almost thermoneutral with SDD basis set. Next step, the addition of water shows again similar energies i.e., ~-26kcal/mol relative to separate educt and substrate, an exothermic reaction with both LANL2DZ & SDD basis sets. The replacement of OH with SH at Gly-87 is an endothermic reaction with the consumption of ~21.7 kcal/mol energy using LANL2DZ basis set and ~27.6 kcal/mol energy using SDD, relative to the separate educt and substrate.

Final step is the addition of another HS group for the formation of enedithiolene group of MPT, this is also an endothermic reaction, with absorption of  $\sim 9.7$  kcal/mol energy using LANL2DZ basis set and  $\sim 13.6$  kcal/mol energy using SDD basis set relative to separate educt and substrate (Figure 5.3).

Quantum chemical calculation results of the reaction mechanism for MPT synthase, where cPMP was considered as an anion, indicate that the formation of educt-substrate complex i.e.,  $EL_{-11}$  is an exothermic reaction,  $\sim -14$  kcal/mol energy with LANL2DZ basis set &  $\sim -18$  kcal/mol energy with SDD basis set, relative to separate educt and ligand (Plot 5.2). Formation of first intermediate i.e.,  $EL_{-11}$ -Int-1 also shows an exothermic reaction ( $\sim -1.5$  kcal/mol energy with LANL2DZ basis set &  $\sim -11$  kcal/mol energy with SDD basis set) relative to separate educt and substrate. The addition of water is also an exothermic reaction with the release of  $\sim 25$  kcal/mol energy relative to separate educt and substrate with the both LANL2DZ & SDD basis sets. The replacement of OH with SH at Gly-87 is an endothermic reaction, with absorption energy of  $\sim 21.7$  kcal/mol using LANL2DZ basis set and  $\sim 27.6$  kcal/mol energy using SDD, relative to the separate educt and substrate. Final step, the formation of enedithiolene group of MPT is again an endothermic reaction, with energy consumption of  $\sim 17.7$  kcal/mol using LANL2DZ basis set and  $\sim 18.6$  kcal/mol energy using SDD basis set relative to separate educt and substrate (Figure 5.4).

After analyzing and comparing the basis sets used for the computational analysis on the reaction mechanism of MPT synthase starting from  $EL_{01}$  and it has been observed that LANL2DZ gives more feasible results, low energy pathway on the potential energy surface as compared to SDD basis set.

On comparing the computational results obtained using B3LYP/LANL2DZ level of DFT for the both reaction mechanisms  $EL_{01}$  &  $EL_{-11}$ , it has been observed that although the



formation of Educt substrate complex, EL<sub>01</sub>, is relatively more exothermic ( $\sim -28.7$  kcal/mol) as compared to EL<sub>-11</sub> ( $\sim -14$  kcal/mol), the formation of first intermediate that is the dehydration reaction is difficult (high energy barrier of  $\sim 38$  kcal/mol) in case of reaction mechanism starting from EL<sub>01</sub> (Plot 5.3). On the other hand, first intermediate formation in reaction mechanism starting from EL<sub>-11</sub> requires only  $\sim 12.5$  kcal/mol of energy. Intermediate-2 formation (EL-Int-2) is endothermic in both reaction mechanisms where EL<sub>-11</sub> required less amount of energy ( $\sim 24$  kcal/mol) for the formation of intermediate-2 as compared to EL<sub>01</sub> ( $\sim 37$  kcal/mol). At next step almost equal amount of energy is required for the formation of intermediate-3 formation (EL-Int-3). Final step (formation of intermediate-4) is also endothermic and EL<sub>-11</sub> required less amount of energy to achieve this step as compared to EL<sub>01</sub> (Figure 5.5).

Therefore, the reaction mechanism starting from EL<sub>-11</sub> is the more preferable reaction mechanism for the formation of MPT from cPMP or we can also say that cPMP in its anionic form (-1 charge) is more active as compared to its neutral form in the formation of MPT from cPMP.

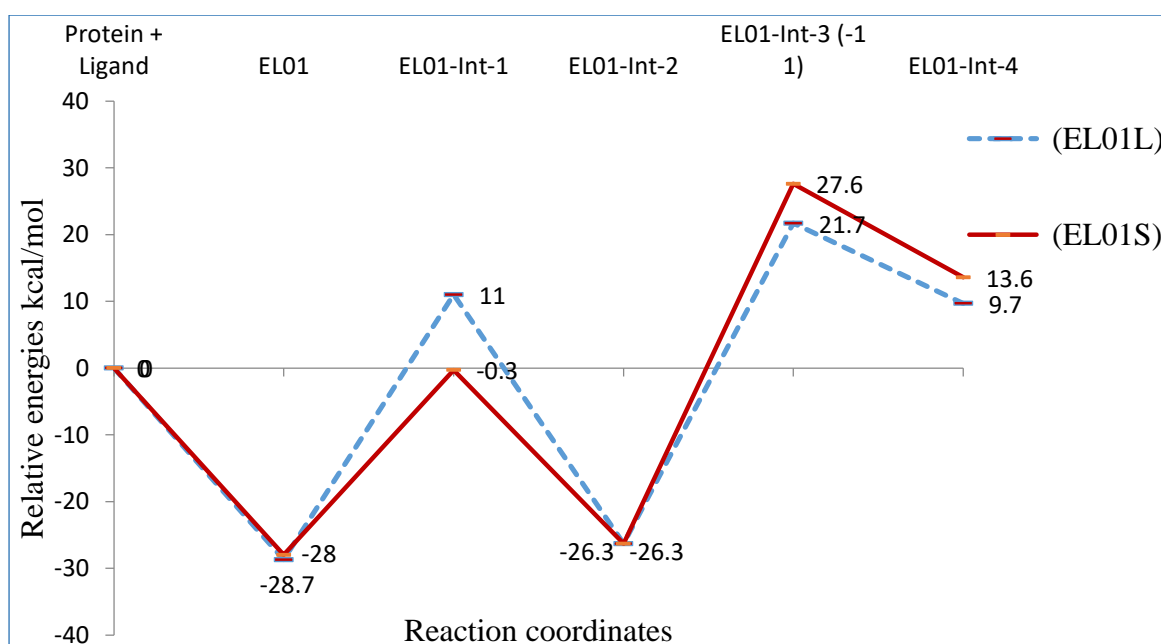


Figure 5.3: Relative energies (kcal/mol) graph in solvent phase of EL<sub>01</sub>L and EL<sub>01</sub>S

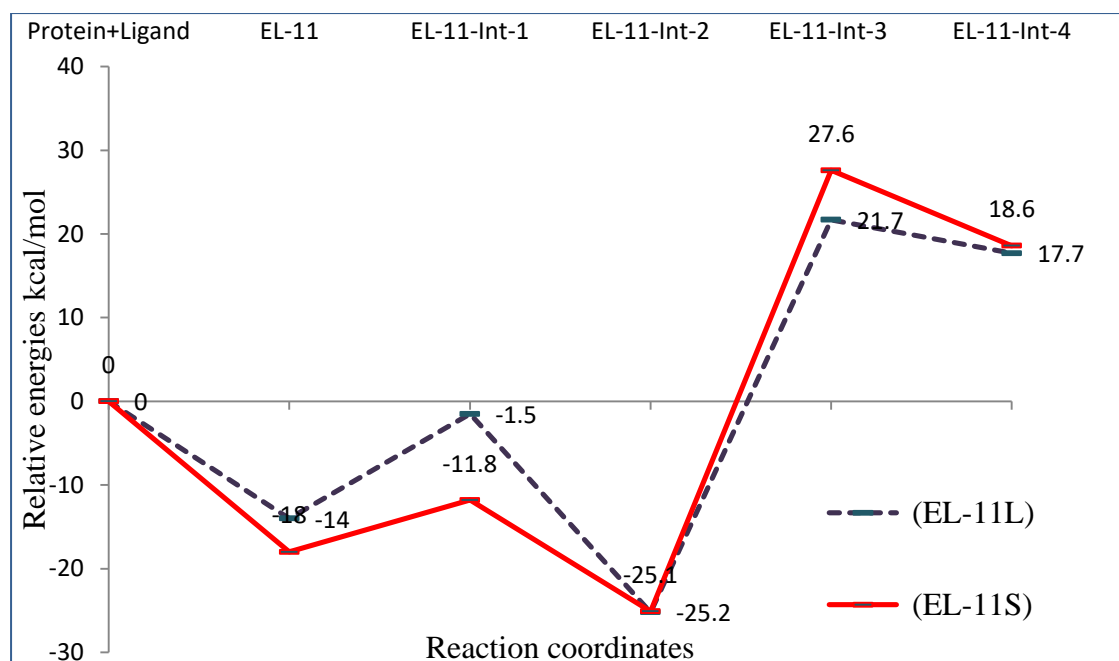


Figure 5.4: Relative energies(kcal/mol) graph in solvent phase (of EL<sub>-11</sub> L) and (EL<sub>-11</sub>S)

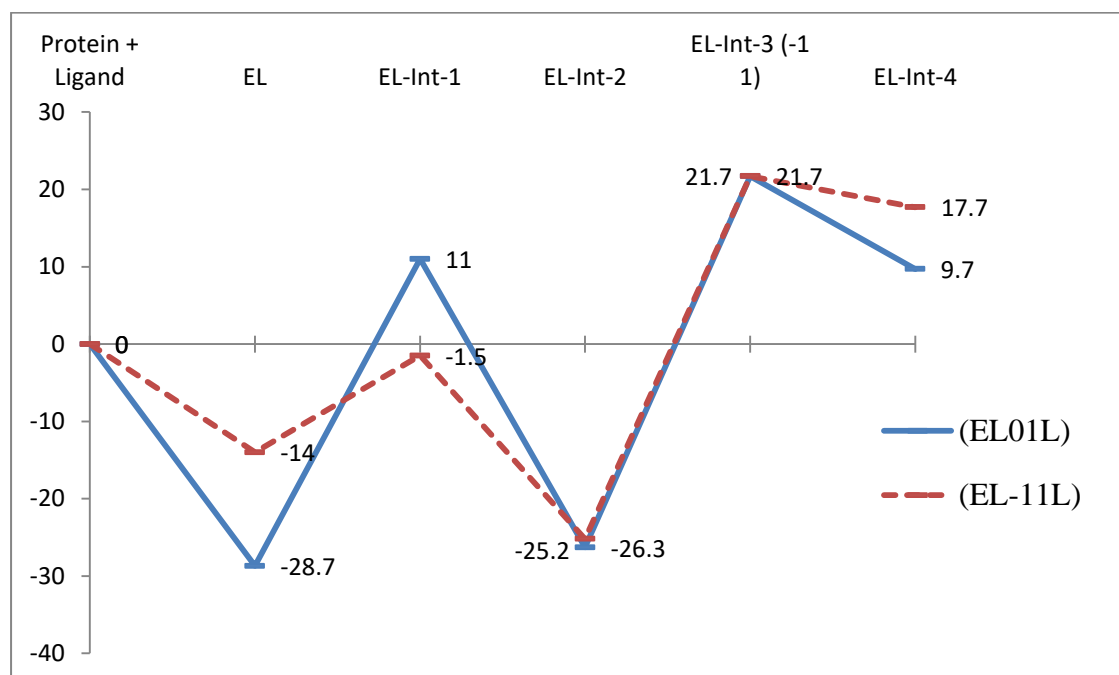


Figure 5.5: Relative energies (kcal/mol) graph of Comparison of LANL2DZ

Final objective of this research was to understand the role of copper in the biosynthesis of molybdenum cofactor. Therefore, optimized intermediate-4 geometry (EL<sub>-11</sub>-Int-4) was modified to study the effect of copper. Two hydrogen atoms attached with two sulfurs at the enedithiolene part of MPT were replaced with a copper ion (Figure 4.50). Computational

analysis was carried out using B3LYP/LANL2DZ & B3LYP/SDD level of DFT. Computational results obtained showed negligible changes when Cu replaces two hydrogen atoms attached with the sulfur atoms, i.e with B3LYP/LANAL2D it required -0.01kcal/mol & with B3LYP/SDD it required -0.02 kcal/mol energy values on PES relative to optimized EL<sub>11</sub>-INT-4. Therefore, it has been revealed that there is no specific role that copper plays, rather only to provide place for the attachment of molybdenum in later step of MoCo biosynthesis.

# CHAPTER 6

## CONCLUSION

Computational analysis was done on the second step of molybdenum cofactor biosynthesis to find the best model complex geometries as well as the reaction mechanism involved in the formation of MPT from cPMP with the help of enzyme MPT synthase in humans. For this purpose, PDB structure was taken with PDB ID 5MPO<sup>27</sup> from X-ray crystal structure data of MPT synthase and active site was predicted for cPMP as no information on binding pocket for cPMP was available in the literature. Once binding pocket was predicted, molecular docking of cPMP in MPT Synthase was performed to identify the binding pocket residues present at the active site and involved in the formation of MPT from cPMP. Quantum Mechanical studies were performed to study the reaction mechanism for MPT formation. Based on the electronic charge on cPMP two reaction mechanisms, differing in charge on cPMP, were studied using B3LYP/LANL2DZ and B3LYP/SDD levels of DFT. As a result of this computational analysis, it has been concluded that cPMP is more efficient/active in its anionic state where phosphate of cPMP was with one negative charge. Also, comparing the LANL2DZ and SDD basis sets, the conclusion is that LANL2DZ provides the more efficient results and can be utilized for the computational analysis of biological systems.

# REFERENCES

1. Habib, U., & Hoffman, M. (2017). Effect of molybdenum and tungsten on the reduction of nitrate in nitrate reductase, a DFT study. *Chemistry Central Journal*, 11(1), 35.
2. Mendel, R. R. (2009). Cell biology of molybdenum. *Biofactors*, 35(5), 429-434.
3. Mendel, R. R., & Kruse, T. (2012). Cell biology of molybdenum in plants and humans. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1823(9), 1568-1579.
4. Leimkühler, S. (2017). Shared function and moonlighting proteins in molybdenum cofactor biosynthesis. *Biological chemistry*, 398(9), 1009-1026.
5. Belaidi, A. A., & Schwarz, G. (2017). Molybdenum Cofactor in Humans: Health, Disease, and Treatment. In *Molecular, Genetic, and Nutritional Aspects of Major and Trace Minerals* (pp. 399-410). Academic Press.
6. Schwarz, G. (2016). Molybdenum cofactor and human disease. *Current opinion in chemical biology*, 31, 179-187.
7. Kisker, C., Schindelin, H., & Rees, D. C. (1997). Molybdenum-cofactor-containing enzymes: structure and mechanism. *Annual review of biochemistry*, 66(1), 233-267.
8. Kisker, C., Schindelin, H., Baas, D., Rétey, J., Meckenstock, R. U., & Kroneck, P. M. (1998). A structural comparison of molybdenum cofactor-containing enzymes. *FEMS Microbiology Reviews*, 22(5), 503-521.

9. Iobbi-Nivol, C., & Leimkühler, S. (2013). Molybdenum enzymes, their maturation and molybdenum cofactor biosynthesis in *Escherichia coli*. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, 1827(8-9), 1086-1101.
10. Yokoyama, K., & Leimkühler, S. (2015). The role of FeS clusters for molybdenum cofactor biosynthesis and molybdoenzymes in bacteria. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1853(6), 1335-1349.
11. Barrière, F. (2003). Modeling of the molybdenum center in the nitrogenase FeMo-cofactor. *Coordination Chemistry Reviews*, 236(1-2), 71-89.
12. Mendel, R. R. (2009). Cell biology of molybdenum. *Biofactors*, 35(5), 429-434.
13. Schwarz, G., & Mendel, R. R. (2006). Molybdenum cofactor biosynthesis and molybdenum enzymes. *Annu. Rev. Plant Biol.*, 57, 623-647.
14. Veldman, A., Santamaria-Araujo, J. A., Sollazzo, S., Pitt, J., Gianello, R., Yaplitto-Lee, J., & Fairweather, J. (2010). Successful treatment of molybdenum cofactor deficiency type A with cPMP. *Pediatrics*, 125(5), e1249-e1254.
15. Schwarz, G. (2005). Molybdenum cofactor biosynthesis and deficiency. *Cellular and Molecular Life Sciences CMLS*, 62(23), 2792-2810.
16. Reiss, J., & Hahnewald, R. (2011). Molybdenum cofactor deficiency: mutations in GPHN, MOCS1, and MOCS2. *Human mutation*, 32(1), 10-18.
17. Reiss, J. (2000). Genetics of molybdenum cofactor deficiency. *Human genetics*, 106(2), 157-163.
18. Lewars, E. (2003). *Computational chemistry. Introduction to the theory and applications of molecular and quantum mechanics*, 318.
19. Young, D. (2004). *Computational chemistry: a practical guide for applying techniques to real world problems*. John Wiley & Sons.

20. Lewars, E. G. (2010). Computational chemistry: introduction to the theory and applications of molecular and quantum mechanics. Springer Science & Business Media.
21. Lewars, E. G. (2016) Computational chemistry: introduction to the theory and applications of molecular and quantum mechanics, Springer.
22. Morris, G. M., & Lim-Wilby, M. (2008). Molecular docking. In Molecular modeling of proteins (pp. 365-382). Humana Press.
23. Daniels, J. N., Wuebbens, M. M., Rajagopalan, K. V., & Schindelin, H. (2008). Crystal structure of a molybdopterin synthase– precursor Z complex: Insight into its sulfur transfer mechanism and its role in molybdenum cofactor deficiency. *Biochemistry*, 47(2), 615-626.
24. Rudolph, M. J., Wuebbens, M. M., Turque, O., Rajagopalan, K. V., & Schindelin, H. (2003). Structural studies of molybdopterin synthase provide insights into its catalytic mechanism. *Journal of Biological Chemistry*, 278(16), 14514-14522.
25. Daniels, J. N., Wuebbens, M. M., Rajagopalan, K. V., & Schindelin, H. (2008). Crystal structure of a molybdopterin synthase– precursor Z complex: Insight into its sulfur transfer mechanism and its role in molybdenum cofactor deficiency. *Biochemistry*, 47(2), 615-626.
26. Crystal Structure of Human Molybdopterin Synthase Catalytic Subunit (Mocs2B)Vollmar, M., Kiyani, W., Krojer, T., Goubin, S., Allerston, C., Froese, D.S., von Delft, F., Burgess-Brown, N., Bountra, C., Arrowsmith, C.H., Edwards, A., Yue, W.W.
27. Crystal structure of human molybdopterin synthase complexKopec, J., Bailey, H., Fitzpatrick, F., Strain-Damerell, C., Oberholzer, A.E., Williams, E., Burgess-Brown, N., von Delft, F., Arrowsmith, C., Edwards, A., Bountra, C., Yue, W.W.

28. Hover, B. M., Tonthat, N. K., Schumacher, M. A., & Yokoyama, K. (2015). Mechanism of pyranopterin ring formation in molybdenum cofactor biosynthesis. *Proceedings of the National Academy of Sciences*, 112(20), 6347-6352.
29. Zhang, N., & Li, D. (2017). Druggability analysis of membrane proteins by DoGSiteScorer (No. e2868v1). *PeerJ Preprints*.
30. Volkamer, A., Kuhn, D., Rippmann, F., & Rarey, M. (2012). DoGSiteScorer: a web server for automatic binding site prediction, analysis and druggability assessment. *Bioinformatics*, 28(15), 2074-2075.
31. Wang, S., Li, W., Liu, S., & Xu, J. (2016). RaptorX-Property: a web server for protein structure property prediction. *Nucleic acids research*, 44(W1), W430-W435.
32. Volkamer, A., Griewel, A., Grombacher, T., & Rarey, M. (2010). Analyzing the topology of active sites: on the prediction of pockets and subpockets. *Journal of chemical information and modeling*, 50(11), 2041-2052.
33. allakyan, S., & Olson, A. J. (2015). Small-molecule library screening by docking with PyRx. In *Chemical Biology* (pp. 243-250). Humana Press, New York, NY.
34. Seeliger, D., & de Groot, B. L. (2010). Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *Journal of computer-aided molecular design*, 24(5), 417-422.
35. Salentin, S., Schreiber, S., Haupt, V. J., Adasme, M. F., & Schroeder, M. (2015). PLIP: fully automated protein–ligand interaction profiler. *Nucleic acids research*, 43(W1), W443-W447.
36. Schwede, T., Kopp, J., Guex, N., & Peitsch, M. C. (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic acids research*, 31(13), 3381-3385.



37. Dennington, R., Keith, T., and Millam, J. (2009) GaussView, version 5, Semichem Inc., Shawnee Mission, KS.
38. Hehre, W., Lathan, W., Ditchfield, R., Newton, M., and Pople, J. (1970) Gaussian 70 Quantum Chemistry Program Exchange, Indiana University, Bloomington, Indiana, Program.
39. Edwards, M., Roeper, J., Allgood, C., Chin, R., Santamaria, J., Wong, F., ... & Whitehall, J. (2015). Investigation of molybdenum cofactor deficiency due to MOCS2 deficiency in a newborn baby. *Meta gene*, 3, 43-49.
40. Huijmans, J. G., Schot, R., de Klerk, J. B., Williams, M., de Coo, R. F., Duran, M., ... & Mancini, G. M. (2017). Molybdenum cofactor deficiency: Identification of a patient with homozygote mutation in the MOCS3 gene. *American Journal of Medical Genetics Part A*, 173(6), 1601-1606.
41. Mendel, R. R. (2013). The molybdenum cofactor. *Journal of Biological Chemistry*, 288(19), 13165-13172.
42. Wuebbens, M. M., Liu, M. T., Rajagopalan, K. V., & Schindelin, H. (2000). Insights into molybdenum cofactor deficiency provided by the crystal structure of the molybdenum cofactor biosynthesis protein MoaC. *Structure*, 8(7), 709-718.
43. Tsepis, A. C. (2014) DFT flavor of coordination chemistry, *Coordination Chemistry Reviews* 272, 1-29.
44. Schaftenaar, G., Vlieg, E., & Vriend, G. (2017). Molden 2.0: quantum chemistry meets proteins. *Journal of computer-aided molecular design*, 31(9), 789-800.
45. Leimkühler, S., Wuebbens, M. M., & Rajagopalan, K. V. (2011). The history of the discovery of the molybdenum cofactor and novel aspects of its biosynthesis in bacteria. *Coordination chemistry reviews*, 255(9-10), 1129-1144.

46. Touthkine, A., Han, W. G., Ullmann, M., Liu, T., Bashford, D., Noodleman, L., & Hahn, K. M. (2007). Experimental and DFT studies: Novel structural modifications greatly enhance the solvent sensitivity of live cell imaging dyes. *The Journal of Physical Chemistry A*, 111(42), 10849-10860.
47. T. H. Dunning Jr. and P. J. Hay, *Modern Chemistry*, Ed. H. F. Schaefer III, Vol. 3 (Plenum, New York, 1977) 1-28.
48. Martin, J. M., & Sundermann, A. (2001). Correlation consistent valence basis sets for use with the Stuttgart–Dresden–Bonn relativistic effective core potentials: The atoms Ga–Kr and In–Xe. *The Journal of Chemical Physics*, 114(8), 3408-3420.
49. Schaftenaar, G., & Noordik, J. H. (2000). Molden: a pre-and post-processing program for molecular and electronic structures. *Journal of computer-aided molecular design*, 14(2), 123-134.
50. DeLano, W. L. (2002). The PyMOL molecular graphics system. <http://www.pymol.org>.