Computational Analysis of Binding Energy and Interactions between Molybdenum *Insertase*: Geph-G and Geph-E Domain with Molybdopterin (MPT) in the Biosynthesis of Molybdenum Cofactor

By

Saba Shabber

NUST00000206907

Supervisor

Dr. Uzma Habib

MS Computational Science and Engineering

Department of Computational Sciences

Research Centre for Modeling and Simulation

National University of Sciences and Technology

Islamabad, Pakistan

July, 2020

Computational Analysis of Binding Energy and Interactions between Molybdenum *Insertase*: Geph-G and Geph-E Domain with Molybdopterin (MPT) in the Biosynthesis of Molybdenum Cofactor

By

Saba Shabber

NUST00000206907

Supervisor

Dr. Uzma Habib

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

in

MS Computational Science and Engineering

Department of Computational Sciences

Research Centre for Modeling and Simulation

National University of Sciences and Technology

Islamabad, Pakistan

July, 2020

TH-4 FORM

THESIS ACCEPTANCE CERTIFICATE

APPROVAL

DECLARATION

I hereby declare that my thesis work titled **— *Computational Analysis of Binding Energy and Interactions between Molybdenum Insertase: Geph-G and Geph-E Domain with Molybdopterin (MPT) in the Biosynthesis of Molybdenum Cofactor*** 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.

Saba Shabber

MS Computational Science & Engineering

Dedicated to…

My MMA Coach Ehtisham Karim Shaheem, my uncle Tahir Raza Naqvi (for financially supporting my education), to my best friend Mashaal Shah and most of all to my mother Syeda Iffat Fakhira Naqvi.

ACKNOWLEDGEMENTS

I am very grateful for having the opportunities, strength and health I have had in order to complete this work. I would also like to extend my gratitude towards 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 the RCMS Super Computer facility and institutional support to complete my work.

I am extremely grateful to my supervisor Dr. Uzma Habib Assistant Professor, Research Center for Modeling and Simulation (RCMS), National University of Sciences and Technology (NUST). I am forever indebted to her patience and kindness; she showed faith in my abilities pushing me forward every step of the way. 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 tireless guidance. I am would also like to appreciate Engineer Hassan for providing tremendous 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 people who always believed in me and cheered me on, my friend Mashaal and my MMA team members from Team Fight Fortress namely Coach Sher Alam, Coach Shah Nawaz, Cherryustaad, Uloomi Karim Shaheen, Haider Satti, Haider Farman, Zeeshan Shah (TNT), Ayub Khan, Najam Khan, Zia Mashwani, Sardar Hassan, Aqib Awan and Mrs. Rehaam Atiya; also and Self Works (mental health program); without them beside me I am sure I would have succumbed to the peer pressure and thrown in the towel. Thank you all.

Saba Shabber

Table of Contents

List of Figures

ABSTRACT

Gephyrin, a postsynaptic protein self-amasses to shape a scaffold that anchor glycine and γ-aminobutyric acid Type A receptors ($GABA_AR$) to the cytoskeleton, guaranteeing the accurate accumulation of postsynaptic receptors at the opportune spot. In vertebrates gephyrin is basically composed into an N-terminal G-domain and a C-terminal E-domain associated by a central linker region or C-domain. All three domains of gephyrin work in a synergic manner and neuronal and non-neuronal functions are performed by this protein. One of the numerous jobs of gephyrin is the biosynthesis of molybdenum cofactor (MOCO). The exact mechanism and binding actions of molybdenum *Insertase* (consisting of gephyrin G (Geph-G) and E (Geph-E) domains) are currently obscure and require further research; so as to have a superior comprehension of the metal exchange reaction between Copper (Cu) particle (ion) and molybdate (MoO_4^{-2}). Docking studies and molecular dynamic simulations were completed to determine the binding energies and interaction patterns as part of a function prediction for gephyrin, as the relative orientation of a ligand with a protein may influence the kind of signal created and anticipate both the quality and type of signal produced. Due of this computational investigation, it has been presumed that the metal containing pterin (MPT) is precarious with or without the copper particle subsequently outlining the job of copper particle only as a decoy utilized for substitution. Once adenylated the metal containing pterin (MPTAMP) stabilizes and copper particle is now able to be discharged and the ligand is passed along to the second protein domain Geph-E, where inclusion of molybdenum (Mo) as molydate (MoO_4^{-2}) happens.

CHAPTER 1

INTRODUCTION

1. Gephyrin – The Neurotransmitter Cytoskeleton Linker Protein

Since its revelation in 1982 by Heinrich Betz's group, Gephyrin is known as a focal component that anchors, clusters and stabilizes glycine and γ-aminobutyric acid type A receptor ($GABA_AR$) at inhibitory synapses of the mammalian cerebrum. It self-assembles into a hexagonal cross section and communicates with different inhibitory synaptic proteins. Gephyrin was originally purified as a 93-kDa protein and was related with glycine receptors (GlyRs) and co-purified with polymerized tubulin. Thus, it was thought to be an anchoring protein associated with glycine receptors.[1,2]

Gephyrin has three significant regions: G-domain, C-domain and the E-domain. In vertebrates gephyrin is fundamentally composed into an N-terminal G-domain and a C-terminal E-domain connected by a central linker region or C-domain. The G and E-domains of gephyrin are homologous to the *Escherichia coli (E.Coli)* enzymes molybdopterin *adenylyltransferase* (MogA) and molybdopterin *molybdenumtransferase* (MoeA) respectively, which are associated with the biosynthesis of molybdenum cofactor (MOCO). MOCO is fundamental for the action of different metabolically significant enzymes such as aldehyde oxidase, xanthine oxidase and sulfite oxidase. Gephyrin contains numerous phosphorylation sites, mainly located within the G-domain.[1,3]

The C-domain of gephyrin is engaged in neurotransmission and goes about as a site for posttranslational modification and binding to other synaptic proteins, such as $GABA_A$ receptor-associated protein ($GABA_ARAP$). Specific sequences of the C-domain contribute to the

formation of the GlyR-gephyrin complexes. Be that as it may, the structure for full-length gephyrin protein has not yet been illuminated, partially in light of the fact that specific properties of the C-domain complicate the crystallization of the full-length protein.

Crystal structure studies so far of individual gephyrin domains have shown that while the G-domain has an inherent inclination to trimerize, the E-domain tends to dimerize. In view of these observations it was proposed that the full-length protein forms a planar hexagonal lattice underneath the synaptic membrane. The gephyrin structure permits the association of microtubule and microfilament-associated hexagonal protein into a grid, which facilitates the spatial dispersion of $GABA_A$ receptors at the postsynaptic membrane.[3,4]

The gephyrin gene (GPHN) encoding for gephyrin is highly conserved and show cases a complicated intron-exon structure. The alternative splicing of GPHN leads to multiple gephyrin splice variants expressed in both neuronal and non-neuronal tissues. Transcript variants of gephyrin are found not only in brain and spinal cord but also in liver, kidney, heart, and lungs. Surprisingly, these are also the organs in which molybdenum dependent enzymes the most elevated enzymatic activities.[2,5]

Apart from the incorporation of molybdenum into the MOCO, the Gephyrin protein is found to interact directly with the (1) actin-binding protein known as profiling, (2) post-synaptic localization of major GABA receptor subtypes, (3) DNA-activating protein kinase RAFT1, (4) two alternative splice variants of collybistin and (5) brain specific guanine nucleotide exchange factors. Hemizygous deletions in GPHN have have been portrayed in various other neuro-formative disorders including autism, schizophrenia and seizures. A common deleted area reported is within exons 3–5 in the G-domain of the protein suggesting this area as key to protein function. [3,6]

1.1. Gephyrin Dysfunction and Neurological Disorders

Mutations in GPHN or modifications in protein expression points have been institutionalized in numerous neuro-pathological states comprising of neuro-developmental and neuro-degenerative disorders. GPHN gene's exon 3–5, which encode for the domain of gephyrin, were branded as unrelated in six cases effected by Autism Spectrum Disorders (ASDs), schizophrenia or temporal lobe epilepsy.

Concerning neurodegenerative disorders, the examination of post mortem brain tissues from Alzheimer Disease (AD) patients has permitted insight towards a connotation between gephyrin dysfunction and neurodegenerative diseases. GPHN gene is amongst the nine exceedingly penetrant genetic risk loci for schizophrenia.[3]

The GPHN hemizygous missense mutation (G375D), discovered in a patient with a type of epileptic encephalopathy is reminiscent of Dravet's syndrome, as gephyrin G375D led to discrepancies in gephyrin clusterization with severe damage to GABAergic synaptic utility. Moreover gephyrin G375D mutation was incapable of synthetizing the molybdenum cofactor, crucial for the activation of MOCO-dependent enzymes. According to a different study gephyrin expression can alter MOCO biosynthesis in MOCO-deficient bacteria, as well as molybdenum-reliant mouse cell line and a MOCO-deficient plant mutant indicative of molybdenum cofactor deficiency (MoCD) comprising of gephyrin is revocable.[1,3,6]

1.2. Molybdenum

Molybdenum (Mo) is a trace element, which is the only second row transition metal mandatory for most living organisms, making it an essential micro-nutrient for microorganisms, plants as well as animals. Molybdenum is catalytically inactive in biological systems until it is assimilated into a scaffold. Molybdenum cofactor precursor also sometimes

known as molybdopterin or pyranopterin (MPT) has an ene-dithiolate coordinating molybdenum which is incorporated in two types of scaffolds: molybdenum-iron-sulfur cluster (FeMoco) found in nitrogenase and the more common pterin-based molybdenum cofactor (MPT) found in xanthine oxidase, sulfite oxidase and dimethyl sulfoxide (DMSO) reductase. Biologically active form of molybdenum is molybdate (MoO_4^{2-}) and in this form it enters the cell by active transport system. Molybdenum along with enzymes plays important role in body. It helps in the digestion of food, production of energy and removal of waste from the body.[7–10]

1.3. Molybdenum Cofactor Deficiency

Molybdenum cofactor deficiency (MoCD) is a severe autosomal recessive innate error of metabolism first defined in 1978. Taking into consideration the occurrence of MoCD, it has been classified as an ultra-orphan disease (1: 100,000). In accordance with the food and Drug Administration (FDA) website, 'an orphan disease is defined as a condition that affects fewer than 200,000 people in the United States of America (U.S.A.)', this makes molybdenum deficiency a rare disorder at the very least. Most rare diseases are difficult to manage, treat as well as diagnose especially in the early stages. The FDA also reported that up until the passage of certain laws, patients determined to have a rare illnesses were generally denied access to effective prescriptions medication. A lot of this had to do with prescription drug manufacturers not being able to make substantial revenue from advertising drugs to such trivial groups of people. Hence funded research for orphan product development is not very common.

MoCD is characterized by a neonatal display of seizures, feeding difficulties, severe developmental deferral, microcephaly with brain atrophy and abrasive facial features. MoCD results in deficiency of the molybdenum cofactor dependent enzymes namely sulfite oxidase, xanthine dehydrogenase, aldehyde oxidase and mitochondrial amidoxime reducing component. Lack of these enzymes results in the buildup of sulfite, taurine, S-sulfocysteine

and thiosulfate in the blood and serum of the body which adds to the severe neurological impairment and can be identified with a dipstick tests for urinary sulphite. Indication from clinical trials has proven early treatment with cPMP can turn MoCD type A from a formerly neonatal lethal condition to a somewhat adaptable condition. [27,29,30]

MoCD is caused by bi-allelic mutagenic variants in MOCS1, MOCS2 or GPHN gene. MOCS1 is responsible for around two- thirds of cases (MoCD Type A), followed by MOCS2 (MoCD Type B) and then GPHN (MoCD Type C). When it comes to Type C, the first case of MoCD due to mutagenic variants in GPHN was called in a Danish family. The second patient confirmed to have GPHN mutagenic variants instigating MoCD was born to parents in Algeria and in the neonatal stage existed all the frequently befalling physiological symptoms of MoCD.[10]

CHAPTER 2

LITERATURE REVIEW

2. Molybdenum Absorption into the Cell

All living organism require Molybdenum for normal functioning, they acquire this micronutrient from external medium in the form of molybdate. In eukaryotes, Molybdate enters the cell *via* a specific transporter system known as the Adenosine Triphosphate (ATP) Binding Cassette (ABC); a known superfamily of membrane transporters, wherein transporters facilitating specific and high-affinity molybdate transport belong to the molybdenum transporter (MOT) family respectively. Two members of the MOT family: MOT1 and MOT2 are known for molybdate transport; MOT1 is a high-affinity and specific molybdate transporter that is crucial for efficient molybdate uptake from the soil in case of plants and from extracellular matrix in animals. MOT1 is present at two different subcellular localizations: plasma membrane and mitochondria. MOT2 is a molybdate transporter of the vacuole that is most likely convoluted in Molybdenum storage and its homeostasis in the cytosol as can be seen in Figure 2.1. [11,12] [13]

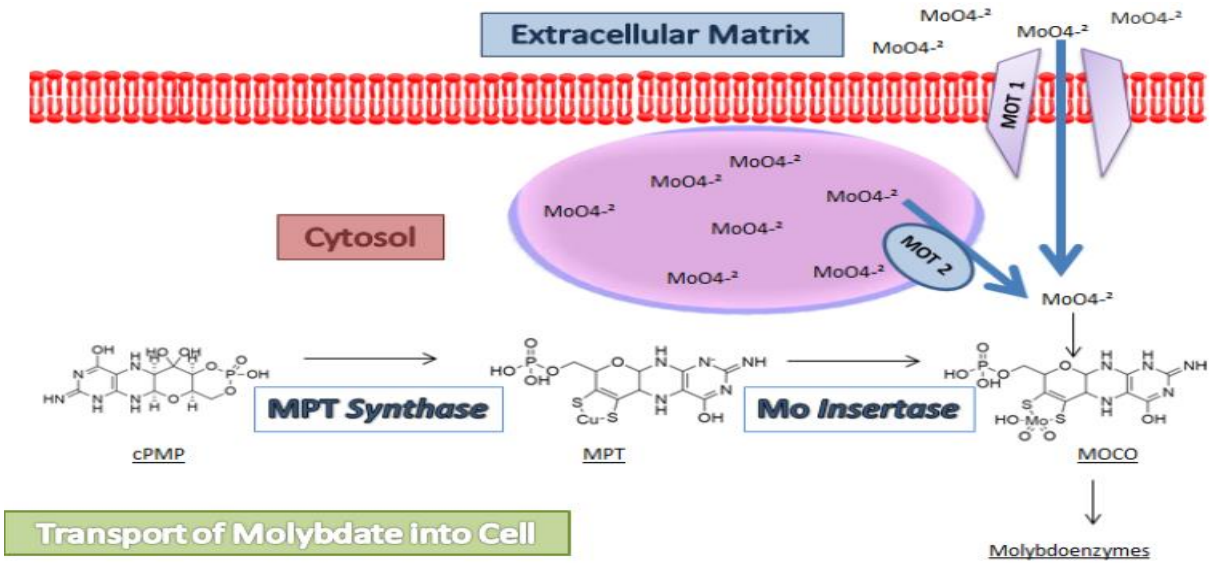


Figure 2.1: Transport of Molybdate into the Cell[14]

2.1. Molybdenum Cofactor Biosynthesis Pathway

MOCO biosynthesis starts in the mitochondria. The amalgamation of most fundamental type of MOCO involves multifaceted association of six proteins, with iron, sulphur, ATP and most importantly copper play a vital role in this process. The procedure comprises of four stages:

(1) Cyclization of guanosine-5'-triphosphate (GTP) to form a cyclo-pyranoperin monophosphate (cPMP) intermediate.

(2) Sulphidation of cPMP to give metal-containing pterin (MPT).

(3) Adenylation of the MPT to generate metal-containing pterin - adenosine mono phosphate (MPT-AMP).

(4) Coordination of molybdate to the enedithiolate group (which can be classified as a metal exchange reaction with copper).

In eukaryotes every one of the six proteins catalyzing the MOCO biosynthesis has been distinguished in plants, fungi and humans. Similar proteins have additionally been

distinguished in prokaryotes like *E.coli*. The nomenclature for the genes and encoded proteins in humans, plants and prokaryotes are molybdenum cofactor sulfurase C-terminal (MOCS), calnexin homolog 1 (Cnx1) and pyranopterin monophosphate *synthase* (Moa) respectively. In *E. coli* the enzymes acyclic pyranopterin monophosphate *synthase* (MoaA) and cyclic pyranopterin monophosphate *synthase* (MoaC) catalyze the conversion of GTP to cPMP i.e. the first step of MOCO biosynthesis, *via* rearrangement of atoms; same reaction was catalyzed by their cognates being Cnx2 and Cnx3 in Arabidopsis thaliana (*A.thaliana*) where as in mammals MOCS1A and MOCS1B perform this task. [8,15,16]

In the second step, sulfidation of cPMP takes place to form the mature MPT, with a functional cis-dithiolene group. This is carried out by MoaB, MoaD and MoaE gene products in the *E.coli* (with the equivalent enzymes in plants and humans being Cnx6 / Cnx7 and MOCS2A / MOCS2B respectively.) [16,17]

The third and fourth step is combined in a bi-functional Cnx1 complex for plants and gephyrin complex for humans, with gene products Cnx1G / Cnx1E and GephG / GephE involved respectively. This bi-functional protein complex is known as molybdenum *Insertase*. In *E. coli*, molybdate is added to MPT cofactor by the gene products MogA and MoeA, which function separately. Hence, during the third step molybdenum *Insertase* (the two-domain protein) catalyzes the Mg-ATP-dependent adenylation reaction of MPT at its G-domain, which is then transferred to its E-domain where the AMP is cleaved and a metal exchange reaction between copper and molybdenum takes place which can be seen in Figure 2.2 .[9,16,18,19]

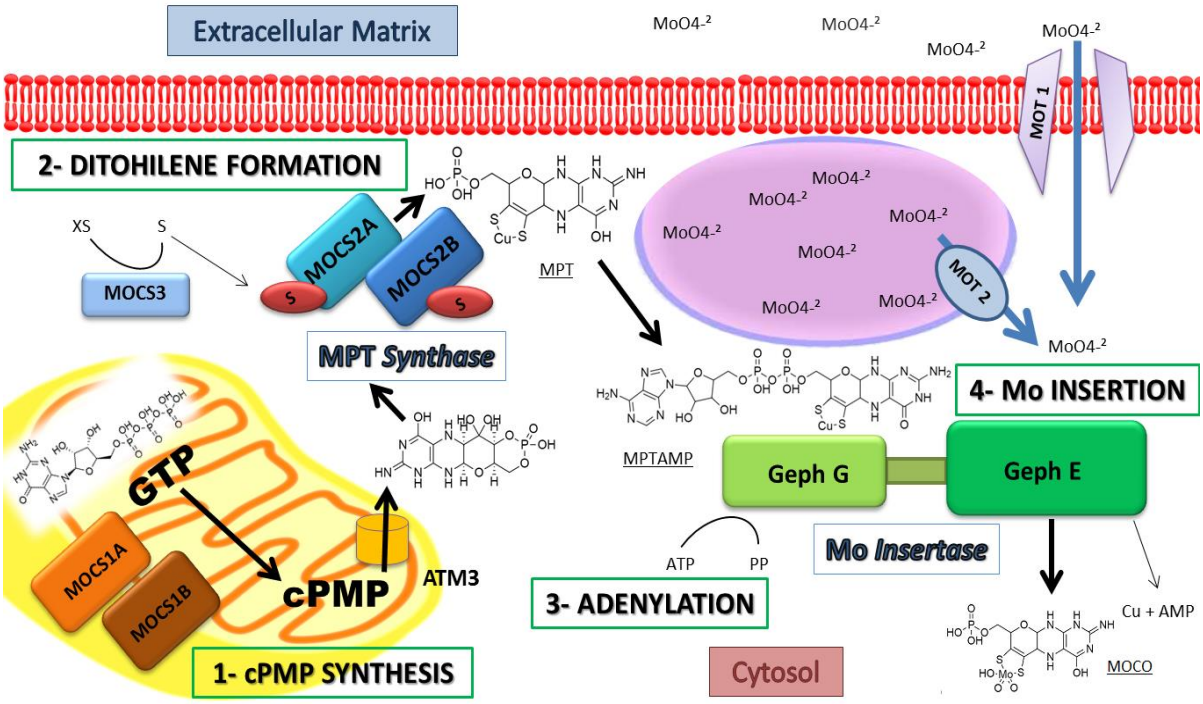


Figure 2.2: Molybdenum Cofactor Biosynthesis Pathway - The biosynthesis starts in the mitochondria and the product of the first step in carried out into the cytosol. The nomenclature for the genes and encoded three homologous proteins in humans, plants and prokaryotes are MOCS, Cnx and Moa respectively. [20,21]

2.2. Molybdenum Cofactor (MOCO)

Focusing on the structure of MOCO, it is eminent that it a one of a kind pterin subsidiary because it has a pterin group, a C6 substituent that coordinates molybdenum *via* an enedithiolate group and a phosphate group (Fig. 2.3). The pterin moiety of MOCO is known for positioning the catalytic metal effectively within the active site; manage the redox activities of the molybdenum atom and electron transmit to or from molybdenum *via* the delocalized electrons within the pterin. X-ray crystallographic investigations of molybdenum enzymes uncover that the cofactor is not bound on the surface of the protein. It is concealed acutely within the center of the enzyme. The exact mechanism of MOCO insertion into the holo-enzyme is not entirely understood; it is uncertain whether MOCO needs to be assimilated before or during completion of the folding process of the apo-enzyme. Once incorporated into the protein the MOCO is permanently bound to the protein and a tunnel-like configuration makes the cofactor available to the substrates. [9,15,22]

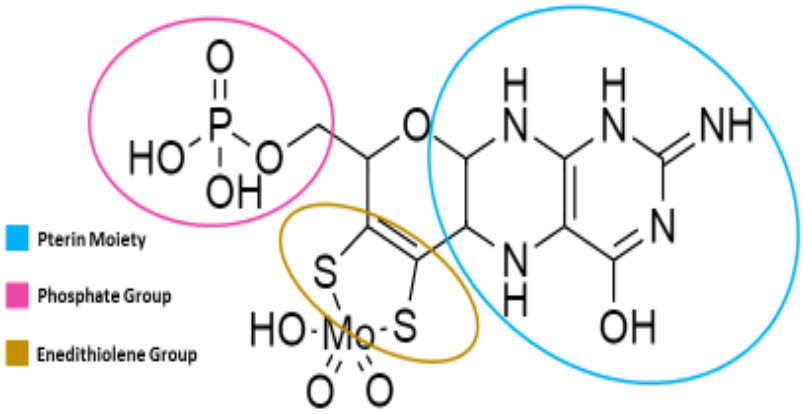


Figure 2.3: Molybdenum Cofactor [20,22]

2.3.Molybdenum Cofactor Dependent Enzymes

Molybdenum based enzymes catalyze the exchange of an oxygen atom to or from a substrate, which is the derivative of a water molecule. These types of reactions are referred to as redox reactions and molybdenum based enzymes can catalyze an extensive array of redox reactions. More than 50 enzymes are reliant on molybdenum, although molybdenum in itself is catalytically dormant in living organisms until it is incorporated in a scaffold. Hence, the metal is not directly attached to the catalytic site but rather incorporated into a low molecular scaffold. These scaffolds are of two types: molybdenum-iron-sulfur cofactor (FeMoco) found in nitrogenase (Figure 2.4) and the more common pterin-based molybdenum cofactor (Figure 2.3) found in xanthine oxidase, sulphite oxidase, DMSO reductase and mitochondrial amidoxime reducing component (mARC).

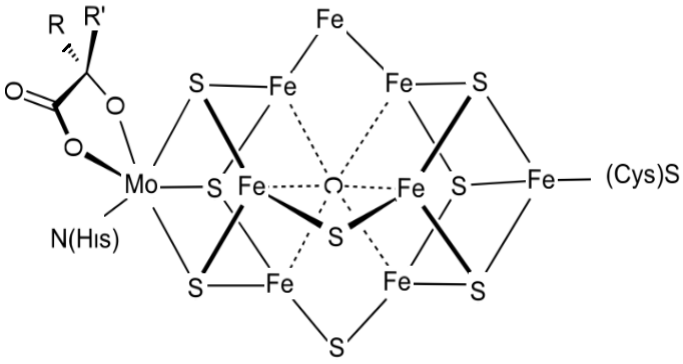


Figure 2.4: Mo-MPT is complex with an [Mo–3Fe–3S] cluster (for nitroganase).[15]

In nitrogenase three homologous systems have been identified so far, out of the three the molybdenum based MoFe-protein has two unique metal centres: P-cluster [4Fe–3S] and FeMo-co, [Mo–3Fe–3S] cluster. The core structure of the other type of molybdenum cofactor, pterin-based molybdenum can be divided into three categories: sulphurated MCD, di-oxo MOCO, bis-MGD. (Figure 2.5) [15,19,22,23]

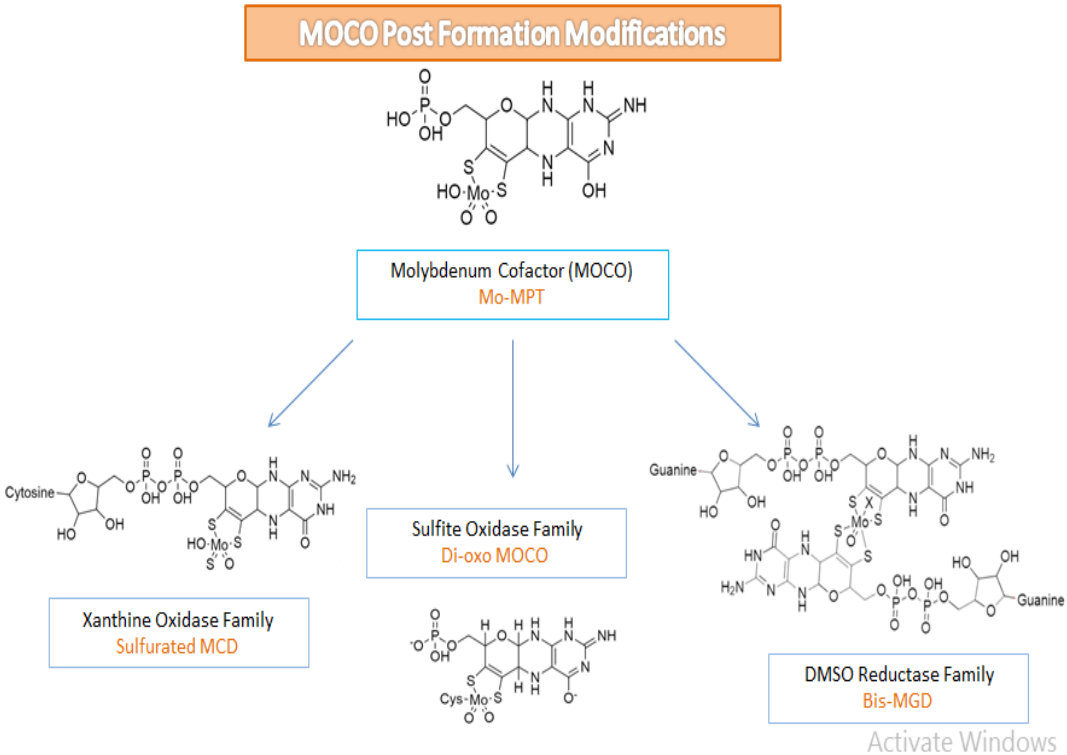


Figure 2.5: The various structures of Mo-MPT in E.*coli* - Mo-MPT can be further altered into three diverse molybdenum-containing enzyme families exist: xanthine oxidase, sulfite oxidase, and DMSO reductase families. In the xanthine oxidase family, Mo-MPT is further modified by the addition of CMP to form sulfurated MCD. Mo-MPT is directly inserted into enzymes of the sulfite oxidase family. For DMSO reductase family, Mo-MPT is further modified by formation of a bis-Mo-MPT intermediate.[19]

2.3.1. Xanthine Oxidase

In xanthine oxidase family (XO), active sites contain a molybdenum atom coordinated with two sulfur atoms of dithiolene of the MPT, one sulfur atom of terminal sulfidic ligand, an oxygen atom and a hydroxide/water molecule. The xanthine oxidase family is fashioned by two homologous members (xanthine oxidase (XO)/xanthine dehydrogenase (XDH)) and Aldehyde Oxidase (AO). XO/XDH is an enzyme convoluted 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.[17,22,24]

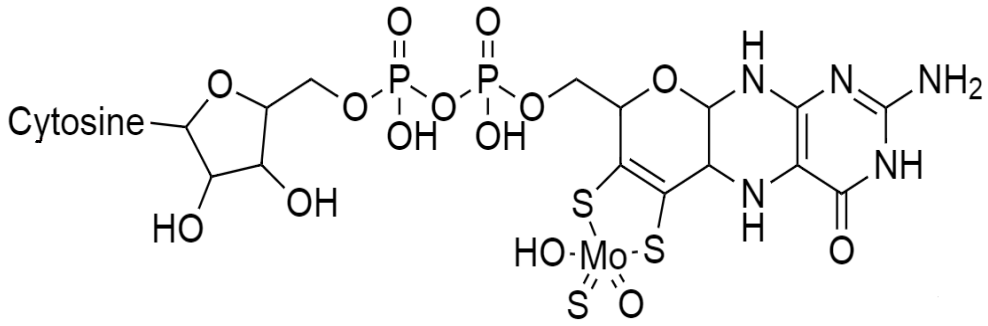


Figure 2.6: Xanthine Oxidase Family[19]

2.3.2. DMSO Reductase

The DMSO reductase family is assorted in both structure and capacity, yet all individuals have two reciprocals of the pterin cofactor bound to the metal. The X can be a serine, a cysteine, selenocysteine, hydroxide, and/or water molecule. The reactions catalyzed by members of this family commonly include oxygen-atom transfer, but sometimes dehydrogenation reactions also occur. Every enzymes of this family has two counterparts of molybdopterin guanine dinucleotide cofactor (MGD) joined together. In addition to this a variant of MPT containing an additional guanosine monophosphate GMP group is also attached

to the phosphate group of MPT. Most other enzymes of the DMSO reductase family also comprise of added subunits and much of the time an iron–sulphur core is present in the MOCO-bound subunit.[8,25,26] (see Figure 2.7)

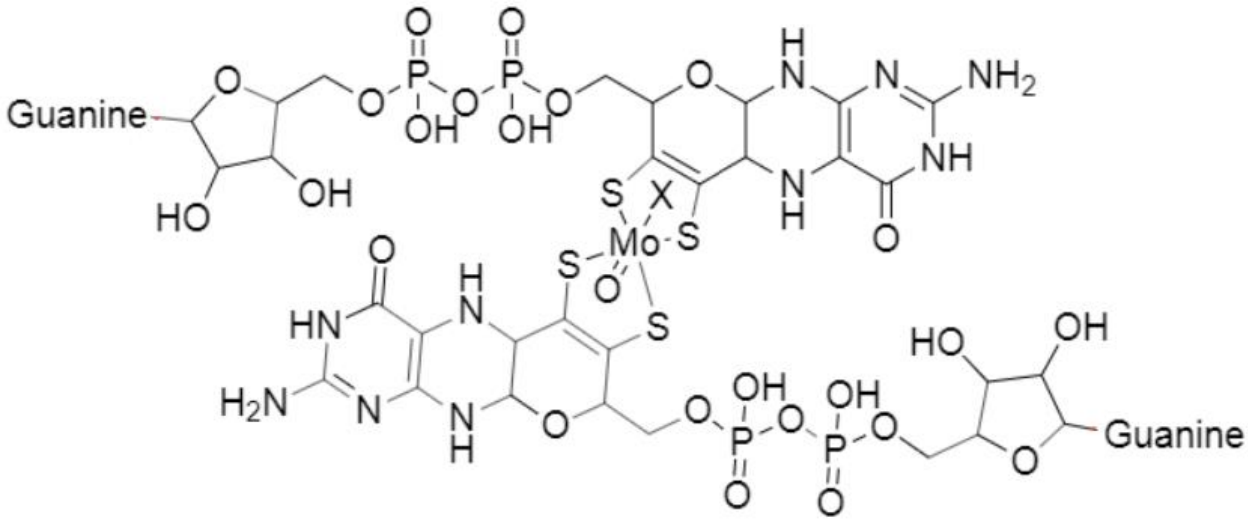


Figure 2.7: DMSO Reductase Family[19]

2.3.3. Sulfite Oxidase

In sulfite oxidase (SO) family, Mo is coordinated with three sulfur and two oxygen atoms. Out of three, two sulfur atoms are from MOCO, whereas, third equatorial sulfur ligand is from conserved cysteine. Apart from sulfite oxidase another recently discovered mitochondrial Amidoxime Reducing Component mARC also makes use of the cofactor. The enzyme was identified by deciphering the mechanism of amidoximes reduction by liver cells' mitochondria. A new MOCO-dependent enzymatic activity was discovered which involves reduction of a pro-drug called benzamidoxime to create its active form benzamidine. The eukaryotic mARC enzymes comprises of two domains: (1) a MOSC domain that is present in the MOS (2) MOCO sulfurases enzymes and its β-barrel domain; mARC enzyme require protein partners to function, as they provide the electrons from NADH to reduce the substrate. The moonlighting proteins formulate a distinct class of multifunctional enzymes which execute

numerous physiological capacities relying upon the concentration of cellular ligands, substrates, cofactors, tissue, cellular localization, oligomeric state and post-translational modification. The human mARCs are located in the mitochondrial membrane with its C-terminal domain protruding in to the cytosol. [8,17,21] (see Figure 2.8)

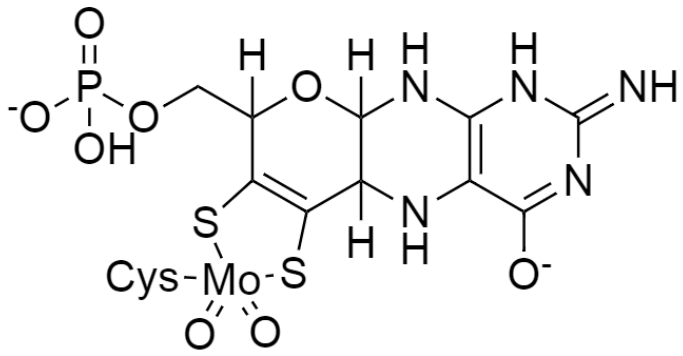


Figure 2.8: Sulfite Oxidase Family[19]

2.4. Mechanism of Molybdenum Insertion – Metal Exchange Reaction

It is at present obscure at which step copper is bound to the dithiolate and was classified as an unexpected finding in crystal structures of Geph-G domain. Once the Molybdenum is inserted into the dithiolate group the molybdenum cofactor is matured and denoted as MOCO. The exact mechanism and binding actions of molybdenum *Insertase* are currently unknown and require further research. Mature MOCO can bind directly to the molybdenum enzymes or domains irreversibly and a tunnel-like structure lets the cofactor become accessible to the substrates. Although the exact mechanism of insertion of the MOCO into the holo-enzyme is not entirely understood; it is uncertain whether MOCO needs to be incorporated before or during completion of the folding process of the apo-enzyme. [17,18]

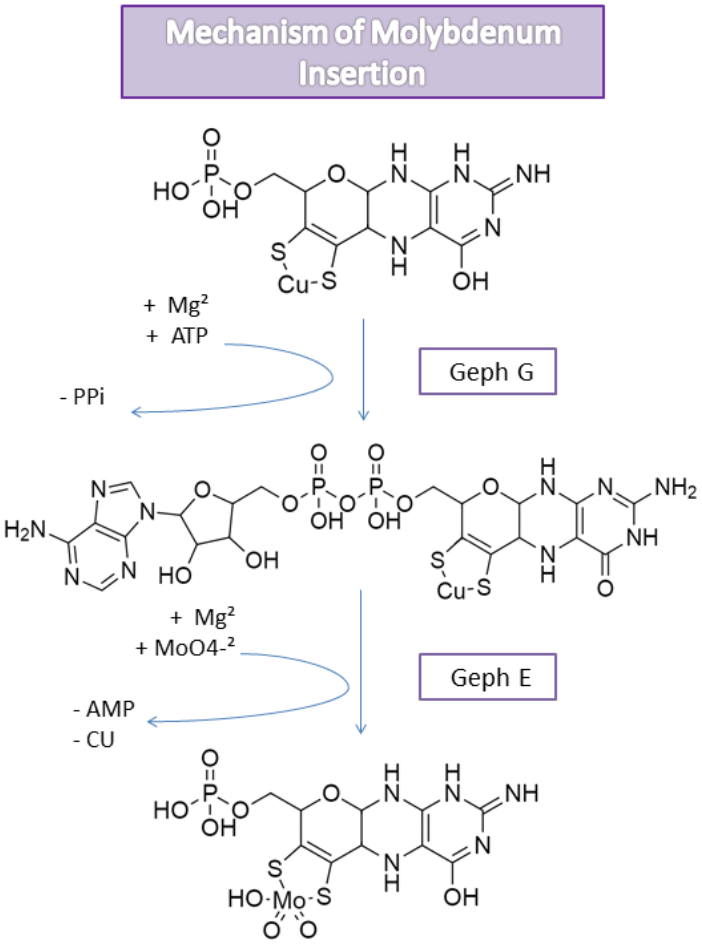


Figure 2.9: Hypothetical mechanism of molybdenum insertion reaction. It is currently unknown at which step copper is bound to the dithiolate, the Cu is exchanged for the Mo in a metal exchange reaction. [23]

2.5. Molybdenum Cofactor Susceptibility to Oxidation / Role of Carrier Proteins

Once the MPT has been successfully converted into MOCO *via* molybdenum *Insertase* in eukaryotes, the newly manufactured cofactor does not occur freely in the cell. It is upheld by being bound to a storage protein called MOCO binding proteins (molybdenum carrier protein (MCP) in Chlamydomonas reinhardtii (*C.Reinhardtii)* and molybdenum binding protein (MoBP) in *A.thaliana* and higher organisms) with a half-life of up to two days, to protect it from oxidation. The MOCO cofactor is extremely sensitive to air-oxidation; hence MoBPs are able to intermingle with both, molybdenum *Insertase* and the Mo-containing enzymes, becoming capable of enclosing the cofactor into the corresponding apo-enzyme. A distinctive

feature of the eukaryotes is that the biosynthesis of the pyranopterin is compartmentalized, meaning that the formation of cPMP taking place in the mitochondrial matrix. After formation the cPMP, it is transported out of the mitochondrion by an ATP-dependent member of the ABC transporter family known as ABC Transporter Membrane protein type 3 (ATM3). In humans, once in the cytosol, sulfuration of cPMP by MOCS2B and MOCS2A takes place with the help of cysteine *desulfurase*. In the cytosol, rest of the synthesis of MOCO takes place, which is then trafficked to members of corresponding apo-enzyme such as sulfite oxidase family with the help of carrier proteins.[11,16]

2.6. Insertion of Molybdenum into the Apo-Enzymes

Sulfuration of the MOCO is very similar to the sulfuration of cPMP and is important for conservation of the cofactor in the cytosol and in some cases is needed for insertion into the apo-enzyme. MOCO sulfurases have been identified from multiple eukaryotes as well as for humans. For eukaryotes abscisic acid deficient 3 (ABA3) performs the sulfuration and for prokaryotes cysteine *desulfurase* subfamilies perform sulfuration. Cofactor inclusion in the case of sulfite oxidase family seems to be kept at a minimal, as the cofactor does not require being deeply buried in the holo-enzyme. The sub-domains could easily move apart transiently to incorporate the cofactor; thermal energy is enough to prompt the essential conformational changes. It is also reported that molybdenum *Insertase* may be convoluted in a specific protein–protein interaction with apo-enzymes to insert the cofactor directly.

In case of most of the bis MGD-utilizing enzymes, such as DMSO reductase family, incorporation of the cofactor into apo-enzymes may be assisted by chaperone proteins as it is quite complex. Not only is the cofactor buried deeply the bis MGD cofactor in itself is a complex version of the MOCO enzyme. The chaperone stabilizes the respective apo-enzyme enabling the cofactor insertion. C-terminal of the MGD-utilizing enzymes has a cap towards

the rear side of protein and exclusion of this cap reveals the totality of the bis MGD cofactor. A widespread funnel offers substrate access to the active site molybdenum center, which is an exceedingly conserved article of DMSO reductase family.

The Xanthine oxidase family has the trickiest insertion pattern of the molybdenum center into enzymes out of all the other families. Needless to say the cofactor needs to be buried deeply and the funnel that gives substrate access to the active site molybdenum center is only 15 Å-long. Enzymes accountable for sulfuration of the cofactor, partake directly in its incorporation of it into the enzymes of xanthine oxidase family. [16,19,20,27]

2.7.Role of Molybdenum *Insertase* in the Biosynthesis of Molybdenum Cofactor: Structural and Functional Comparison in both Plants and Animals

The bi-functional Cnx1 complex for plants and gephyrin complex for humans, with gene products Cnx1G / Cnx1E and GephG / GephE domains, are associated by a linker region comprising of at least 160 residues in Geph and only 25 residues in Cnx1. When compared to MogA, the eukaryotic Cnx1G and GephG gene products have four major differences in their three-dimensional structures namely: (1) In Geph-G and Cnx1G insertion of an additional a-helix in between the first b-strand and a-helix of MogA is observed. (2) The loop between a-helix2 and b-strand2 has conformational changes for the eukaryotic proteins (3) A b-hairpin loop found in MogA is absent from the eukaryotic gene products respectively. (4) The C terminus of Geph-G tracks a different path from that in MogA.

Previously, the plant Mo-insertase Cnx1 has aided as a search model to detect its human homolog. What was [3]revealed was: (1) the protein was already known but with a completely Mo-unrelated function, namely as ‘Gephyrin’ which is ubiquitously expressed mainly in the central nervous system, where it is essential for clustering of inhibitory neuro-receptors in the

postsynaptic membrane (2) plants have the E domain on the N-terminus, in mammals it is the G domain, hence making the two homologues protein have inverse domains.[17]

Apart from the incorporation of Molybdenum into the MOCO, the Gephyrin protein is found to interact directly with the (1) actin- binding protein known as profiling, (2) post-synaptic localization of major GABA receptor subtypes, (3) DNA-activating protein kinase RAFT1, (4) two alternative splice variants of collybistin and (5) brain specific guanine nucleotide exchange factors. Deletion in the Gephyrin coding gene was identified to have played a role in symptoms typical of MOCO deficiency. [13,20,22]

Cnx1-G and Geph-G form stable trimers in the crystals. The trimerization is mediated by residues in a5, b4, a7, the 310-helix and the loop connecting b3 with a4. The packing of GephG in crystals is characterized by a layer-like arrangement of its trimers. Phe90 residue is known for multiple hydrophobic interactions with residues of the adjacent monomer, which is characteristic only in eukaryotes, consequently establishing bonding between individual monomers of the trimer. Electrostatic comparison of the properties of the eukaryotic G-domains and prokaryotic MogA reveal that the distribution of charged residues is well conserved for all three homologous gene products. Although MogA is more closely related to GephG than it is to Cnx1G, with 34.1% and 30.5% similarity respectively. [11,12,28]

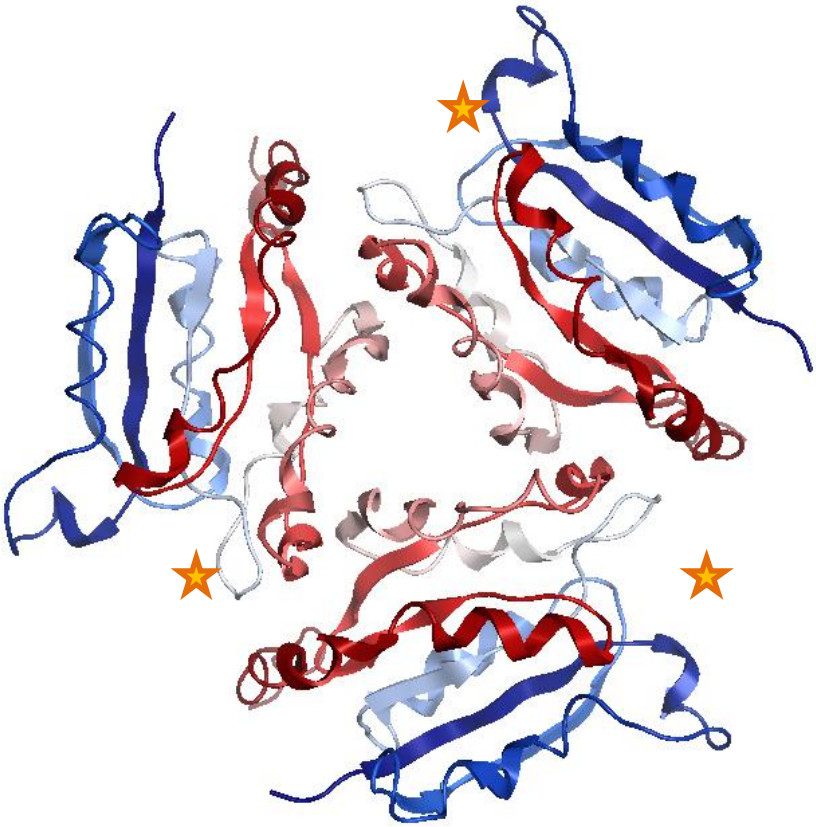


Figure 2.10: Geph-G trimer observed along 3-fold axis: – ★ mark the possible binding sites for MOCO. [9]

2.8.Binding Pockets of the two Domains of Molybdenum *Insertase*

The binding pocket for MPT is highly conserved for the three homologous proteins; statistics from site-directed mutagenesis studies indicate that in the binding site of Cnx1, it is anticipated that MPT binds into the cleft between the functionally vital b3 and a4; encompassing the highly conserved Gly-Gly-Thr-Gly motif (residues 86-89), Leu-Pro-Gly motif (144-146) and aspartate residues (Asp61 and Asp94).

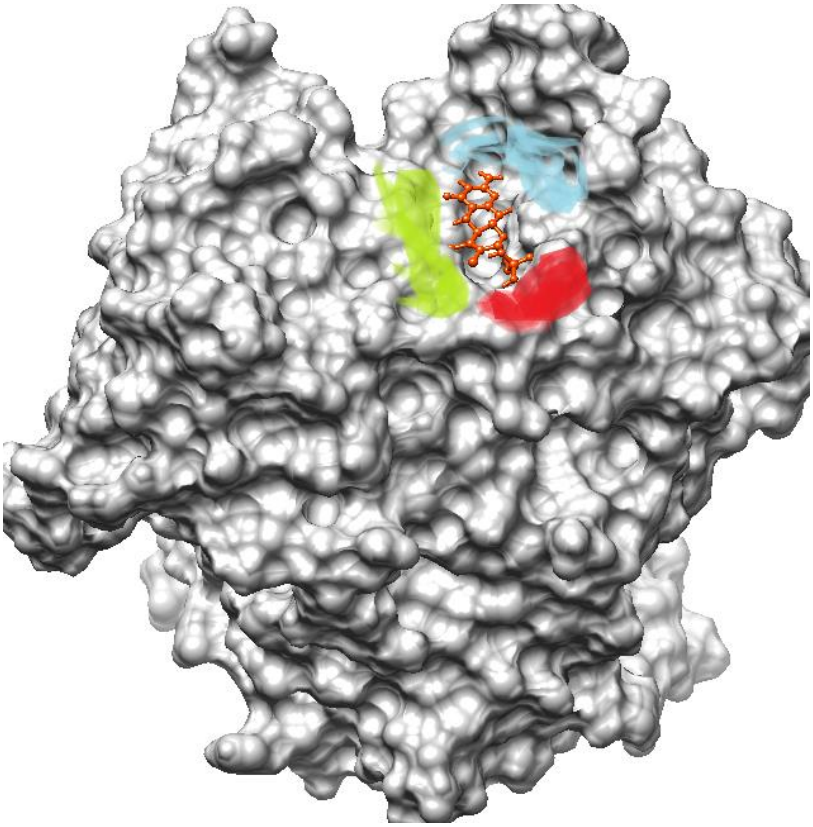


Figure 2.11: Gly-Gly-Thr-Gly motif in red (residues 86-89); Leu-Pro-Gly motif in green (144-146) and aspartate motif in blue (Asp61 and Asp94) within the binding pocket. [9]

The dithiolene group of MPT which is involved in the metal exchange reaction and Mo is in close proximity of two catalytically important aspartate residues. The bend amongst the pterin and the pyran ring compliments the shape of the binding pocket. For the second domain Geph-E Serine (Ser630) is known to play an important role in the insertion of molybdate into the dithiolene group of MPT replacing Cu in a metal exchange reaction. [9]

Problem Statement

In the central nervous system, gephyrin has been shown to anchor neurotransmitter receptors to the cytoskeleton. It is also the only known mammalian homolog to MogA and MoeA. The exact mechanism and binding actions of molybdenum *Insertase* (consisting of gephyrin G and E domains) are currently unknown. Hence it is vital to understand the detailed reaction mechanism for synthesis of MOCO from MPT (which is the last step) and its intimidator MPTAMP, in MOCO biosynthesis.

Objectives

- Finding the role of Cu in MOCO biosynthesis.
- Finding the binding interaction energies and function prediction of both molybdenum *Insertase* domains with their respective substrates namely MPT and MPT-AMP.
- Finding the metal exchange reaction between copper and molybdenum
- Determining a possible linkage between molybdenum *Insertase* domain Geph-E and MoCD type C.

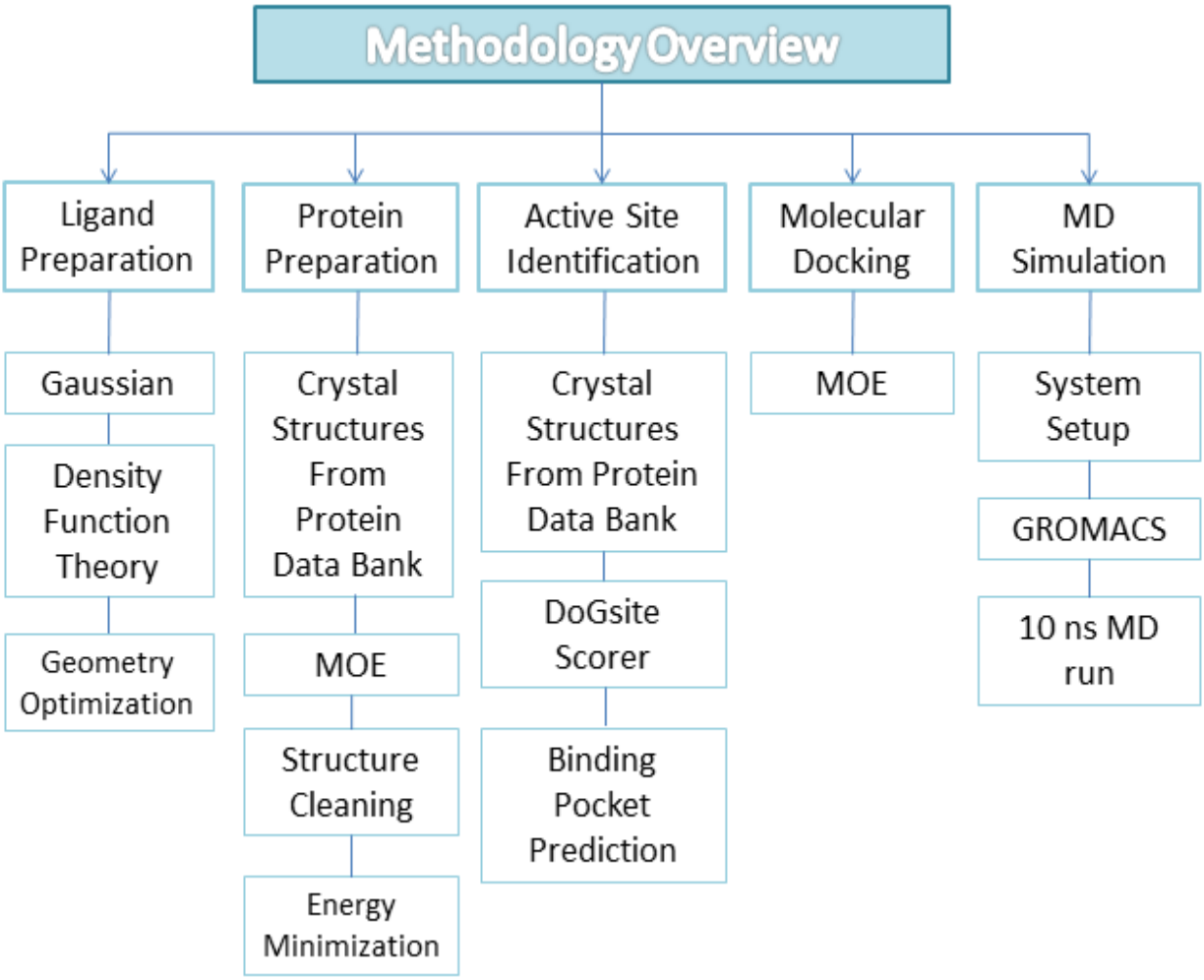
CHAPTER 3

METHODOLOGY

3. Computational Methods

Medicinal chemist uses computational biochemistry in synergy with results of theoretical chemistry, assimilated into effective computer programmers, to predict structure and properties of molecule. Computational biochemistry analyzes the characteristics of molecule such as: charge distribution, relational energy, dipole moment, structure, vibrational frequency and reactivity. Computational analyses have three ranges: *ab* initio methods (high accuracy), semi-empirical methods (moderately accurate) and lastly the molecular mechanics method (low accuracy).

This research work was conducted to comprehend the reaction mechanism of MOCO synthesis from MPT with the help of enzyme Mo *Insertase*, which consists of two protein domains Geph-G and Geph-E. The project is completed in 05 steps mentioned below:



3.1.Ligand Preparation:

Formation of MOCO from MPT is a two-step process; (1) formation of MPT-AMP from MPT and (2) formation of MOCO from MPT-AMP. As there are two ligands MPT and MPT-AMP involved hence they were chosen for geometry optimization:

(1) Molybdopterin - MPT

(2) Adenylated Molybdopterin - MPTAMP

The structure of MPT and MPT-AMP were modeled using Gaussview6[31]. Zinc database[32] entries ZINC40164370 and ZINC62235906 were used as reference. The molecular geometries

of MPT and MPT-AMP, in their ground state, were optimized using the Gaussian 09 program package .[33].

3.1.1. Gaussian-09

Gaussian software was introduced in 1970 by John Pople.[33] The Gaussian package is used for finding energies, molecular geometries, lowest and highest occupied molecular orbital (LUMO and HOMO), spectroscopy as well as the transition states involved in reaction mechanisms. The input files for it are generated using a graphic user interface (GUI) called Gaussview[31] whereas output files are viewed in the linux environment using GUI Molden[34].

3.1.2. Density functional theory (DFT)

Density Functional Theory (DFT) is a method typically used nowadays to calculate ground-state electronic structure calculations of atoms, molecules and solid state materials. This method is good for studying structure, electronic state and different spectra of 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.[35]

3.1.3. Geometry Optimization

Geometric optimization is a method of taking rough geometric approximations and making them as exact as possible, in order to do this DFT calculations were performed on modeled structures using B3LYP exchange correlation functional and Lanl2DZ basis set. [31] All simulations were performed on the High Performance Computer Cluster servers (Ganglia

FING-Rocks-Cluster)[36] available at the Research Center for Modeling & Simulations – National University of Science and Technology (RCMS-NUST); using Gaussian 09 program package [33] *via* Linux. To visualize input, output and simulation steps of the Gaussian09 job run, MOLDEN[34] was used. The calculated electrostatic potential (ESP) charges were taken as partial charges for the ligands.

3.2. Protein Preparation

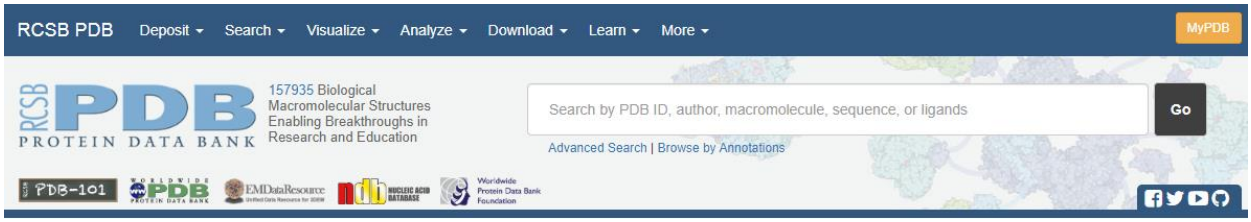
For protein preparation, first the PDB structures are taken from online platform and then the PDB structures are cleaned using software to remove any pre attached ligands. Next the protein's energy is minimized to mimic natural state of protein in the body.

3.2.1. Protein Data bank (PDB)

The crystal structures of the two domains of Molybdenum *Insertase*: Geph-G and Geph-E were obtained from Protein Data bank (PDB)[37] having PDB ID:

- 1jlj (GephG)[9]
- 4tk1(GephE)[38]

It is noteworthy that in case of Geph-G crystal structure (PDB ID: 1jlj), the reported organism is *Homo sapiens*. Whereas for Geph-E crystal structure (PDB ID: 4tk1) the reported organism is *Rattus norvegicus*; for the sake of consistency same organism for both domains should be preferred. Hence, upon template search for sequence alignment of human gene GEPH_HUMAN (Uniport ID: Q9NQX3) *via* Swiss model, it was discovered that PDB ID: 4tk1 showed 100% sequence similarity to human gene inferring that the protein is entirely conserved within mammals. [38,399]



3.2.2. Molecular Operating Environment (MOE)

After obtaining both protein structures from PDB, they were prepared for molecular active site identification and molecular docking using Molecular Operating Environment (MOE) software (2019 version issued by the Chemical Computing Group).[40] For both protein domains, additional use of software called 'University of California, San Francisco' or UCSF Chimera[41] was used. In case of Geph-G it was to import the Cu ion in the binding pocket from its plant homolog Cnx1 (PDB ID: 1uux)[23]. In case of Geph-E a molybdenum cluster was imported into the binding site, from another Geph-E protein (PDB ID: 5eru)[42] of lower resolution.

3.2.3. Structure Cleaning

Now the protein crystal structures (individually) were imported into Molecular Operating Environment (MOE)[40] software and edited to remove all water molecules along with

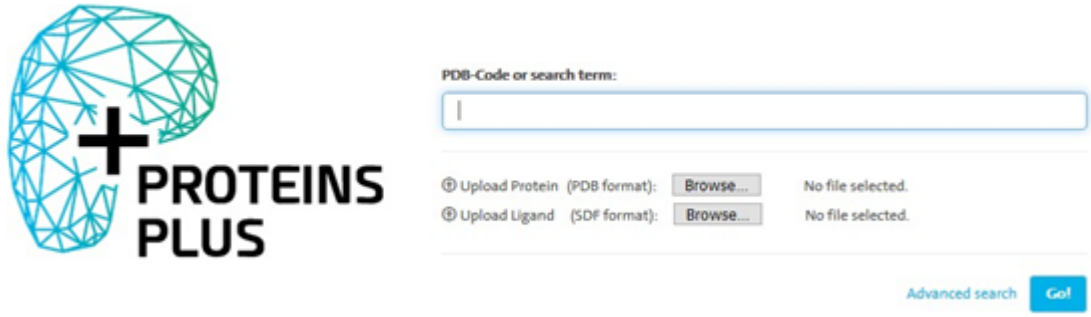
any pre-attached ligands. The tautomeric states of residues were adjusted *via* protonation for which the pH was kept at 7.0, temperature was maintained at 300K and the salt (ionic) concentration retained at 0.1mol/L (according to the Generalized Born-GB electrostatic model). The electrostatic function chosen to be 'GB/VI' (Generalized Born/ Volume Integral) while the van der Waals (vdW) function was depicted to be '800R3' (800R3 simulates the repulsive part of vdW energies), the electrostatic interactions cut off was kept at 15Å.

3.2.4. Energy Minimization

After protonation the protein crystal structures were energy minimized. Protein energy minimization methods are ingrained in observations that native protein structures resemble a system at thermodynamic equilibrium with the least free energy, hence for the energy minimization 'Amber10'[41] force field was used while allowing adjustment to 'H-atoms'. The root-mean-square deviation (RMSD) gradient was kept at '0.1 kcal/mol/A^2'. (Energy minimization is terminated when RMSD gradient falls below this value)

3.3. Active Site Identification

After obtaining protein from PDB and optimized ligand, next step was the prediction of protein binding pocket for the two domain of Molybdenum *Insertase*. Hence the well-known software for binding pocket prediction tool DoGsite Scorer[43] was used, chosen from previously reported literature, to check the drug score of the binding pockets. DoGsite Scorer is a web-server for automatic binding pocket prediction and analysis. It is a 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 the ligand.



3.4. Molecular Docking

Molecular modeling is a field in which methods like docking are used to calculate the preferred orientation of one molecule to a subsequent molecule, when bound to one another to form a steady complex. Understanding of the preferred orientation may in turn be used to calculate the strength of association or binding affinity between two molecules, using for example ‘scoring functions’. Molecular docking is used for two purposes: one is signal transduction and the other is drug design. The signal transduction is the association between biologically significant molecules such as nucleic acids, proteins, lipids and carbohydrates perform an essential role in signal transduction. Furthermore, the relative orientation of the two cooperating partners may affect the type of signal formed (agonism, antagonism). Therefore docking is useful for calculating both the type of signal produced and strength of the signal.

In the case of drug discovery and design docking is needed to calculate the orientation of small molecular candidates of drugs, to their protein targets so that predictions on the affinity and action of the small molecule can be made. Therefore, docking is vital for rational drug design. [4445]

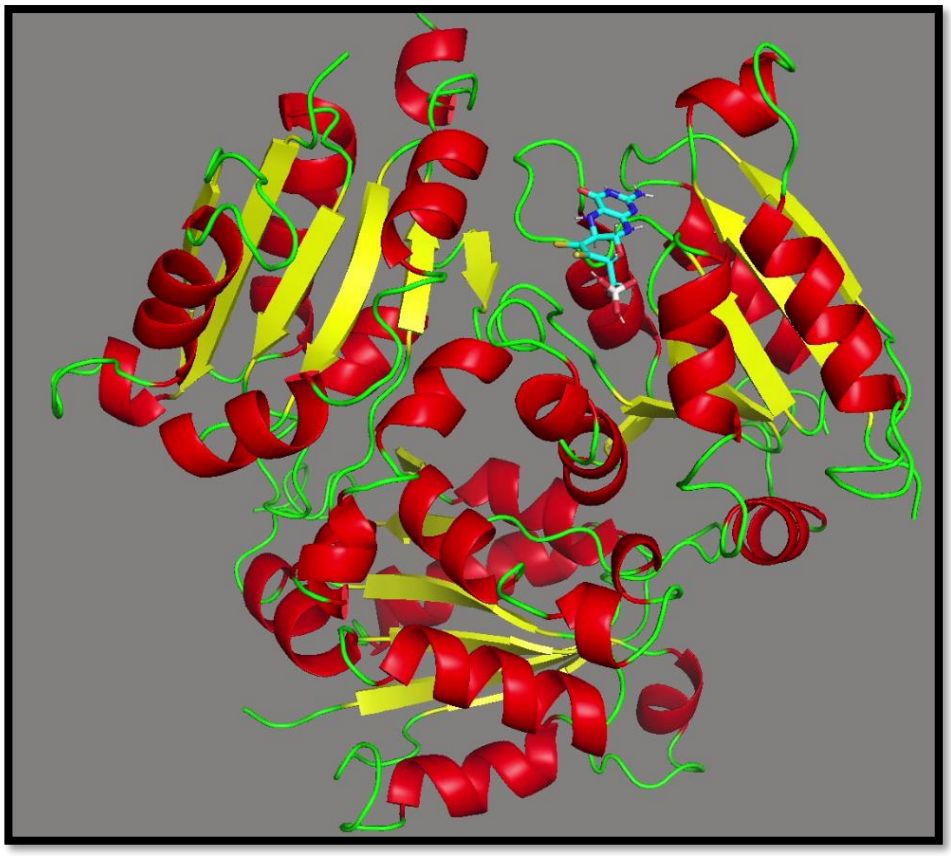


Figure 3.1: Protein-Ligand Complex - Molecular Docking

After active site is predicted both ligands MPT and MPT-AMP were docked into their respective protein domain's orthosteric binding site with the help of Molecular Operating Environment (MOE) software[40] (MPT into GephG / MPT-AMP into GephE). The placement method chosen was 'Alpha Triangle' (which drives poses by random superposition of ligand atom triplets and alpha sphere dummies as receptor cites) with post placement refinement kept as 'induced fit'(changing receptor pocket to cater to the ligand). The initial scoring function was kept as 'London dG' (is a hybrid of empirical + knowledge-based functions, which estimates the binding free energy (ΔG°) of the ligand from a given pose) and 'GBVI / WSA dG' ((Generalized Born/ volume integral/ Weighted Surface area) a force field-based scoring function, which estimates the free energy of binding of a ligand from a given pose) as the final scoring function. No 'wall constraints' were used. The receptor was kept at 'Receptor' while receptor site was chosen to be 'selected atoms'. The electronic density was depicted at 1σ with

a resolution of 2.5 (for simulated electron densities). The search efficiency was kept at 100% (software provides 1-500% as an option). A total of 150 poses were generated in each case, out of which 134(MPT/GephG) and 100(MPT-AMP/GephE) poses were retained by the software. In addition, MPT (with an attached molybdate) was docked in GephE to replicate the MOCO cofactor formation with a total of 90 poses retained out of the 100 generated.

3.5. Molecular Dynamics Simulations

Molecular dynamics (MD) is a molecular mechanics program intended to imitate the movement of atoms within a molecule. MD can be done on a molecule to produce altered conformation which upon energy minimization, give a range of steady conformations. The MD method is used for calculation of the molecular motion of several particles that are interacting in classical manner. In MD, motions of molecules are within some specific time-frame i.e. one time step (fs); atoms have some velocities and are subjected to some forces. The MD simulation is used to observe the conformational changes in molecules by assimilating Newton's second law of motion (F=ma). Motion of a single molecule or of large amount of molecules can be studied by this method. [40,46]

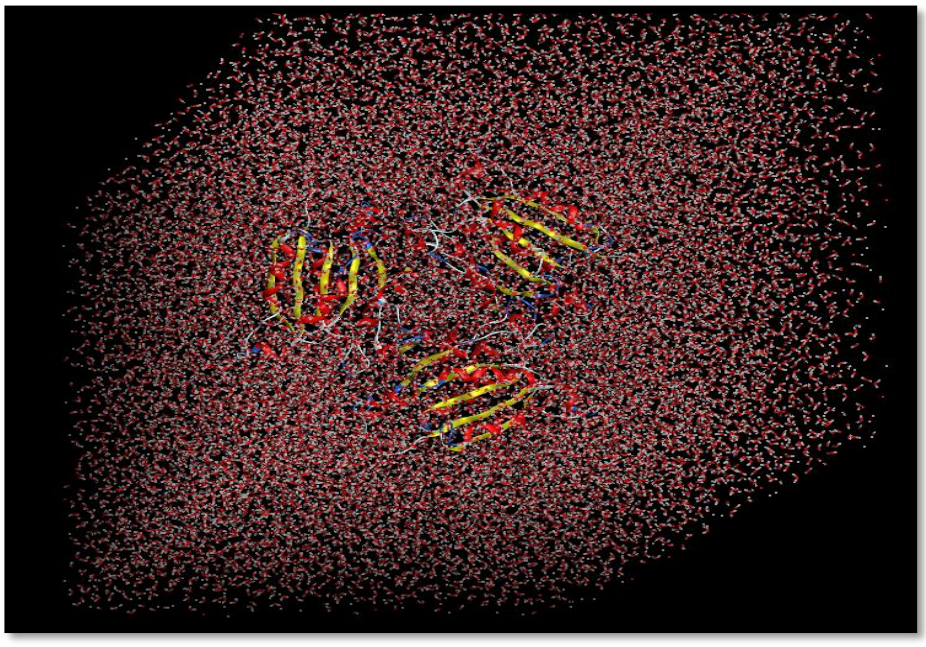


Figure 3.2: Protein-Ligand Complex in a solvent box - MD simulation[47]

3.5.1. GROMACS

All MD simulations were performed using GROningen MAchine for Chemical Simulations (GROMACS) v5.1.0[48], following GROMACS tutorial-5[49] (Protein-ligand Complex) while using the parameter files provided within the tutorial, with the CHARMM36 force field for all simulations. The force field parameters for the ligands were acquired using ParamChem with Chemistry at Harvard Macromolecular Mechanics (CHARMM) general Force Field (CGenFF)[50]. Periodic boundary conditions were used. Newton's equations of motion were incorporated by means of leap- frog algorithm. A 2fs (femtosecond) integration time step was utilized, the bonds amongst hydrogen atoms and every heavy atom constrained to their equilibrium lengths using the Linear Constraint Solver (LINCS) algorithm. The temperature was sustained at 310K (kelvin) using a Nosé-Hoover thermostat with a coupling time constant of 1.0ps. The system box was allowed to fluctuate under 1atm (atmospheric pressure) using a semi-isotropic Parrinello-Rahman barostat. For both Van der Waals (vdW) and electrostatic interactions, cutoffs of 1.2nm (nanometer) were applied. Long-range electrostatic interactions were calculated using the particle mesh Ewald (PME) method. All systems were minimized and then equilibrated for a total of 10 ns, including NVT (constant number, volume and temperature) and NPT (constant number, volume and pressure) with the Berendsen weak coupling method. All these parameters were used in accordance with the GROMACS tutorial-5[49] (Protein-ligand Complex) hence making them grossly generalized and applicable on average to any MD simulation. [49]

In order to perform molecular dynamics simulations a Linux environment was preferred, hence Ubuntu v2.02 was installed whereas the Python v3.7.3 was made use of. The GUI used to manipulate input files were PyMol v2.0[51] and Avogadro v1.2.0[52], similarly in order to visualize the output files VMD v1.9.3[53] was used. The simulations were carried out on

supercomputing facility of research and education center (ScREC) at Research Center for modeling & Simulation (RCMS) NUST.

CHAPTER 4

RESULTS

4. Computed Results

MOCO biosynthesis is a stepwise process that consists of four steps. This study is focused on the computational analysis of its last step i.e., the formation of MOCO from the MPT with the help of enzyme Molybdenum *Insertase* (Geph-G and Geph-E domain).

This study was completed in following four stages:

Ligand Preparation / Geometry Optimization

Active Sites / Binding Pocket Search

Molecular Docking / Binding Energies

Molecular Dynamics/ Interaction Patterns

4.1.Ligand Preparation / Geometry Optimization

According to the literature, formation of MOCO from MPT is a two-step process; (1) formation of MPT-AMP from MPT and (2) formation of MOCO from MPT-AMP. As there are two ligands MPT and MPT-AMP involved, geometry optimization was performed on both ligands, the reference of these ligands were taken from ZINC database[32], geometries were modeled using GaussView[31] and geometry optimization (energy minimization) was performed using Gaussian 09 program.[33] DFT calculations were performed with optimize model geometries of ligand. The optimized geometries of the two ligands can be seen in Figure 4.1 and 4.2 for MPT and MPT-AMP respectively.

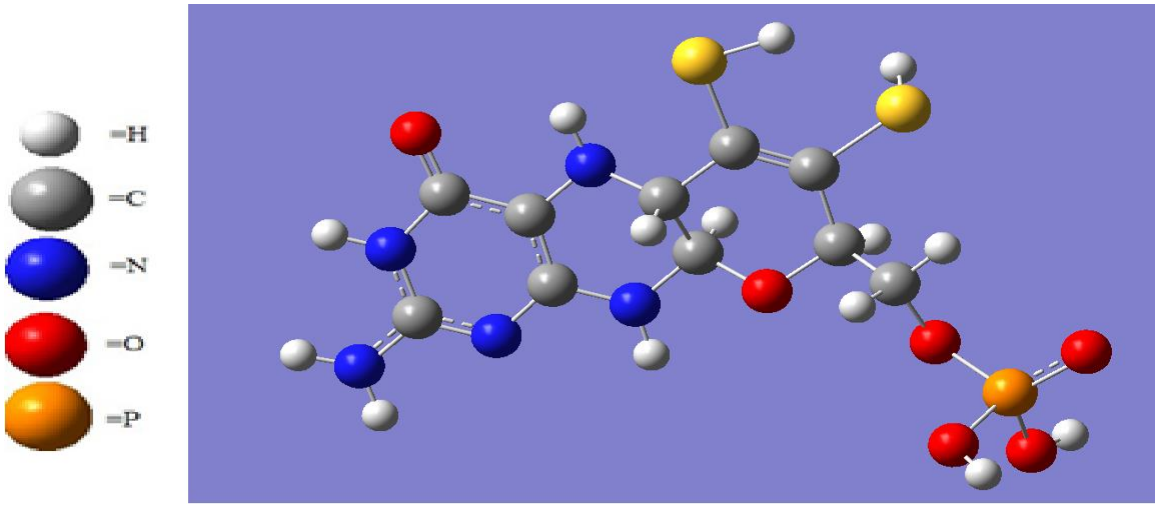


Figure 4.1: Final conformation achieved for MPT *via* DFT geometric optimization

It is to be noted that during energy minimization the overall charge of 'zero' and multiplicity was kept at 'one'. In order to keep charge at zero the sulfur atoms in the dithiolene group have hydrogen atoms which were removed prior to docking so that the sulfur atoms could interact with the copper Cu ion. It was required to achieve the first objective of this research i.e., calculate the interaction energy of metal with the sulfur atoms of dithiolene.

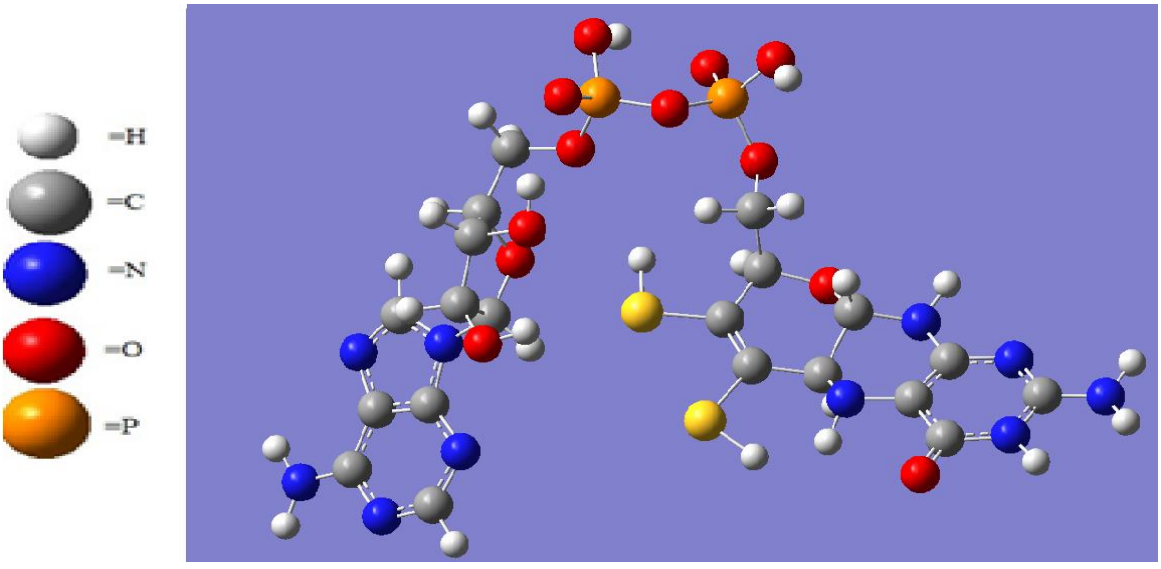


Figure 4.2: Final conformation achieved for MPT-AMP *via* DFT geometric optimization

In case of MPT-AMP the overall charge was also at kept at 'zero' and multiplicity is 'one'. In the case of MPT-AMP no alterations were made to the ligand prior to docking.

4.2. Active Sites / Binding Pocket Search

As mentioned in chapter 3 DoGsite Scorer was used to check whether the active site search predicts a binding pocket in the vicinity previously reported in literature and the drug score of the binding pocket. As mentioned earlier the plant homolog of Geph-G known as Cnx1 has a predicted binding site in the cleft between the functionally important b3 and a4; comprising of the highly conserved glycine glycine threonine glycine (Gly-Gly-Thr-Gly) motif (residues 86-89), leucine proline glycine (Leu-Pro-Gly) motif (144-146) and aspartate residues (Asp61 and Asp94).[9]

In Geph-G the DoGsite Scorer found a similar pocket reported in literature for the plant homolog with the the highly conserved Gly-Gly-Thr-Gly motif, Leu-Pro-Gly motif and aspartate residues present (Figure 4.3). The Pocket predicted by the software had the following residues:

VAL10 SER11 ASP12 SER13 ASP49 GLY74 GLY75 THR76 GLY77 PHE78 ALA79 ARG81 ASP82 SER107 VAL110 THR111 LEU113 GLY114 LEU116 SER117 LEU132 PRO133 GLY134 SER135 LYS137 GLY138 GLU141 CYS142

The pocket had the drug score of 0.81 (on a scale of 0 – 1) which is a clear indication that this was in fact the best possible binding site. The binding pocket has a depth of 21.4 Å, surface area of 836.1 Å^2 and volume of 863.2 Å^3. (Figure 4.3)

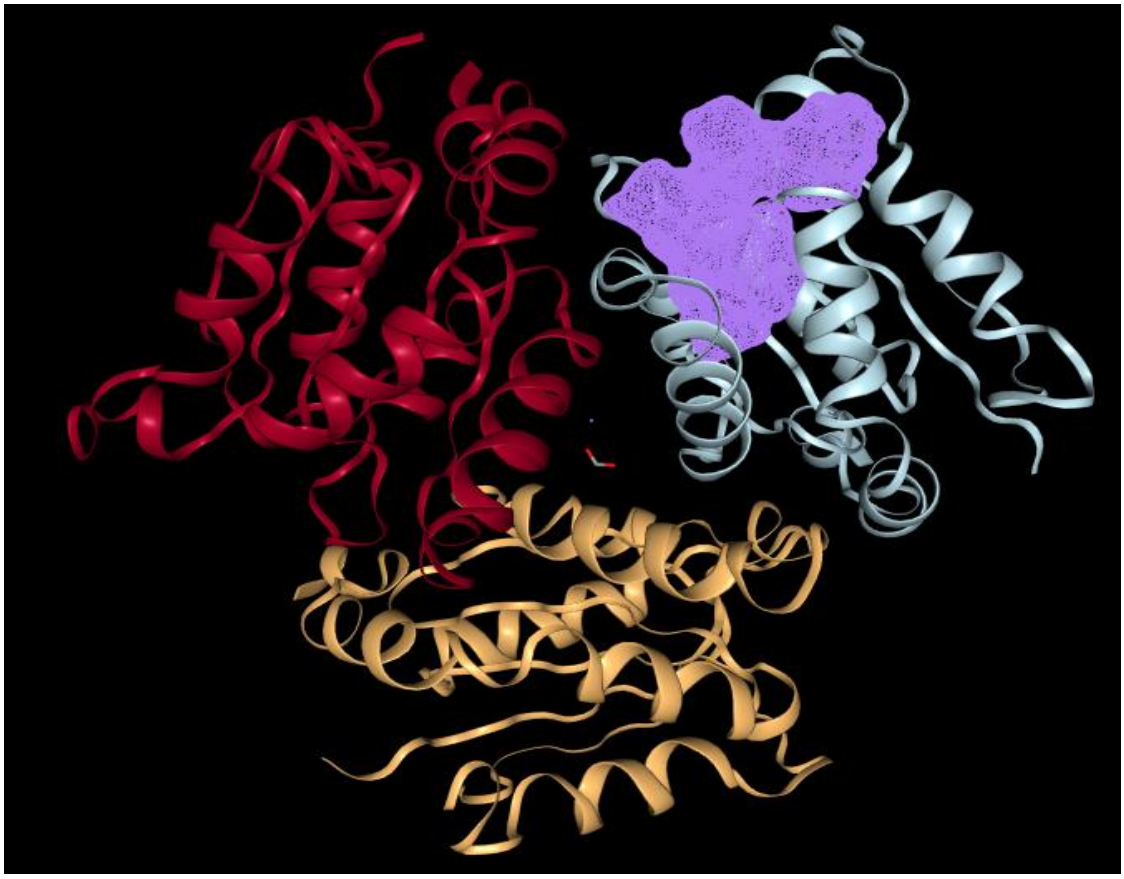


Figure 4.3: Highlighted cavity of Geph-G Trimer (colored chains) predicted by the DoGsite Scorer as the binding pocket for the active site with a drug score of 0.81[43]

In Geph-E the only report residue is Ser 630 that plays a vital role in the insertion of molybdate into the dithiolene group of MPT replacing Cu in a metal exchange reaction, as it reportedly interacts with the molybdenum cluster according to literature.[42] In Geph-E the DoGsite Scorer found a pocket with Ser 630 present with a drug score of 0.82 (on a scale of 0 – 1); (Figure 4.4) the residues present in this pocket are as follows:

LEU365 PRO366 PHE368 ALA370 SER371 VAL372 LYS373 THR413 GLY414 ALA415 ARG458 ASP463 GLU509 ILE522 GLY572 GLY573 VAL574 MET601 LYS602 PRO603 GLY604 LEU605 PRO606 PRO625 GLY626 ASN627 VAL629 SER630 ARG670 GLU672 HIS674 GLY693 LEU700

The drug score of 0.82 is a clear indication that this was in fact the best possible binding site. The binding pocket has a depth of 30.2 Å, surface area of 2894.6 Å^2 and volume of 2364.1 Å^3. (Figure 4.4)

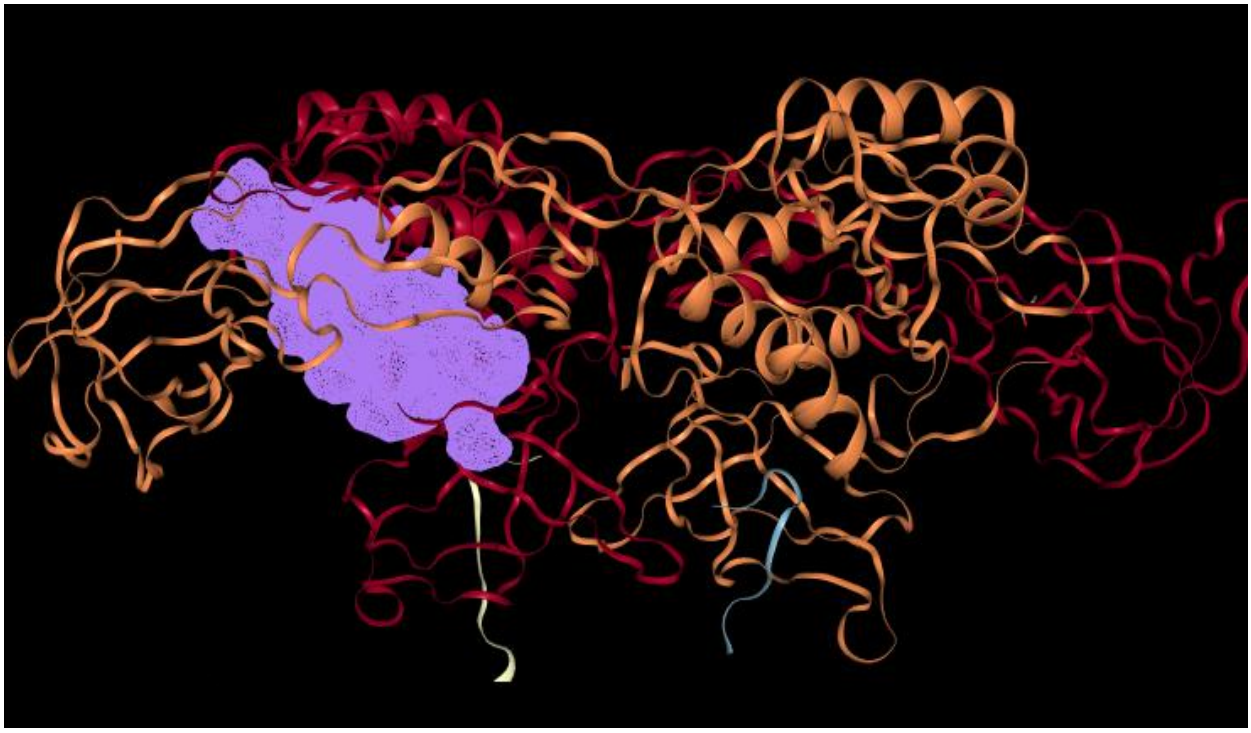


Figure 4.4: Highlighted cavity of Geph-E dimer (colored chains) predicted by the DoGsite Scorer as the binding pocket for the active site with a drug score of 0.82[43]

4.3. Molecular Docking / Binding Energies

A total of three docking jobs were prerformed; both ligands were docked into their respective protein domain's orthosteric binding site, namely MPT into Geph-G, MPT-AMP into Geph-E and for the final docking job Mo-MPT (MOCO) was docked into Geph-E in order to determain ligand binding poses and the binding energies for those poses, with the help of Molecular Operating Environment (MOE) software.

As stated earlier, literature reports that the binding pocket for MPT is highly conserved for the three homologous proteins; data from site-directed mutagenesis studies show that in the binding site of Cnx1 (*in vivo*), MPT is predicted to bind into the cleft between the functionally

important b3 and a4; comprising of the highly conserved Gly-Gly-Thr-Gly motif (residues 86-89), Leu-Pro-Gly motif (144-146) and aspartate residues (Asp61 and Asp94).[9] The bend between the pterin and the pyran ring compliments the shape of the binding pocket. It can be seen from the Figure 4.5 that the docking results are in agreement with the previously reported literature. However, it is to be noted that in the findings below, the motifs in the binding pocket of Geph-G (*in silico*) are displaced by 12 residues. The results depicted the Gly-Gly-Thr-Gly motif at residues 74 -77, Leu-Pro-Gly motif at 132-134 and aspartate residues at Asp49 and Asp82 respectively.

The dithiolene group of MPT, which is involved in the metal exchange reaction, is showing interactions with the Cu ion in the binding pocket and can be seen almost forming a furanose ring (see Figure 4.6) with the ion. The pterin moiety is seen interacting with the aspartate Asp82 residue while being in close proximity to the dithiolene group. The phosphate group is seen interacting with serine Ser135. The glysine Gly134 of the Leu-Pro-Gly motif at 132-134 is also seen interacting with the pterin moiety of MPT.

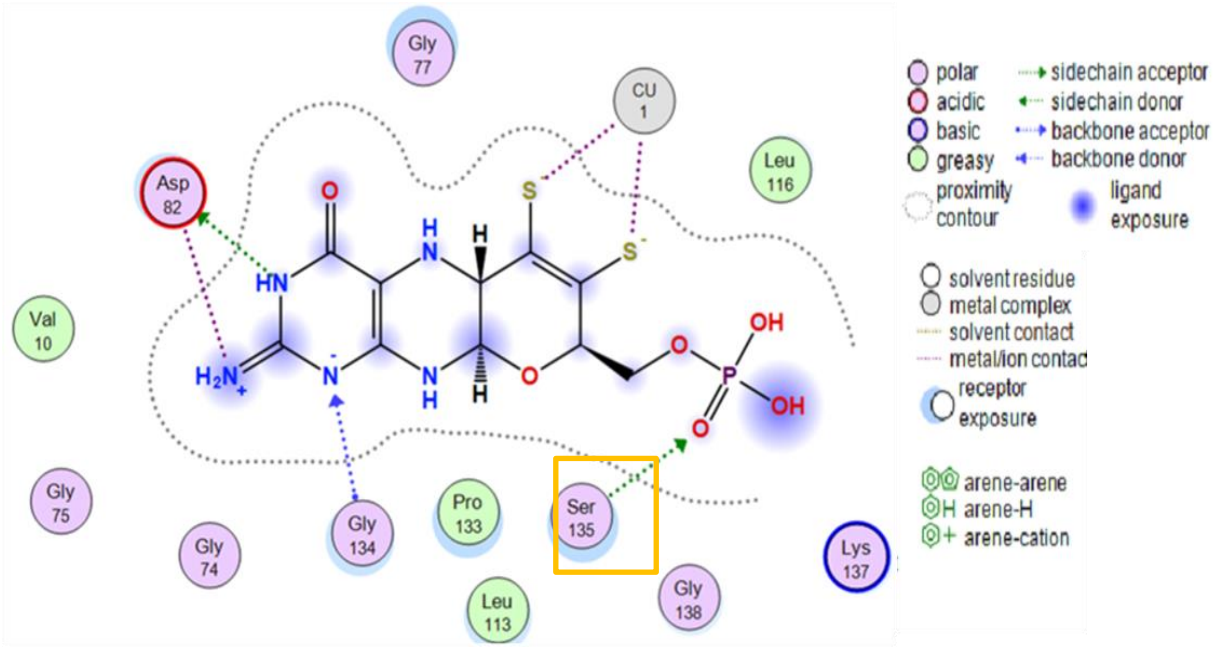


Figure 4.5: 2D ligand interactions of MPT docked with Geph-G domain of Mo *Insertase*, in the presence of Cu

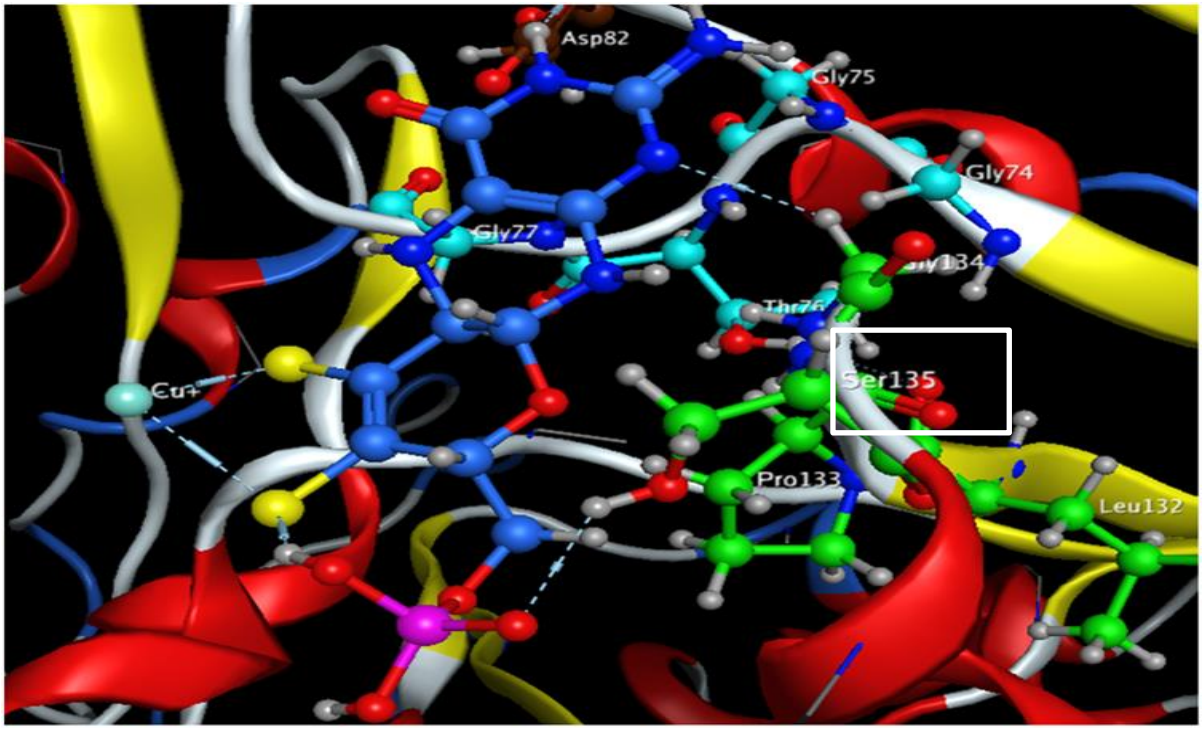


Figure 4.6: 3D ligand interactions of MPT docked with Geph-G domain of Mo *Insertase*

During the docking job for MPT into Geph-G orthosteric binding site, the placement method chosen was 'Alpha Triangle' with post placement refinement kept at 'induced fit'. The initial scoring function 'London dG' and the final scoring function 'GBVI / WSA dG' retained a total of 134 poses out of 150 generated as the early termination cutoff was defined at 3 identical poses. The correlation plot of the Generalized Born/ volume integral 'GB/VI' binding free energy (ΔG˚) and root mean square deviation 'RMSD' is depicted in Figure 4.7 for all 134 generated poses. Highlighted is the ligand binding pose 107 with binding free energy of -5.1 kcal/mol. (The ligand interactions can be seen in Figure 4.5 and 4.6 respectively.)

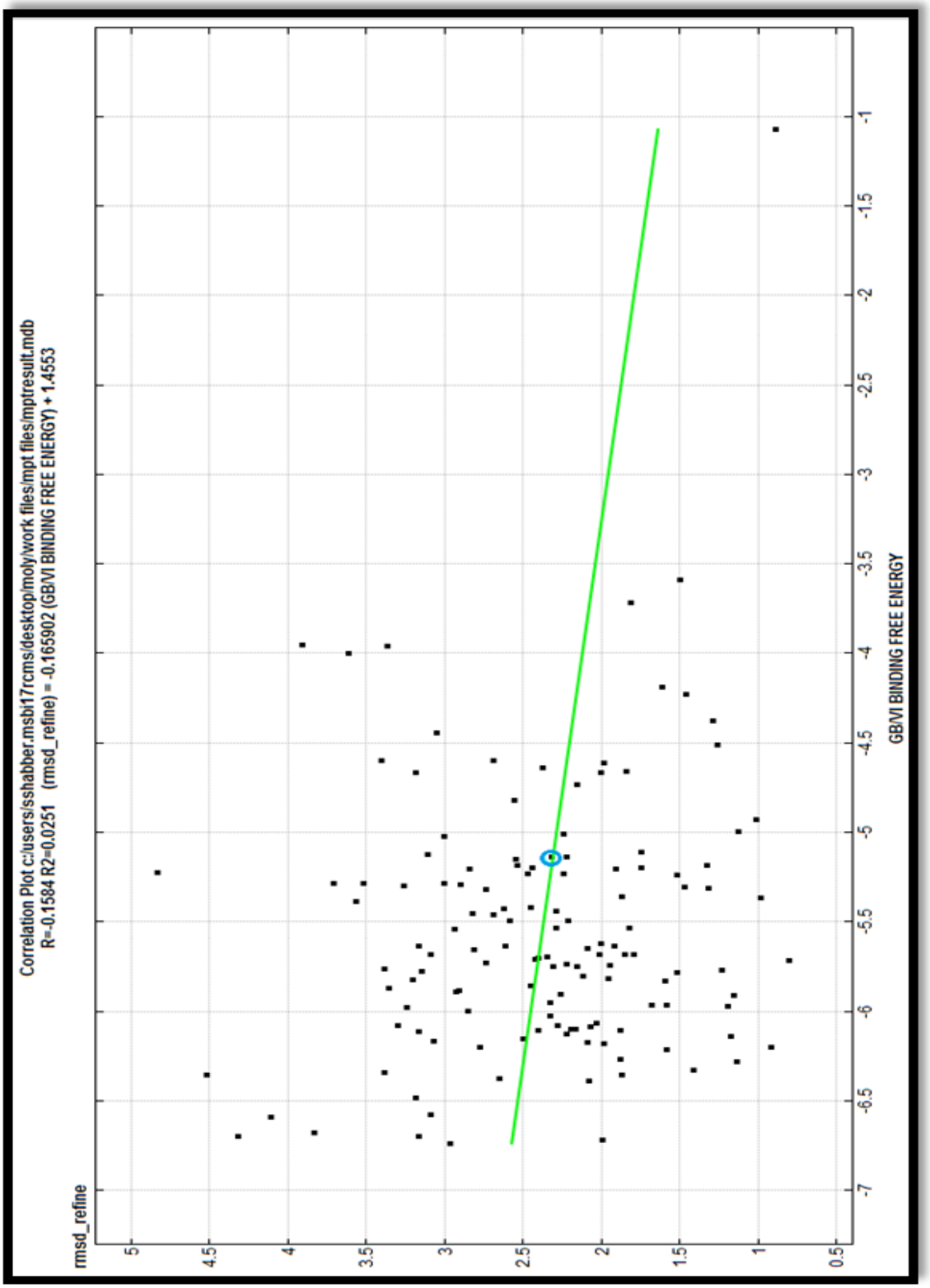


Figure 4.7: Correlation plot of GB/VI Binding Free Energy and RMSD for MPT/Geph-G

Next, the binding energy and interaction patterns seen for MPT-AMP with Geph-E are previously reported to have the Ser630 as having an important role in the in cooperation of molybadate into the final Molybdenum Cofactor. Upon being imported (from the low

resolution protein PDB ID: 5eru) with the help of UCSF Chimera, into the binding site of the Geph-E to check interactions, the molybdenum cluster and molybdate showed interactions with the reported serine Ser360 residue. The molybdenum cluster and molybadate were however removed prior to the docking of MPT-AMP in order to prevent the binding pocket from over cluttering.

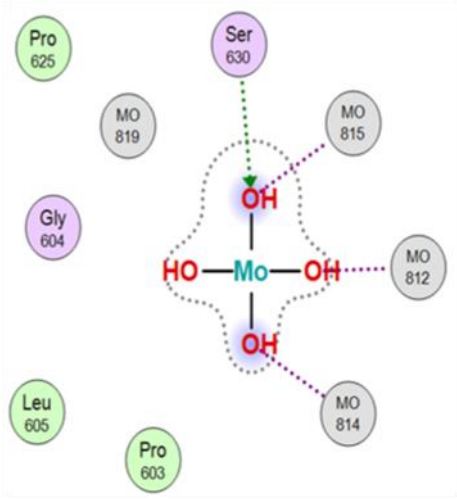


Figure 4.8: 2D depiction of molybadate / Mo cluster interaction with Ser 630 in Geph-E

Docking was preformed of MPT-AMP into Geph-E orthosteric binding site, for which the placement method chosen was 'Alpha Triangle' with post placement refinement kept at 'induced fit'. The initial scoring function 'London dG' and the final scoring function 'GBVI / WSA dG' retained a total of 100 poses out of 150 generated as the early termination cutoff was defined at 3 identical poses.

The only previously reported residues known to interact with MPT-AMP were glutamic acid Glu509 and arginine Arg670. From Figure 4.9 it is evident that this study succeeded in validating those findings as Glu509 can be seen as interacting with the ligand as a side chain acceptor. Furthermore, Aspartate Asp 463 is also seen as a side chain acceptor, whereas leucine Leu605 can be seen interacting with the dithiolene group as a backbone donor. However, interestingly it appears to be that the most important residue in this whole scenario is yet again

the Ser630, as it can be seen right in front of the sulphur part of the dithiolene group further validating the speculations of this residues involvement in the insertion of molybadate into MPTAMP. As previously discussed Ser630 is known to play an important part in the insertion of molybadate into the dithiolene group for the formation of final MOCO. It can be seen in Figure 4.9 and 4.10 that the Ser630 residue is in fact in the right vicinity in order to carry out this reaction.[42]

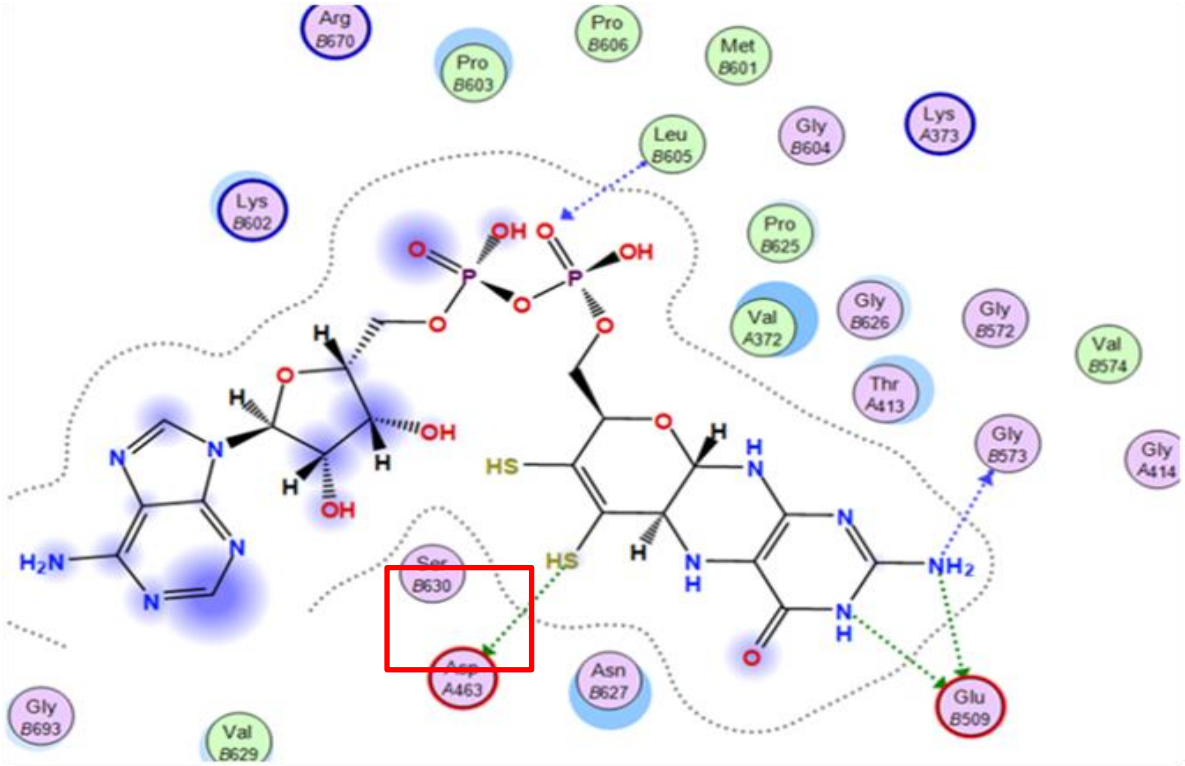


Figure 4.9: 2D ligand interactions of MPT-AMP docked with Geph-E domain of Mo *Insertase*

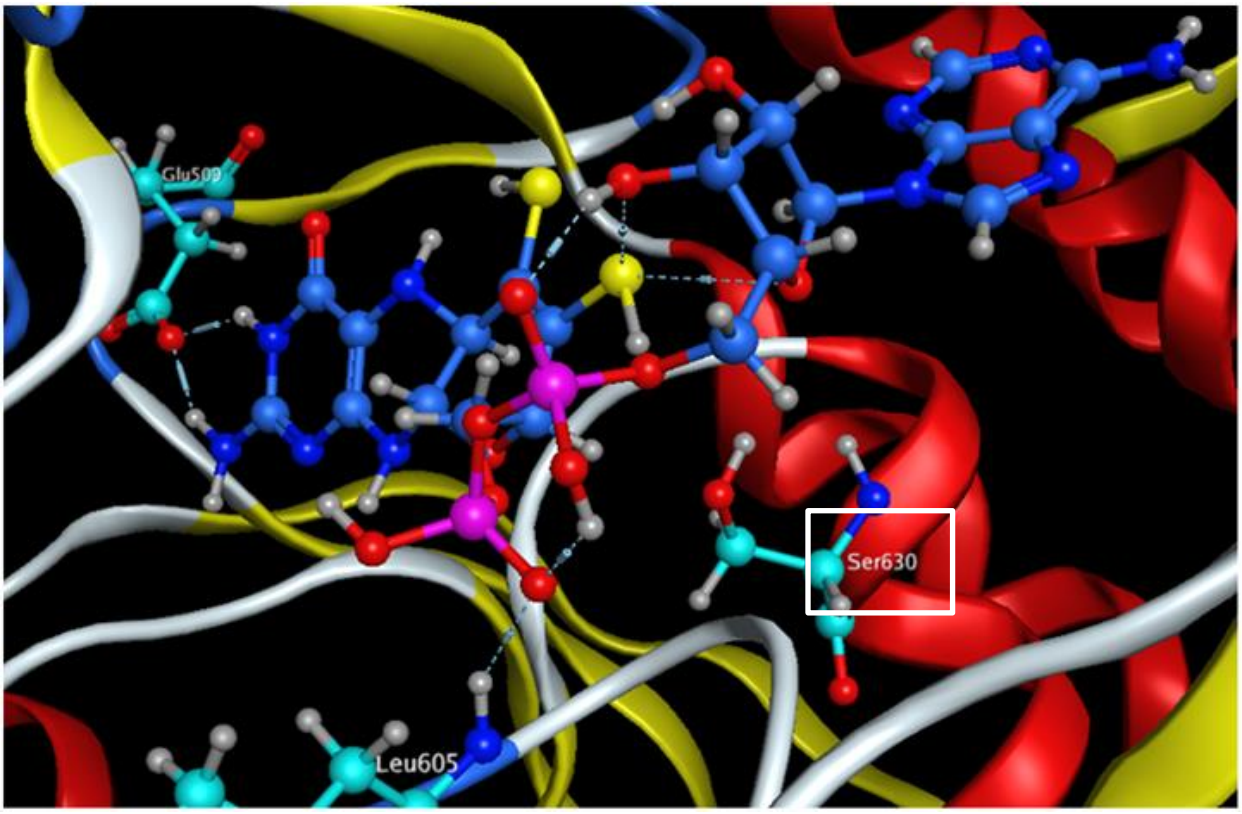
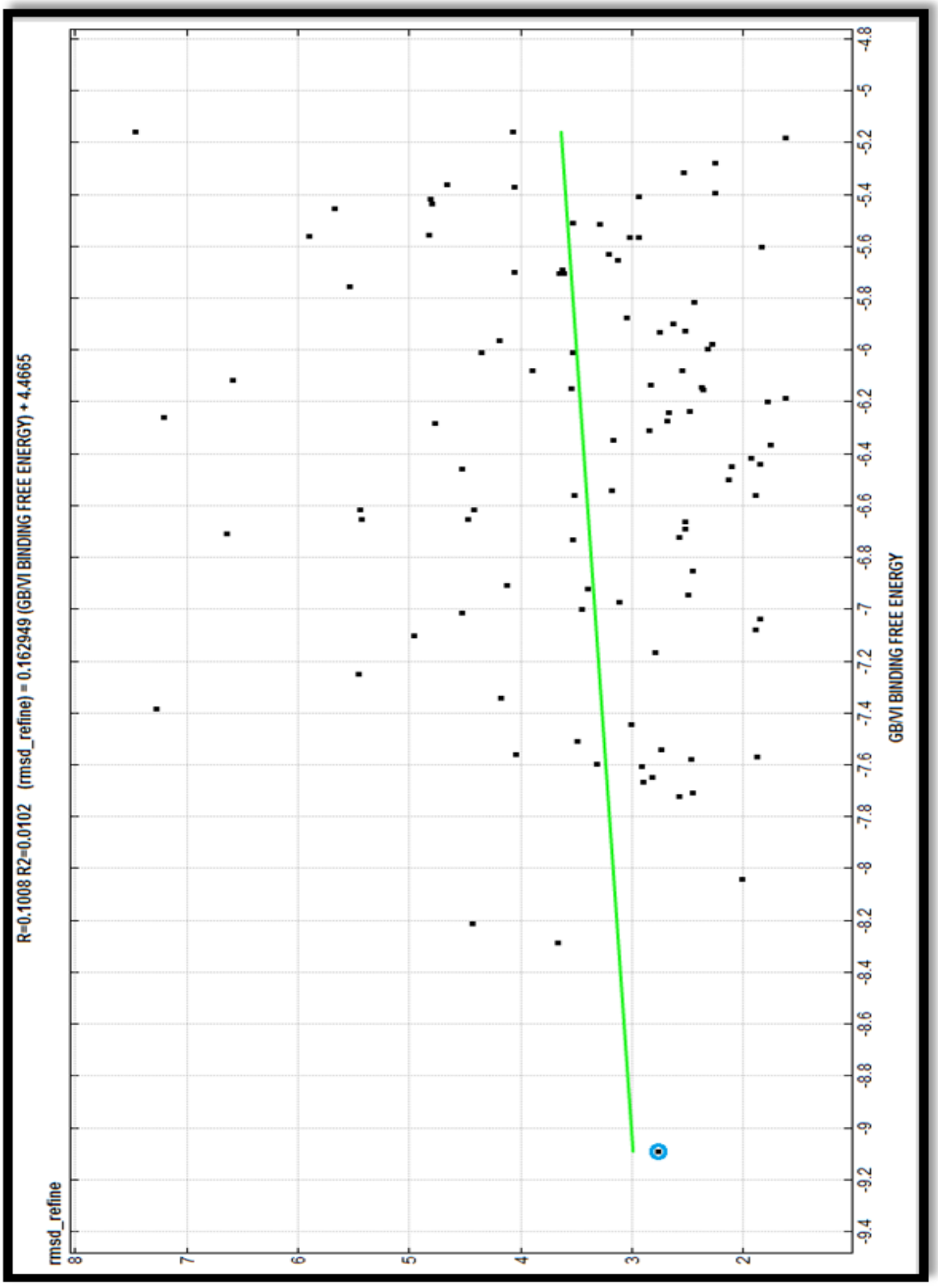


Figure 4.10: 3D ligand interaction of MPT-AMP docked with Geph-E domain of Mo *Insertase*

The correlation plot of the Generalized Born/ volume integral 'GB/VI' binding free energy (ΔG˚) and RMSD is depicted in Figure 4.11 for all 100 generated poses. The first ligand binding pose is with binding free energy of -9.1 kcal/mol (the ligand interactions of which can be seen in Figure 4.9 and 4.10 respectively).

Figure 4.11: Correlation plot of GB/VI Binding Free Energy and RMSD for MPT-AMP/Geph-E



A final docking procedure was performed for additional validation of the insertion procedure of molybdenum in the form of molydate (MoO_4^{2-}) into the MPT to form the final Mo-MPT (MOCO) molybdenum cofactor. This can be seen in Figure 4.12 and 4.13. The Aspartate Asp 463, glutamic acid Glu672 and Asparagine Asp 463 are seen interacting with the ligand with binding free energy of -7.9 kcal/mol.

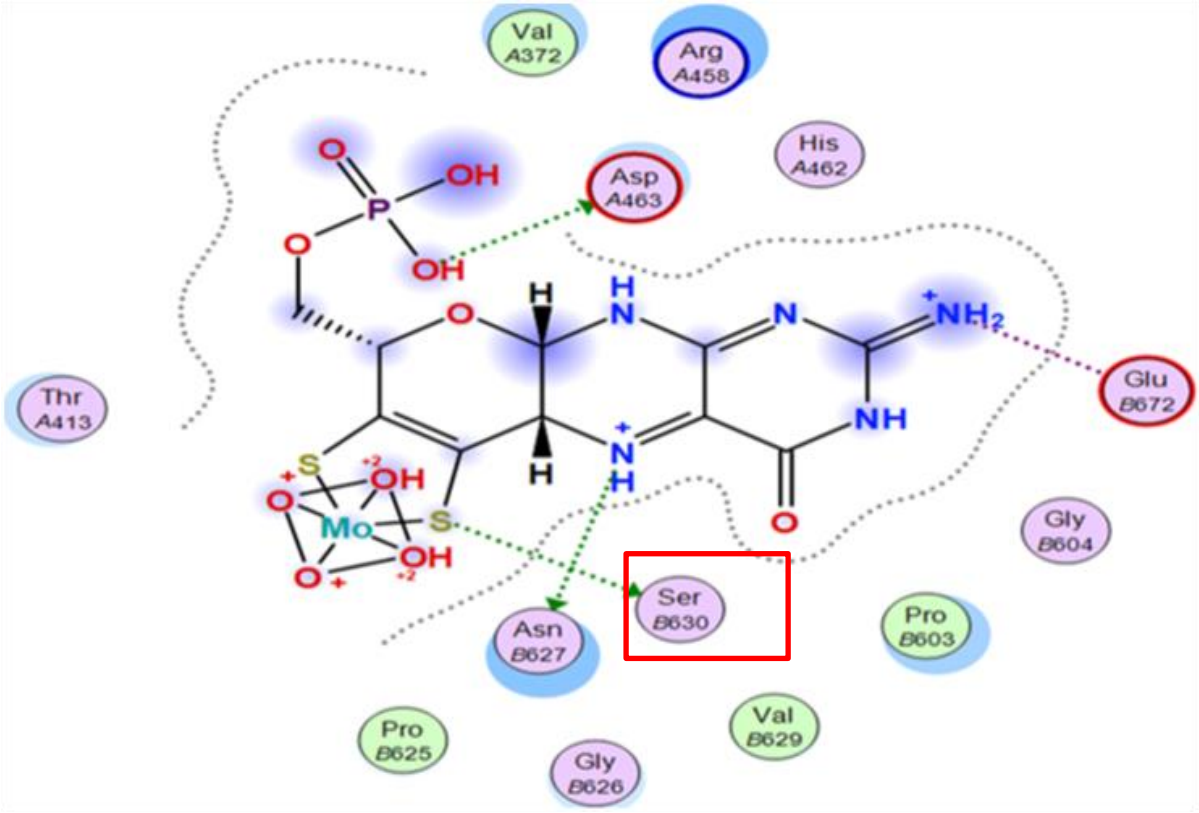


Figure 4.12: 2D ligand interactions of MOCO (Mo-MPT or MPT with attached molybadate) docked with Geph-E domain of Mo *Insertase*

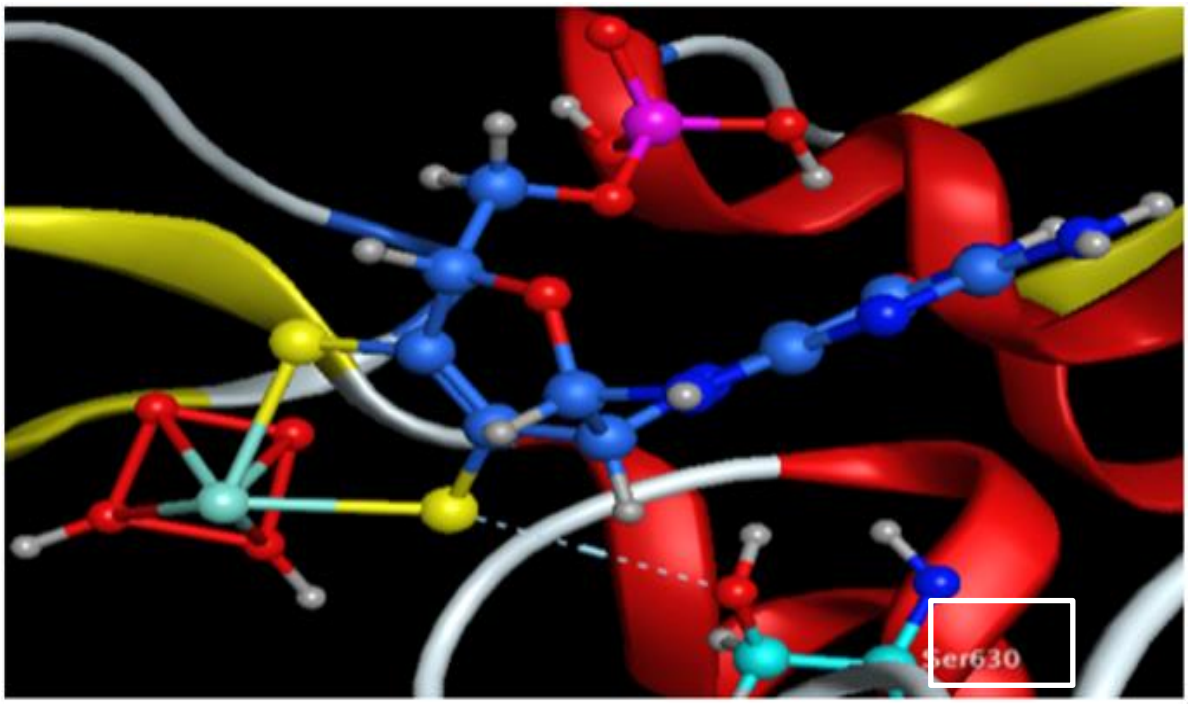


Figure 4.13: 3D ligand interactions of MPT with attached molybadate docked with Geph-E domain of Mo *Insertase*

Although depicted in the diagrams is an attached molybdate to the MPT once the molybdenum is inserted into MPT only two oxygen molecules remain bonded to the molybdenum; in post modifications either a hydroxy or cysteine group can be attached to the molybdenum depending on the enzyme it's inserted in; the RMSD of Mo-MPT/Geph-E is in positive correlation with Gibbs free energy (ΔG°) once again. This shows that necessary fluctuation of energy required in terms of how stable MPT becomes once adenylated (going from negative correlation just for MPT to positive correlation) and then consequently the positivity of the correlation maintain itself even after the insertion of molybdate. Note that the negative correlation of RMSD and Gibbs free energy (ΔG°) denotes an unfavorable reaction in nature while a positive correlation denotes a favorable reaction.

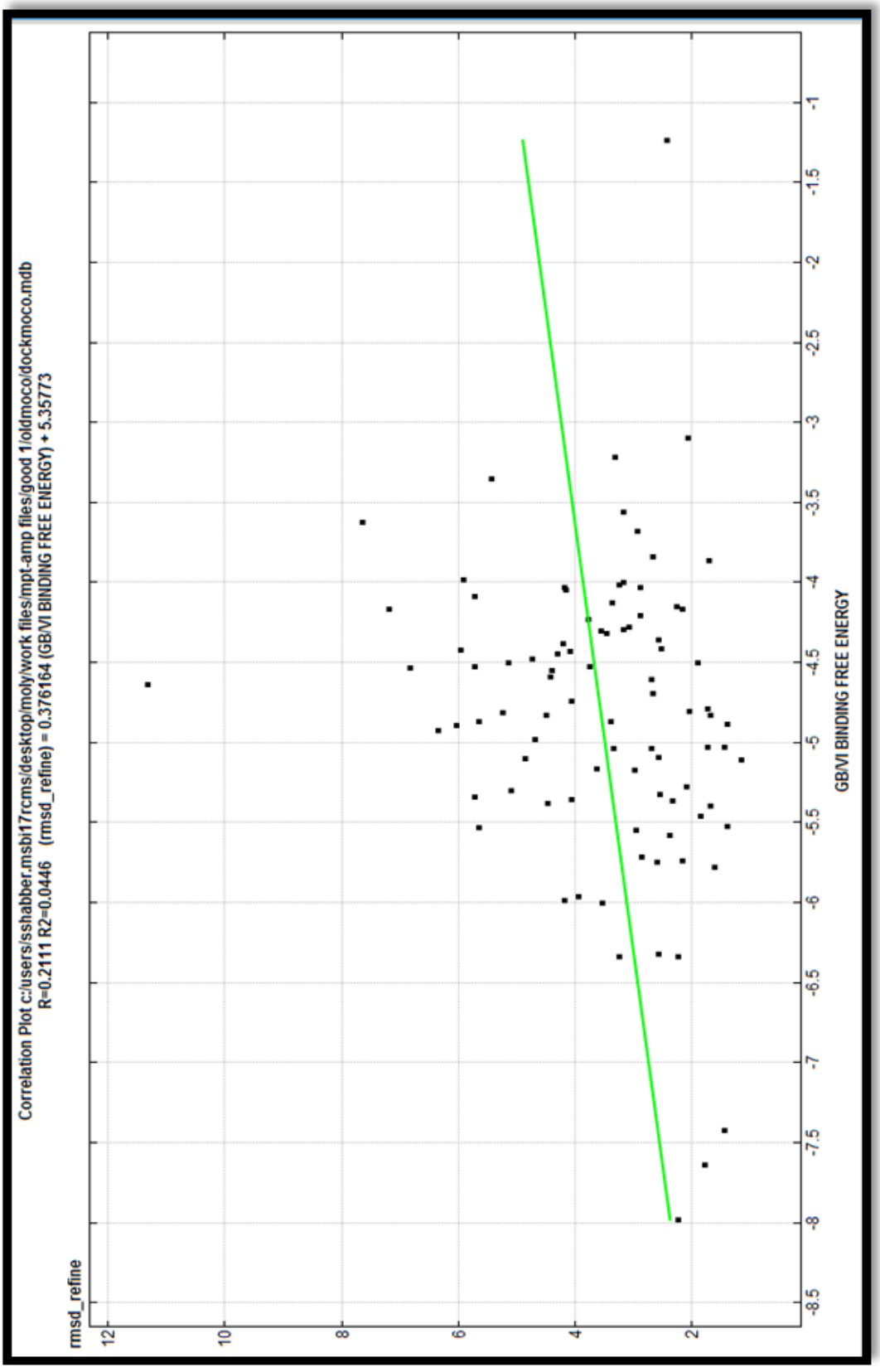


Figure 4.14: Correlation plot of GB/VI Binding Free Energy and RMSD for Mo-MPT/GephE

4.4.Molecular Dynamics/ Interaction Patterns

MD simulations were carried out by using protein-ligand complexes generated by molecular docking. The dynamic behaviors of the structures were examined; in order to clarify the dynamic stability of these structures, RMSD of the trajectories were obtained using VMD[53]. The RMSD was calculated per frame; in the case of MPT docked with Geph-G the MD simulation yielded 93 frames for 10 ns where as in the case of MPT-AMP docked with Geph-E 67 frames for 10 ns as can be seen in figure 4.15 and 4.16 respectively.

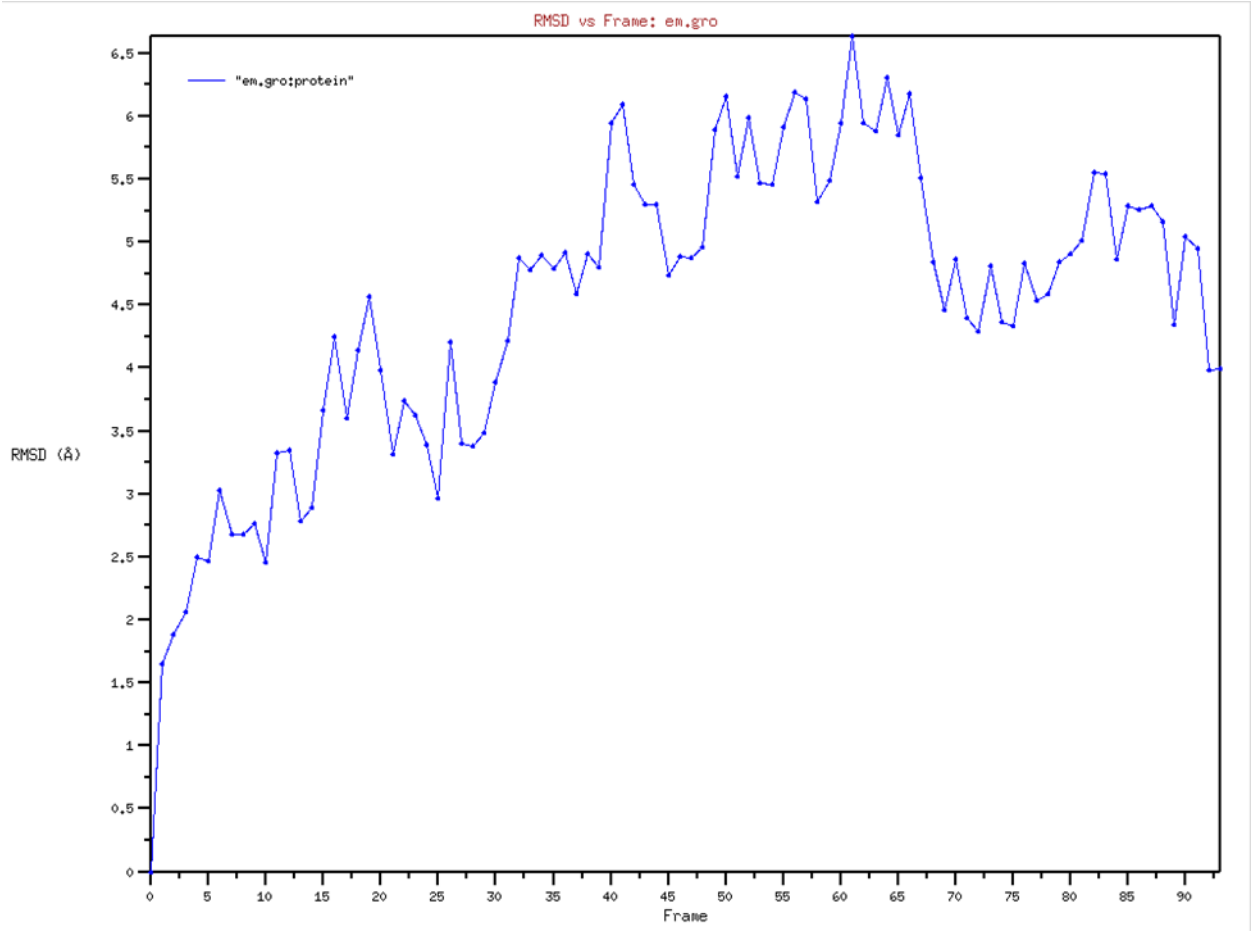


Figure 4.15: RMSD plot of MPT/Geph-G for 93 frames in 10 ns

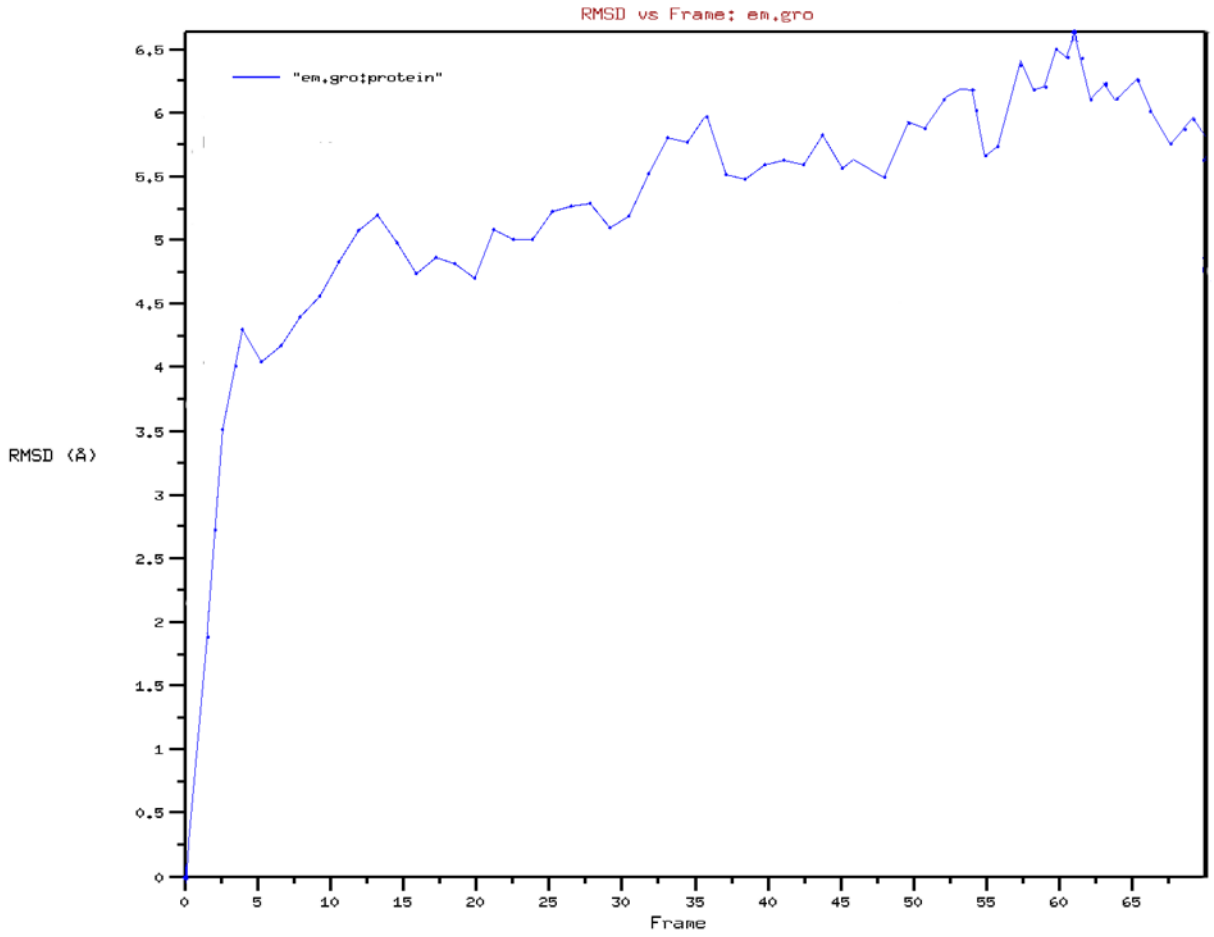


Figure 4.16: RMSD plot of MPTAMP/GephE for 67 frames in 10

The RMSD plot showed that the RMSD of the protein-ligand complex with respect to the initial structures increased for 10 ns in both cases. In case of MPT with Geph-G the RMSD increased steadily while peaking at 63rd frame at 6.5 and then declining with the last frame at an RMSD of 4.0, whereas in the case of MPTAMP with Geph-E the RMSD is seen at a constant upraise and peaking at the 60th frame at 6.5 and then declining mildly with the last frame at an RMSD of 5.7 respectively. The trajectories remain relatively stable until the end of the simulation. Thus, the trajectories of the MD simulations of these structures seem reliable. The short range Lennard-Jones potential as well as coulomb's energy was also calculated for both complexes (Table 1 and 2)

Energy Type	Average Value (kJ/mol)	RMSD
Coul-SR: Protein-MPT	-187.145	35.01
LJ-SR: Protein-MPT	-82.677	14.05

Table 1: Short Range Lennard-Jones potential Energy & coulomb's energy for MPT/Geph-G

Energy Type	Average Value (kJ/mol)	RMSD
Coul-SR: Protein-MPTAMP	-193.989	22.91
LJ-SR: Protein-MPTAMP	-214.334	13.89

Table 2: Short Range Lennard-Jones potential Energy & coulomb's energy for MPTAMP/Geph-E

The negative values on average for both complexes show that the interaction patterns of the respective ligands with their proteins were quite favorable. Upon further investigation of the interactions taking place at the end of the MD simulations matched closely to that of the preexisting docking results obtained, further fortifying the validity of the obtained docked results.(see Figure 5.3 and 5.4)

The interactions reported for the last frame of the MD simulation for MPT with Geph-G are depicted in the following Figures 4.19 and 4.20 respectively. MPT is seen interacting with Ser135 and Asp82. The Gly-Gly-Thr-Gly motif can clearly be seen in the binding pocket. The results are in sync with the docking results which are in tune with the literature.

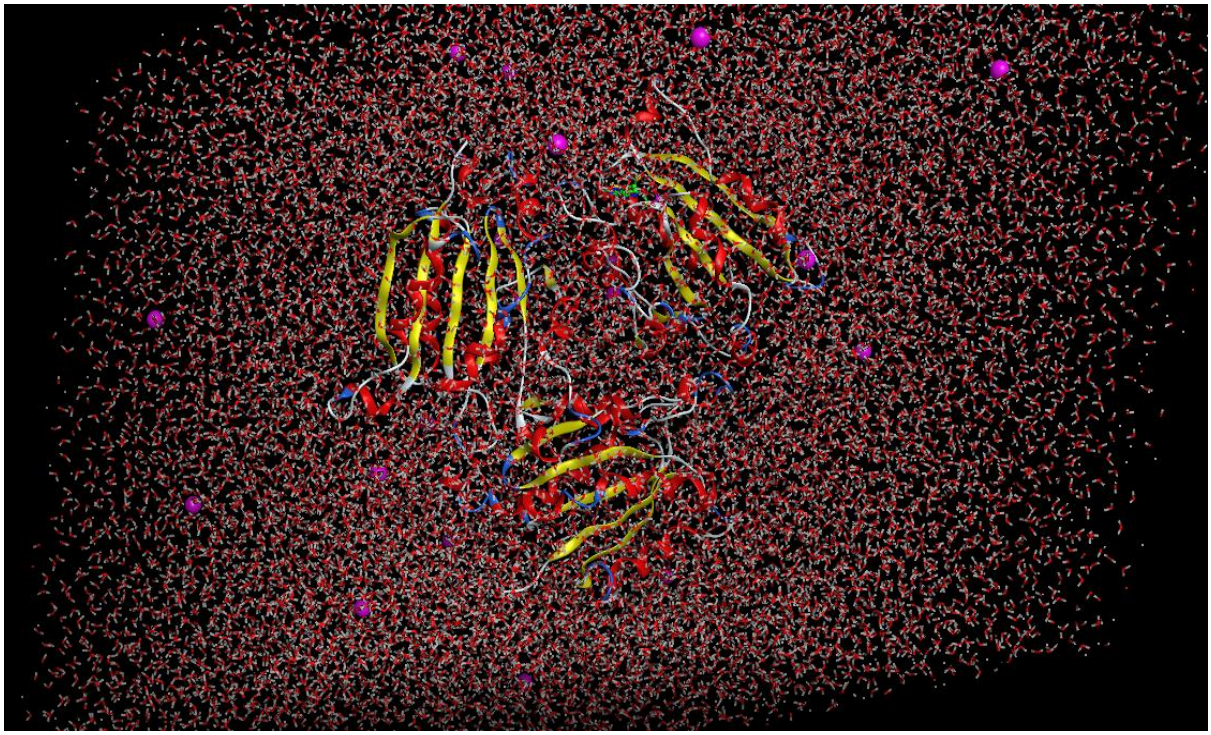


Figure 4.17: MD simulation of Geph-G - finite sized simulation box with equilibrated conditions

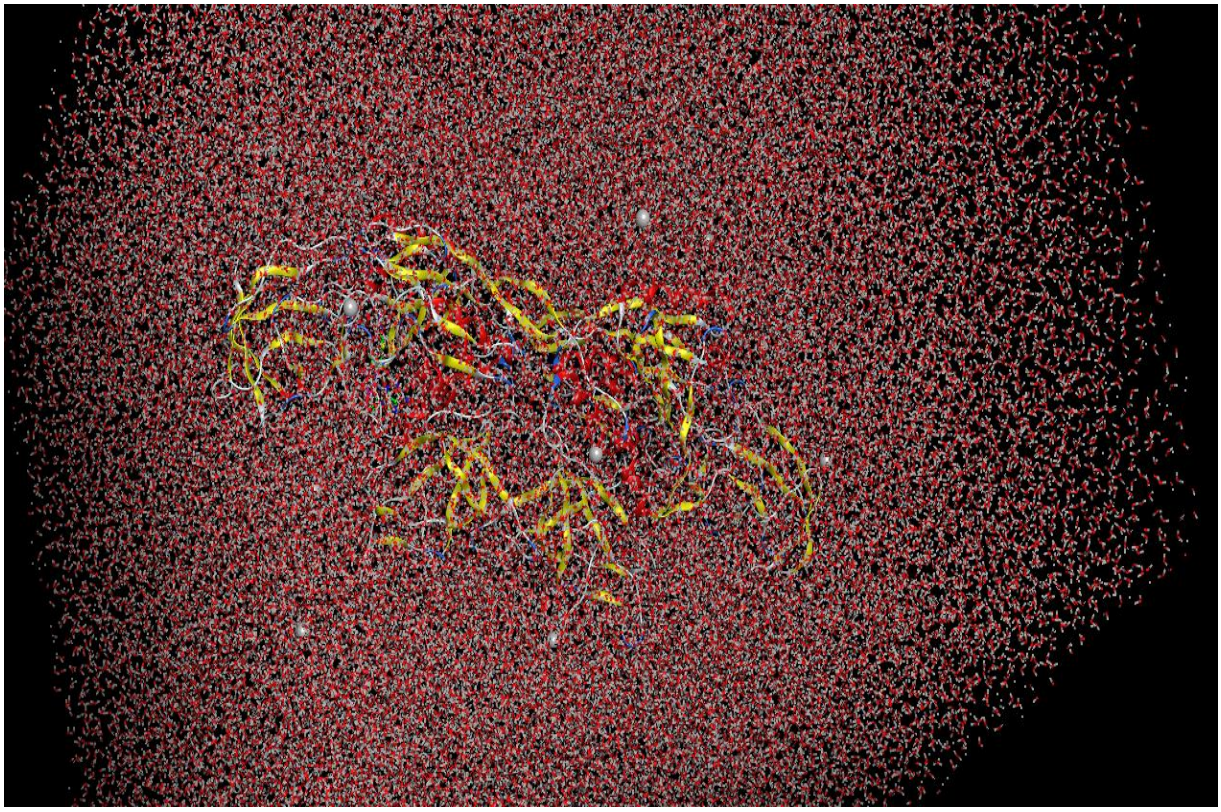


Figure 4.18: MD simulation of Geph-E - finite sized simulation box with equilibrated conditions

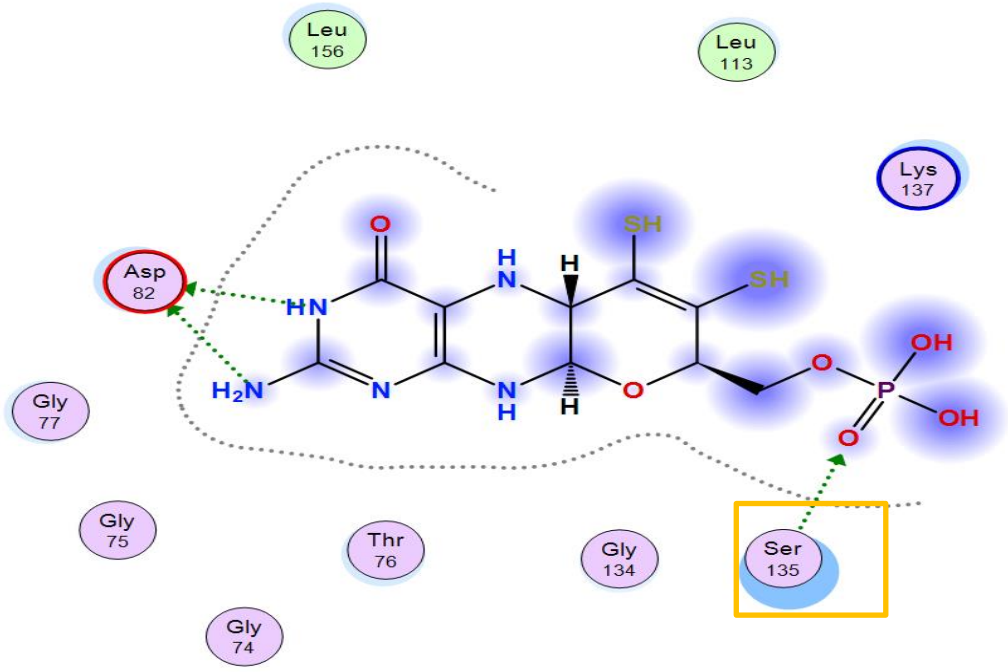


Figure 4.19: 2D ligand interactions of MD sim Last frame for MPT with Geph-G

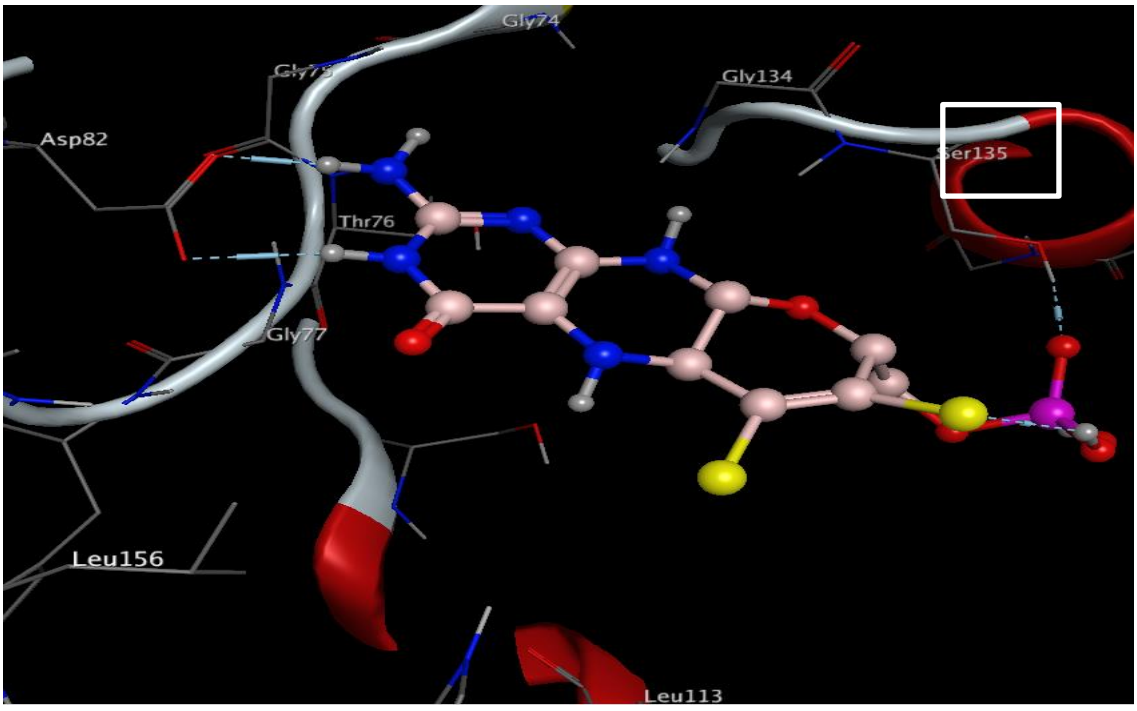


Figure 4.20: 3D ligand interactions of MD sim Last frame for MPT with Geph-G.

The interactions reported for the last frame of the MD simulation for MPTAMP with Geph-E are depicted in the following Figures 4.21 and 4.22 respectively. MPTAMP is seen within the same vicinity in the binding pocket as in the docked results. The ligand is seen interacting with residues Glu 509, Gly 604, Val 372 and Gly 626 while Ser 630 remains in the vicinity of the ditheoline group getting in position to aid molybadate insertion; it is clearer in the 3D depiction rather than the 2D depiction.

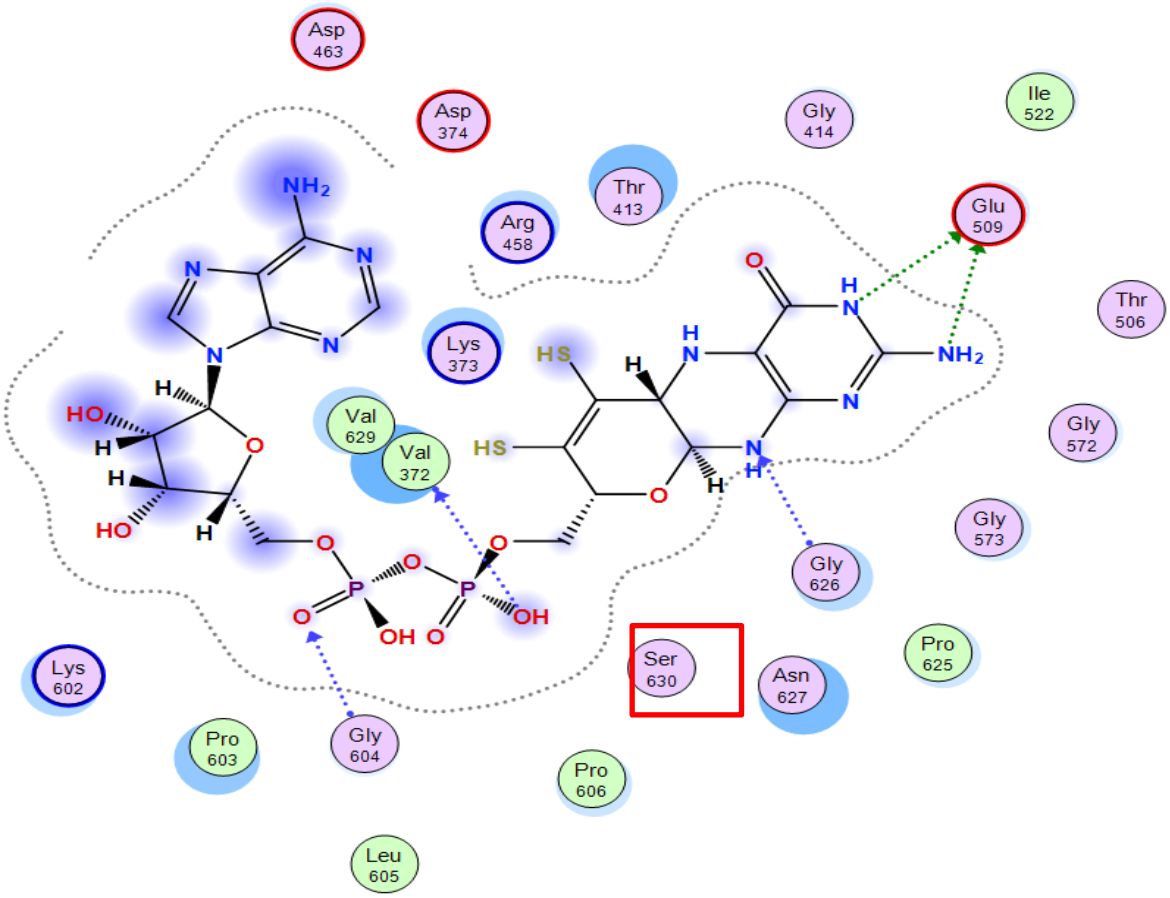


Figure 4.21: 2D ligand interactions of MD sim Last frame for MPTAMP with Geph-E

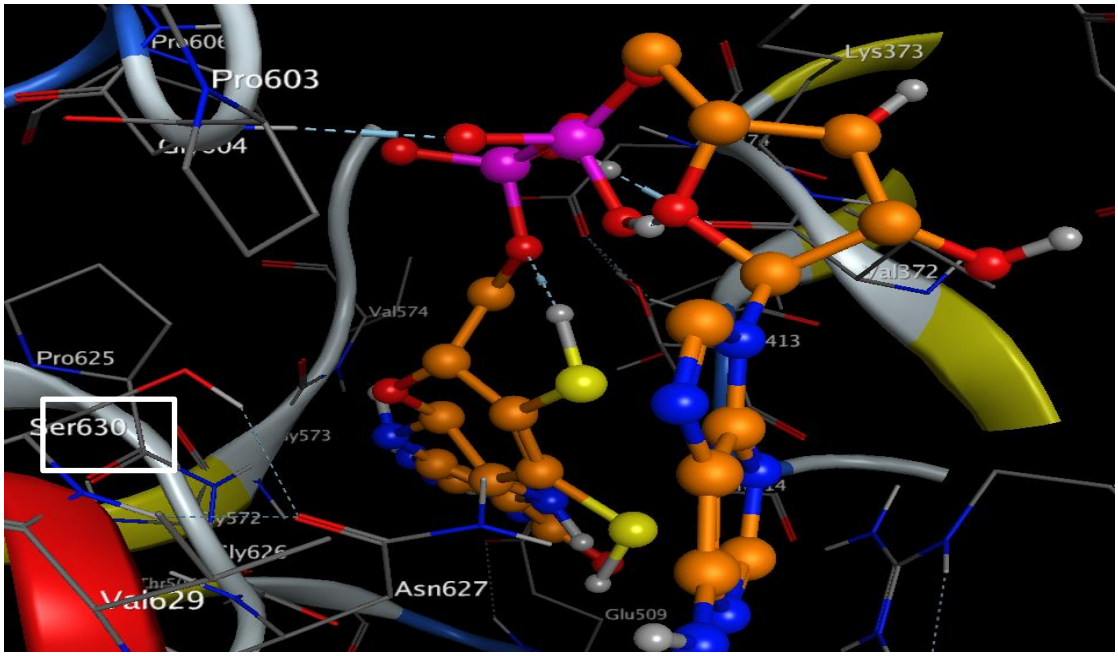


Figure 4.22: 3D ligand interactions of MD sim Last frame for MPTAMP with Geph-E

CHAPTER 5

DISCUSSION

Molybdenum cofactor (MOCO) is comprised of a diet derived molybdate (MoO_4^{-2}) ion and a dithiolate pterin moiety known as molybdopterin (MPT). This study was conducted keeping in mind (1) the functional role of gephyrin in MOCO biosynthesis (2) to understand the detailed reaction mechanism for synthesis of MOCO from MPT and its intimidator MPTAMP. Understanding the function of gephyrin is crucial in order to form an opinion on the exact metal exchange reaction which allows replacement of Cu with Molybadate.

In order to have a better understanding of reaction mechanism for synthesis of MOCO from MPT, binding energy and interaction patterns between Molybdenum *Insertase*: Geph-G and Geph-E domains with MPT and MPTAMP respectively, were calculated *via* molecular docking and molecular dynamic simulations.[54–56]

Docking was performed to calculate the Gibbs free binding energy (ΔG°) within the binding pockets of the two protein domains of Molybdenum *Insertase,* located using pocket prediction software (DoGSite Scorer). The purpose of MD after docking calculations was to confirm the poses found by docking. MD simulation was useful for confirming whether the major contacts found during docking will be maintained during the MD. In general, the combination of two *in-silico* techniques (Molecular docking, MD) improves the reliability of the results.[45]

For this study the consistency in results when it came to (1) the active site defined by the binding pocket prediction software to the (2) docking results and (3) the trajectories casted by the MD simulation has helped draw some conclusions about the ‘reaction mechanism’ of the last step of MOCO ‘metal exchange reaction’.

Drawing comparison from the results it can be seen in Figure 5.1 and 5.2, both MPT and MPTAMP got docked within the binding sites predicted for the DoGSite Scorer for Geph-G and Geph-E respectively. The DoGsite Scorer gave a drug score of 0.81 for Geph-G and 0.82 for Geph-E, which is a clear indication that these were in fact the best possible binding sites for the protein domains. This matters as it validates the binding pockets mentioned in the literature and those reported *via* experimental data.

Figure 5.1: Side by Side Comparison – Geph-G Docking Results VS DoGsite Scorer Prediction

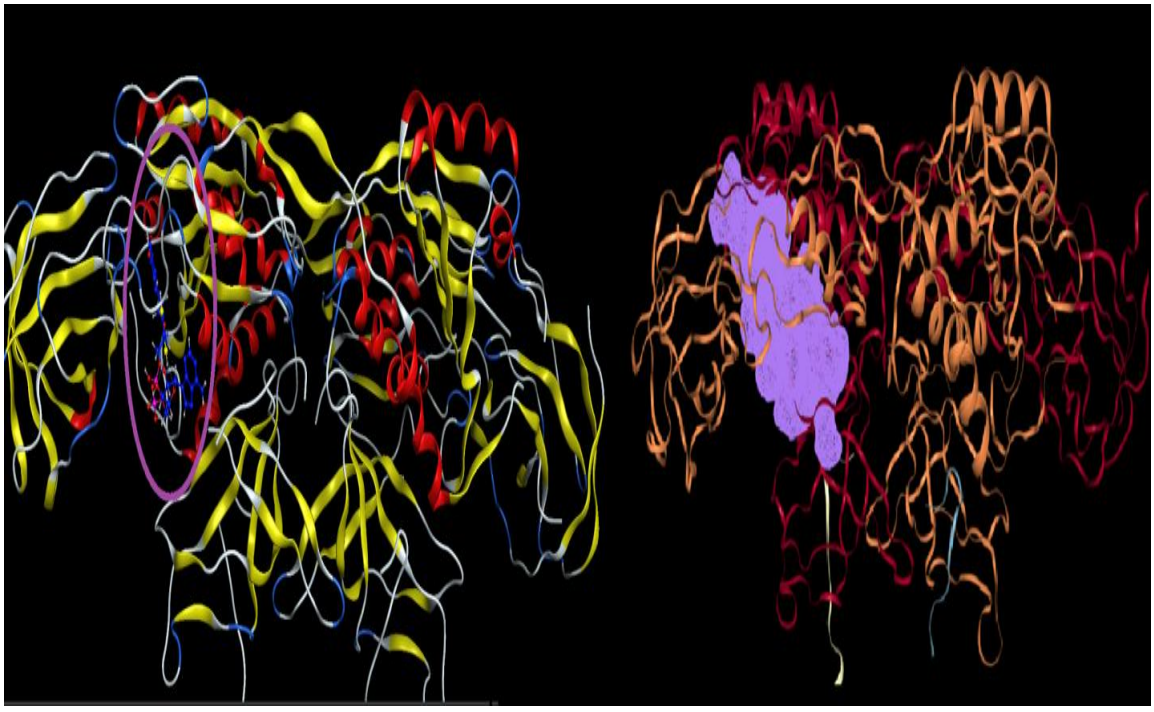


Figure 5.2: Side by Side Comparison – Geph-E Docking Results VS DoGsite Scorer Prediction

Secondly, results obtained from docking are in the form 'Gibbs free binding energy' (ΔG°), which is a binding energy between the ligand and the residues of the binding pocket. Binding energy is negative potential energy which happens whenever the interaction is attractive (like a positive charge with a negative charge). It is called "binding" because it represents the amount of energy required to completely free the two interacting objects from each other.[57,58] Furthermore the relationship between 'Gibbs free binding energy' (ΔG°) and the 'RMSD' of the ligand's binding pose is important. If both values are directly proportional, (meaning RMSD increases with the increase in the binding free energy) then that shows the ligand's binding pose is favorable at a lower energy state. If the two values are inversely proportional, (meaning RMSD decreases with the increase in the binding free energy) then that indicates that the respective ligand's binding pose is unfavorable at a lower energy state.

Observing the results of MPT being docked with Geph-G, it is evident from the plot (Figure 4.5) that there exsists an inverse relation between 'Gibbs free binding energy' (ΔG°) and the 'RMSD' of the ligand's binding pose and a higher energy state of MPT was required for favorable interactions with the Cu ion; the ligand interactions of which can be seen in Figure 4.5 and 4.6 respectively, as the binding pose of MPT ranked 107 by RMSD with binding free energy of -5.1 kcal/mol is seen interacting with the Cu ion (serine Ser 135).)

As a previously conducted research shows that two hydrogen atoms attached with two sulfurs at the enedithiolene part of MPT were replaced with Cu ion. Computational analysis was carried out using DFT, the results of which showed negligible changes when Cu replaces two hydrogen atoms attached with the sulfur atoms, i.e. only -0.01 kcal/mol energy differences was seen. This is a clear indication that MPT is a high energy molecule from the time of its formation and possibly high energy of this molecule is what aids incorporation of Cu ion into this molecule. The incorporation of Cu ion into MPT has negligible impact on the molecules energy indicating that Cu ion is in fact just a transient commodity meant only for a metal

exchange reaction and the docking results of this study back that up. Other examples of such metal exchange reactions, which naturally occur in the world, are Aluminum and magnesium Al/Mg exchange reaction in metalloenzymes, with the same surrounding ligand. [59]

Overall MPT is unstable with or without the copper ion hence depicting the role of Cu ion as just a commodity to be used for replacement. It is noteworthy that upon adenylation, MPT is stabilized making it more receptive of molydate even though the energy rises again (going from -9.1 to -7.9 kcal/mol) but it is evident that Molybdenum is far more stable than the Cu ion and MPT duo. Moreover, it makes sense that the transient Cu ion which is meant to be replaced by molybdenum is less stable and at a higher energy state making it easily replaceable. In case molybdenum is unable to replace Cu ion it can lead to MoCD type B; observed inhibition of MOCO synthesis by Cu ion suggests a link to copper transport disorders leading to MoCD Type B. In some studies the role of Cu ion is also to help save MPT from rapid oxidization.[232215]

It is seen that in contrast to MPT being docked in Geph-G, when MPT-AMP being docked in Geph-E, (Figure 4.9) the first ligand binding pose has the least RMSD with binding free energy of -9.1 kcal/mol showing a desired interaction within the binding pocket; moreover their relationship is directly proportional, meaning that MPTAMP showed more stable interactions in a lower energy state, possibly due to the adenylation process. Serine (Ser 630) is seen in the vicinity of the ditheoline group getting in position to aid molybadate insertion into the dithiolene group; this can be seen in Figure 4.9 and 4.10.

A final docking procedure was performed for additional validation of the insertion procedure of molybdenum (in the form of molydate) into the MPT to form the final MOCO cofactor (Mo-MPT) (Figure 4.12 and 4.13). The plot in Figure 4.14 shows RMSD of Mo-MPT binding pose with Geph-E is also in positive correlation with Gibbs free energy (ΔG°), just like in the case of for MPTAMP with Geph-E; further fortifying the fact that Cu ion coupled MPT

make an unstable duo, whereas molybdenum is far more stable in its place. The Ser630 residue plays an important role in the insertion of molybdenum into the cofactor; it can be seen interacting with the dithiolene group as a side chain acceptor. Aspartate (Asp463), glutamic acid (Glu672) and Asparagine (Asp463) are also seen interacting with the ligand with binding free energy of -7.9 kcal/mol.

Lastly, MD simulations were performed, which were useful for confirming whether the major contacts found during docking will be maintained during MD simulation giving insight on the pattern of interaction of the docked ligands with their respective protein, in order to produce more reliable results.

In contrast to docking which gave us binding poses as well as binding energy (Gibbs free binding energy (ΔG°)), MD simulations calculates the interaction pattern as well as interaction energies of ligands with the pocket residues called 'Short Range Lennard Jones' (LJ) (lowest-energy arrangement of an infinite number of atoms)[60] and Coulomb energy (the energy associated with the electrostatic forces of a system of particles, especially with that of the electrons of a covalent bond). (See Table 3 below.)

CALCULATED POTEINTIAL ENERGIES OF INTERACTION (kJ/mol)	MPT	MPTAMP
TOTAL ENERGY OF SYSTEM	-705491	-2497330
SR-LJ: LIGAND	-12.3	-48.2
SR-COUL: PROTEIN-LIGAND	-187.1	-193.9
SR-LJ: PROTEIN-LIGAND	-82.7	-214.3
RMSD-LJ: PROTEIN-LIGAND	14.0	13.9

Table 3: Summery of all the interaction energy terms calculated for MD simulation for both MPT and MPTAMP for their respective protein domains. In case MPT, SR-COUL is higher meaning more work was required for making this interaction possible, whereas SR-LJ is lower indicating a less favorable bond. In case of MPTAMP, SR-COUL is lower meaning less work was required for making this interaction possible, whereas SR-LJ is higher indicating a more favorable bond.

In order to make sense of the interaction energy values calculated it is necessery to be familiar with terms such as Short Range Lennard Jones (SR-LJ) and Short Range Coulomb Energy (SR-COUL) and the implications these energies have when positive or negative. (Short range = less than 1 Å)

For Lennard Jones potential consider any two molecules in a protein and they'll both have attractive and repulsive forces between them. The attractive forces are between dipoles on the molecules whereas the repulsive forces occur when the electron clouds of the two molecules are close enough to repel one another. These attractive and repulsive forces determine the potential energy of the two molecules. The total potential energy varies as the separating distance between the molecules changes. This total potential energy is called the Lennard Jones potential energy; attractive forces correspond to negative potential energy on the other hand repulsive forces correspond to positive potential energy (meaning the more negative Lennard Jones potential value is the stronger the bond is between the corresponding molecules). Hence

the tendency would be for the two molecules to move to a position where their potential energy is at a minimum (the most negative). [60–62]

Short Range Coulomb Energy (SR-COUL), is basically the electric potential energy of a system of point charges defined as the work required assembling this system of charges by bringing them close together (meaning how much energy was required to make this interaction happen).[57]

Looking at the interaction energy terms calculated for MD simulation for both MPT and MPTAMP for their respective protein domains. For MPT the Short Range Coulomb Energy (SR-COUL) for Geph-G and MPT is -187.1 kJ/mol and Short Range Lennard Jones (SR-LJ) is -82.7 kJ/mol. As observed Lennard Jones (SR-LJ) is far less than Coulomb Energy (SR-COUL) indicating that MPT is unstable and work is required to achieve this interaction, which is far greater than how strong a bond is present between the protein and ligand (greater Coulomb Energy than the Lennard Jones Potential). This also explains the requirement of ATP (see Figure 2.9) for the conversion of MPT to MPTAMP. The ATP not only provides the Adenine for adenylation to stabilize the MPT, it also provides the necessary energy required to make this interaction possible.

The Short Range Coulomb Energy (SR-COUL) for Geph-E and MPT-AMP is -193.9 kJ/mol and Short Range Lennard Jones (SR-LJ) is -214.3 kJ/mol. It can be observed that the Lennard Jones (SR-LJ) is far greater than Coulomb Energy (SR-COUL) indicating that MPTAMP has a favorable interaction and the greater negative value of Lennard Jones indicates that the interaction between MPT-AMP and Geph-E are stable.

Hence for this study performing MD after docking further fortified the poses found by docking. MD simulation proved to be useful for confirming whether the major contacts found during docking were correct in terms of energies being put out. For interaction patterns the

comparison between final poses suggested by MD simulation and docking it is quite evident that the results are consistent. Following is a comparison of the docking and MD simulation in Figure 5.3 and Figure 5.4 and it can be seen that nearly all functional groups for both ligands are only mildly displaced by the end of the MD simulation.

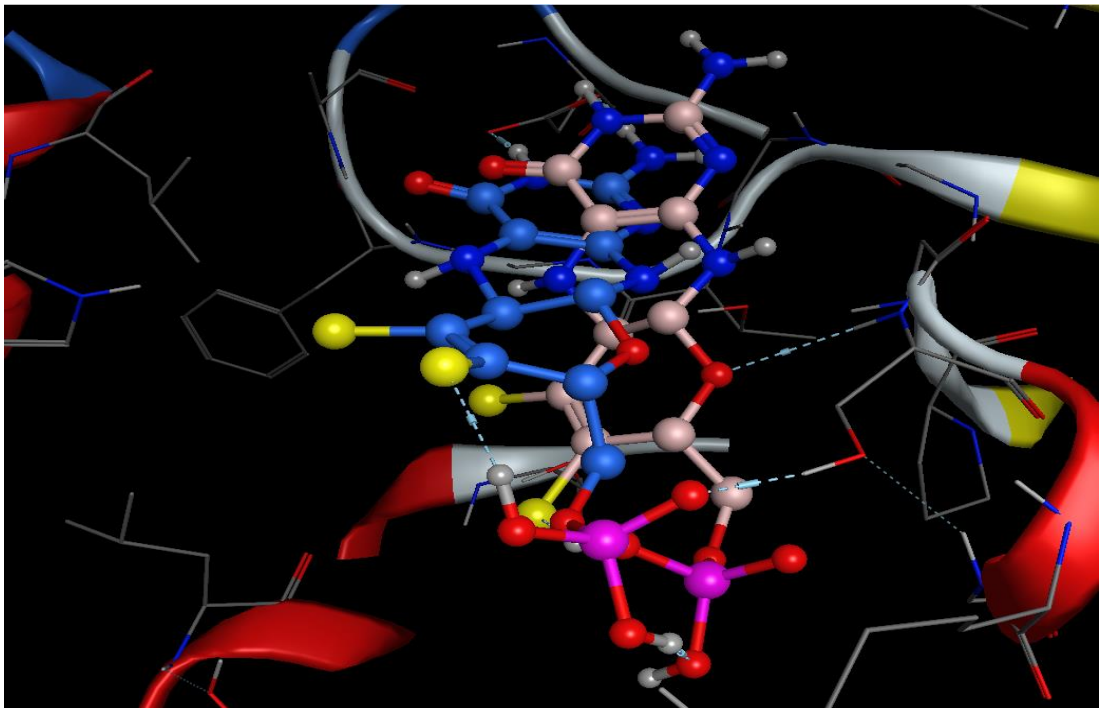


Figure 5.3: MPT - Initial Docked Pose (BLUE) vs Final Pose of MD Sim (PINK)

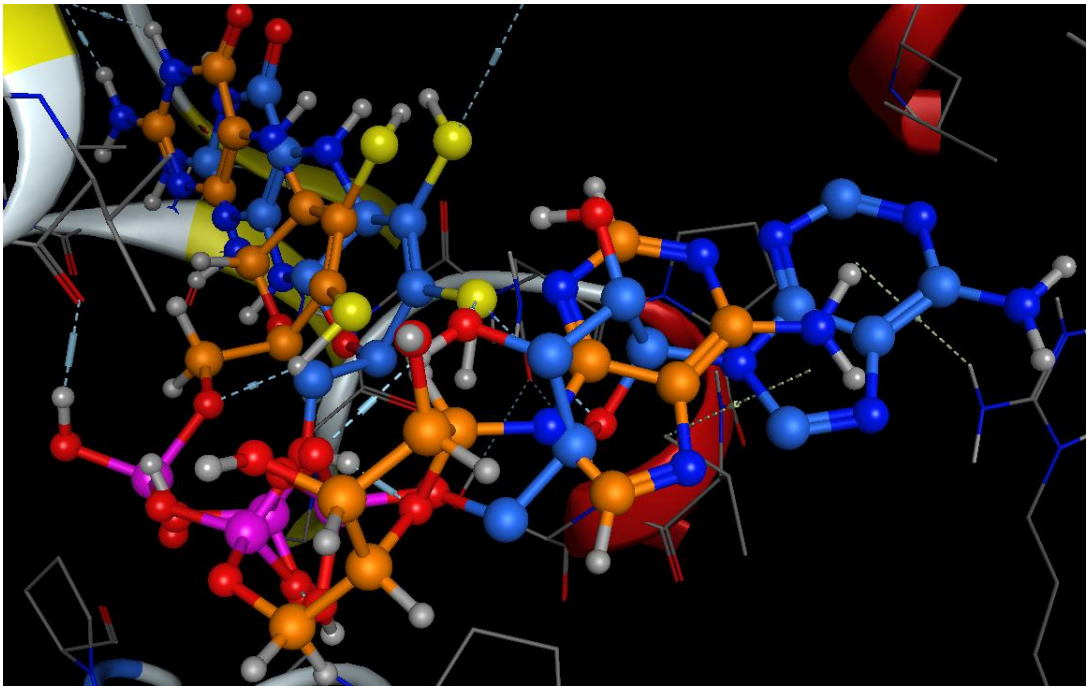


Figure 5.4:MPTAMP - Initial Docked Pose (BLUE) vs Final Pose of MD Sim (ORANGE)

So far the results of docking and MD simulation conducted in this study have been able to depict the role of Cu in MOCO biosynthesis. Its mechanism in the light of binding energies and interaction energies calculated hence far, giving insight on the metal exchange reaction between copper and molybdenum. This takes place in order to incorporate molybdenum into the mature cofactor.

GPHN gene encodes for gephyrin, which is exceedingly conserved and displays a complex intron-exon structure. The alternative splicing of GPHN leads to numerous gephyrin splice variants expressed both in neuronal and in non-neuronal tissues. Transcript variants of gephyrin are found not only in brain and spinal cord but also in liver, kidney, heart, and lungs. Interestingly, these are also the organs in which molybdenum dependent enzymes show the highest enzymatic activities. The gephyrin domains convoluted in molybdenum cofactor biosynthesis are Geph-G and Geph-E; dysfunction in the activity of these domains leads to symptoms of MOCO-deficient patients, which have severe neurological abnormalities such as increased muscles, rigid posturing, myoclonus, abnormal movements and refractory seizures, although these neurological functions are that of the C-domain. This indicates that all three domains of gephyrin work in a synergic fashion and the neuronal and non-neuronal functions performed by gephyrin are correlated to each other.

The role of molybdenum *Insertase* domain Geph-E specifically in type C is still up for debate as it is already known that MoCD is an autosomal recessive disorder. More knowledge is needed on the genetics of the enzymes involved in MOCO biosynthesis. Several types of mutations including nonsense, missense, frame shift and exon splicing are reported for MoCD. The genetic disposition of this disease runs far deeper than just mutations as maternal uniparental disomy for chromosome 6 (which in itself is a very rare condition) is also reported for MOCS1 which too leads to MoCD.[55,63]

CHAPTER 6

CONCLUSION

Computational analysis was done on the third and fourth (final) step of molybdenum cofactor biosynthesis to find the binding energies and interaction patterns as well as the reaction mechanism involved in the formation of MOCO from MPT with the help of enzyme Molybdenum *Insertase* in humans. For this purpose, PDB structures of the two protein domains of the enzyme were taken with PDB ID: 1jlj and PDB ID: 4tk1 from X-ray crystal structure data of Geph-G and Geph-E respectively; while the active site was predicted for both protein domains as little information was available for the binding site of Cnx1 (plant homolog of Geph-G) and almost no information on binding pocket for Geph-E was available in the literature. Once binding pocket was predicted, molecular docking of MPT and MPTAMP was performed with Geph-G and Geph-E respectively to identify the binding pocket residues present at the active site and involved in the formation of MOCO from MPT. Furthermore Molecular Dynamic simulations were performed to study the interaction pattern for both ligands with their respective protein domains. As a result of this computational analysis, it has been concluded that the MPT is unstable with or without the copper ion hence depicting the role of Cu ion as just a commodity to be used for replacement upon reaching the first domain Geph-G. Since the MPT is unstable ATP in required to not only provide the Adenine for adenylation which stabilizes the MPT, it also provides the necessary energy required to make this interaction possible. Once adenylated the MPT stabilizes and Cu ion can now be released and the ligand is passed along to the second protein domain Geph-E, where insertion of molybdenum in the form of molydate takes place while the adenine monophosphate is cleaved off to form the final MOCO cofactor (Mo-MPT) with a final binding energy of -7.9 kcal/mol. The interaction patterns were defined by comparing ligand bind pose of docking and MD

simulation which confirmed that the major contacts found during docking were maintained during the MD, hence fortifying the reliable of the results obtained.[9,23,38,54,59]

REFERENCE

1. Choii G, Ko J. Gephyrin: a central GABAergic synapse organizer. *Exp Mol Med*. 2015:47. doi:10.1038/emm.2015.5

2. Kneussel M, Betz H. Receptors, gephyrin and gephyrin-associated proteins: Novel insights into the assembly of inhibitory postsynaptic membrane specializations. *J Physiol*. 2000;525(1):1-9. doi:10.1111/j.1469-7793.2000.t01-4-00001.x

3. Pizzarelli R, Griguoli M, Zacchi P, et al. Tuning GABAergic Inhibition: Gephyrin Molecular Organization and Functions. *Neuroscience*. 2019. doi:10.1016/j.neuroscience.2019.07.036

4. Mele M, Leal G, Duarte CB. Role of GABAAR trafficking in the plasticity of inhibitory synapses. *J Neurochem*. 2016;139(6):997-1018. doi:10.1111/jnc.13742

5. Stallmeyer B, Schwarz G, Schulze J, et al. The neurotransmitter receptor-anchoring protein gephyrin reconstitutes molybdenum cofactor biosynthesis in bacteria, plants, and mammalian cells. *Proc Natl Acad Sci U S A*. 1999;96(4):1333-1338. doi:10.1073/pnas.96.4.1333

6. Smolinsky B, Eichler SA, Buchmeier S, Meier JC, Schwarz G. Splice-specific functions of gephyrin in molybdenum cofactor biosynthesis. *J Biol Chem*. 2008;283(25):17370-17379. doi:10.1074/jbc.M800985200

7. Nagappa M, Bindu PS, Taly AB, Sinha S, Bharath RD. Child Neurology: Molybdenum cofactor deficiency. *Neurology*. 2015;85(23):e175-e178. doi:10.1212/WNL.0000000000002194

8. Leimkühler S, Leimkühler L, Iobbi-Nivol C. Bacterial molybdoenzymes: old enzymes

for new purposes. *FEMS Microbiol Rev*. 2016;043:1-18. doi:10.1093/femsre/fuv043

9. Schwarz G, Schrader N, Mendel RR, Hecht H-J, Schindelin H. Crystal structures of human gephyrin and plant Cnx1 G domains: comparative analysis and functional implications. *J Mol Biol*. 2001;312(2):405-418. doi:10.1006/JMBI.2001.4952

10. Atwal PS, Scaglia F. Molybdenum cofactor deficiency. *Mol Genet Metab*. 2016;117(1):1-4. doi:10.1016/j.ymgme.2015.11.010

11. Tejada-Jimenez M, Chamizo-Ampudia A, Calatrava V, Galvan A, Fernandez E, Llamas A. From the eukaryotic molybdenum cofactor biosynthesis to the moonlighting enzyme MARC. *Molecules*. 2018;23(12). doi:10.3390/molecules23123287

12. Srivastava VK, Srivastava S, Srivastava S, Arora A, Pratap JV. Structural insights into putative molybdenum cofactor biosynthesis protein C (MoaC2) from Mycobacterium tuberculosis H37Rv. *PLoS One*. 2013;8(3):e58333. doi:10.1371/journal.pone.0058333

13. Tejada-Jiménez M, Chamizo-Ampudia A, Galván A, Fernández E, Llamas Á. Molybdenum metabolism in plants. *Metallomics*. 2013;5(9):1191-1203. doi:10.1039/c3mt00078h

14. Zhang Y, Rump S, Gladyshev VN. Comparative Genomics and Evolution of Molybdenum Utilization. *Coord Chem Rev*. 2011;255(9-10):1206-1217. doi:10.1016/j.ccr.2011.02.016

15. Schwarz G, Mendel RR, Ribbe MW. Molybdenum cofactors, enzymes and pathways. *Nature*. 2009;460(7257):839-847. doi:10.1038/nature08302

16. Hille R, Hall J, Basu P. The mononuclear molybdenum enzymes. *Chem Rev*. 2014;114(7):3963-4038. doi:10.1021/cr400443z

17. Hille R. The Mononuclear Molybdenum Enzymes †. *Chem Rev*. 1996;96(7):2757-2816. doi:10.1021/cr950061t

18. Mendel RR, Zwerschke D, Hercher TW, et al. The functional principle of eukaryotic molybdenum insertases. *Biochem J*. 2018;475(10):1739-1753. doi:10.1042/bcj20170935

19. Leimkühler S. The Biosynthesis of the Molybdenum Cofactor in Escherichia coli and Its Connection to FeS Cluster Assembly and the Thiolation of tRNA . *Adv Biol*. 2014;2014(Figure 1):1-21. doi:10.1155/2014/808569

20. Mendel RR, Kruse T. Cell biology of molybdenum in plants and humans. *Biochim Biophys Acta - Mol Cell Res*. 2012;1823(9):1568-1579. doi:10.1016/J.BBAMCR.2012.02.007

21. Leimkühler S. Shared function and moonlighting proteins in molybdenum cofactor biosynthesis. *Biol Chem*. 2017;398(9):1009-1026. doi:10.1515/hsz-2017-0110

22. Mendel RR. The molybdenum cofactor. *J Biol Chem*. 2013;288(19):13165-13172. doi:10.1074/jbc.R113.455311

23. Kuper J, Llamas A, Hecht H-J, Mendel RR, Schwarz G. Structure of the molybdopterin-bound Cnx1G domain links molybdenum and copper metabolism. *Nature*. 2004;430(7001):803-806. doi:10.1038/nature02681

24. Niks D, Hille R. Molybdenum- and tungsten-containing formate dehydrogenases and formylmethanofuran dehydrogenases: Structure, mechanism, and cofactor insertion. *Protein Sci*. 2019;28(1):111-122. doi:10.1002/pro.3498

25. Hover BM, Tonthat NK, Schumacher MA, Yokoyama K. Mechanism of pyranopterin

ring formation in molybdenum cofactor biosynthesis. doi:10.1073/pnas.1500697112

26. Burgess BK, Lowe DJ. Mechanism of Molybdenum Nitrogenase. *Chem Rev*. 2002;96(7):2983-3012. doi:10.1021/cr950055x

27. Mechler K, Mountford WK, Hoffmann GF, Ries M. Ultra-orphan diseases: a quantitative analysis of the natural history of molybdenum cofactor deficiency. *Genet Med*. 2015;17(12):965-970. doi:10.1038/gim.2015.12

28. Wang J, Krizowski S, Fischer-Schrader K, et al. Sulfite Oxidase Catalyzes Single-Electron Transfer at Molybdenum Domain to Reduce Nitrite to Nitric Oxide. *Antioxidants Redox Signal*. 2015;23(4):283-294. doi:10.1089/ars.2013.5397

29. Reiss J, Christensen E, Dorche C. Molybdenum cofactor deficiency: First prenatal genetic analysis. *Prenat Diagn*. 1999;19(4):386-388. doi:10.1002/(SICI)1097-0223(199904)19:4<386::AID-PD550>3.0.CO;2-#

30. Hubert L, Sutton VR. Molybdenum Cofactor Learn more about Molybdenum Cofactor Disorders of purine and pyrimidine metabolism Genetic-Metabolic Disorders Presenting as Acute , but Reversible , Severe Epilepsies. 2017.

31. Umar IK, Samir B. Cumhuriyet Science Journal CSJ. *Cumhur Sci J*. 2019;40(1):197-203. doi:http://dx.doi.org/10.17776/csj.356185

32. Irwin JJ, Shoichet BK. ZINC – A Free Database of Commercially Available Compounds for Virtual Screening. *J Chem Inf Model*. 2005;45(1):177. doi:10.1021/CI049714

33. Li P, Chen S. A review on Gaussian Process Latent Variable Models. *CAAI Trans Intell Technol*. 2016;1(4):366-376. doi:10.1016/j.trit.2016.11.004

34. Schaftenaar G, Vlieg E, Vriend G. Molden 2.0: quantum chemistry meets proteins. *J*

Comput Aided Mol Des. 2017;31(9):789-800. doi:10.1007/s10822-017-0042-5

35. Ahmed MZ, Habib U. DFT studies of temperature effect on coordination chemistry of Cu (II) -trimethoprim complexes. *J Coord Chem*. 2018;8972(Ii):1-12. doi:10.1080/00958972.2018.1447667

36. Rakesh Bhatnagar1 DJP. (PDF) Performance Analysis of A Grid Monitoring System-Ganglia. *Int J Emerg Technol Adv Eng*. 2013;3(8):5. doi:ISSN 2250-2459, ISO 9001:2008

37. Berman HM, Battistuz T, Bhat TN, et al. The protein data bank. *Acta Crystallogr Sect D Biol Crystallogr*. 2002;58(6 I):899-907. doi:10.1107/S0907444902003451

38. Maric HM, Kasaragod VB, Hausrat TJ, et al. Molecular basis of the alternative recruitment of GABAA versus glycine receptors through gephyrin. *Nat Commun*. 2014;5:5767-5767. doi:10.2210/PDB4TK1/PDB

39. Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res*. 2003;31(13):3381-3385. doi:10.1093/nar/gkg520

40. Istyastono EP. Docking Studies of Curcumin As a Potential Lead Compound To Develop Novel Dipeptydyl Peptidase-4 Inhibitors. *Indones J Chem*. 2019;9(1):132-136. doi:10.22146/ijc.21574

41. Chimera U. Autodock tutorial. 2014:1-5. doi:10.1016/j.physa.2012.01.046

42. Kasaragod VB, Schindelin H. Structural Framework for Metal Incorporation during Molybdenum Cofactor Biosynthesis. *Structure*. 2016;24(5):782-788. doi:10.1016/j.str.2016.02.023

43. Volkamer A, Kuhn D, Rippmann F, Rarey M. Dogsitescorer: A web server for automatic binding site prediction, analysis and druggability assessment. *Bioinformatics*. 2012;28(15):2074-2075. doi:10.1093/bioinformatics/bts310

44. Suresh C V. DOCKING-A Review. *J Appl Chem*. 2012(2):167-173. doi:2278-1862

45. Meng X-Y, Zhang H-X, I MM, Cui M. Molecular docking: a powerful approach for structure-based drug discovery. *Curr Comput Aided Drug Des*. 2011;7(2):146-157. doi:ISSN 1875-6697

46. Sgrignani J, Magistrato A. QM/MM MD Simulations on the Enzymatic Pathway of the Human Flap Endonuclease (hFEN1) Elucidating Common Cleavage Pathways to RNase H Enzymes. 2015. doi:10.1021/acscatal.5b00178

47. Hess B, Kutzner C, Van Der Spoel D, Lindahl E. GRGMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput*. 2008;4(3):435-447. doi:10.1021/ct700301q

48. Apol E, Apostolov R, Berendsen HJC, et al. GROMACS 2016 Reference Manual. 1991. doi:9781597181600

49. Protein-Ligand Complex. http://www.mdtutorials.com/gmx/complex/index.html. Accessed November 6, 2019.

50. CGenFF Home. https://cgenff.umaryland.edu/. Accessed November 6, 2019.

51. Yuan S, Chan HCS, Hu Z. Using PyMOL as a platform for computational drug design. *Wiley Interdiscip Rev Comput Mol Sci*. 2017;7(2). doi:10.1002/wcms.1298

52. Hanwell MD, Curtis DE, Lonie DC, Vandermeerschd T, Zurek E, Hutchison GR. Avogadro: An advanced semantic chemical editor, visualization, and analysis platform.

J Cheminform. 2012;4(8). doi:10.1186/1758-2946-4-17

53. Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. *J Mol Graph*. 1996;14(1):33-38. doi:10.1016/0263-7855(96)00018-5

54. Atwal PS, Scaglia F. Molybdenum cofactor deficiency. *Mol Genet Metab*. 2016;117(1):1-4. doi:10.1016/j.ymgme.2015.11.010

55. Gümüş H, Ghesquiere S, Per H, et al. Maternal uniparental isodisomy is responsible for serious molybdenum cofactor deficiency. *Dev Med Child Neurol*. 2010;52(9):868-872. doi:10.1111/j.1469-8749.2010.03724.x

56. Bender D, Kaczmarek AT, Santamaria-Araujo JA, et al. Impaired mitochondrial maturation of sulfite oxidase in a patient with severe sulfite oxidase deficiency. *Hum Mol Genet*. 2019;28(17):2885-2899. doi:10.1093/hmg/ddz109

57. Ghahramany N, Gharaati S, Ghanaatian M. New approach to nuclear binding energy in integrated nuclear model. *Phys Part Nucl Lett*. 2011;8(2):97-106. doi:10.1134/S1547477111020087

58. Allori V, Goldstein S, Tumulka R, Zanghì N. Many worlds and Schrödinger's first quantum theory. *Br J Philos Sci*. 2011;62(1):1-27. doi:10.1093/bjps/axp053

59. Haroon K. Computational Analysis on the Formation of Metal-Binding Pterin (MPT) from cyclic Pyranopterin Monophosphate (cPMP) in Molybdenum Cofactor Biosynthesis. 2019:5-10.

60. Sidorenkov A V, Kolesnikov S V, Saletsky AM. *Molecular Dynamics Simulation of Graphene on Cu(111) with Different Lennard-Jones Parameters*.; 2016. doi:1605.01594v1

61. Baowan D, Hill JM. Mathematical modeling of interaction energies between nanoscale objects: A review of nanotechnology applications. *Adv Mech Eng*. 2016;8(11):168781401667702. doi:10.1177/1687814016677022

62. Faro TMC, Thim GP, Skaf MS. A Lennard-Jones plus Coulomb potential for Al3+ ions in aqueous solutions. *J Chem Phys*. 2010;132(11). doi:10.1063/1.3364110

63. Mayr SJ, Sass JO, Vry J, et al. A mild case of molybdenum cofactor deficiency defines an alternative route of MOCS1 protein maturation. *J Inherit Metab Dis*. 2018;41(2):187-196. doi:10.1007/s10545-018-0138-7