Identification and Computational Analysis of 3D-Structure Of MOCS1

Maham Hamid
National University of Sciences and Technology
uzma habib (✉ uzma.habib@rcms.nust.edu.pk)
National University of Sciences and Technology https://orcid.org/0000-0001-6592-6419
Javeria Batool
National University of Sciences and Technology
Arshemah Qaisar
National University of Sciences and Technology
Rehan Zafar Paracha
National University of Sciences and Technology

Research Article

Keywords: MOCS1, molybdenum cofactor, tertiary structure, Docking, mutations1

Posted Date: June 14th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-473901/v1

Abstract

Cyclic pyranopterin monophosphate (cPMP) is one of the most stable intermediates in Molybdenum cofactor (MoCo) biosynthetic pathway. In humans, synthesis of cPMP from Guanosine triphosphate (GTP) requires functional genes i.e. Molybdenum Cofactor Synthesis-1 (MOCS1) genes that contains for two catalytic proteins MOCS1A and MOCS1B. Importance of MOCS1A and MOCS1B for biosynthesis of MoCo reveals from the fact that its deficiency leads to MoCo type A deficiency. As there is no structure available for MOCS1 genes in the literature, tertiary structure of MOCS1 genes were investigated in this research via threading or folds recognition method by i-TASSER and validation was done using ERRAT, Verify3D and Ramachandran plots. Binding sites were predicted and validated. Docking of MOCS1A with GTP and MOCS1B with 3, 8 dihydroguanosine was done using Autodock via PyRx. Apart from this, highly confident mutations were also predicted using SIFT and polyphen2 that can alter the structure and function of MOCS1 gene.

1. Introduction

Molybdenum cofactor biosynthesis protein 1 (encoded by MOCS1 gene) is a human protein involved in the formation of molybdenum cofactor (MOCO). It is also present in animals, fungi and some cellular slime molds[1]. Cytogenetic location of this gene is 6p21.2, which means that it is present on short arm p of chromosome 6 at position 21.2. Molecular Location is from base pairs 39,904,170 to 39,934,462 on chromosome 6 [2]. The most interesting feature of the MOCS1 gene (MOCS1A and MOCS1B proteins) is that it is responsible in the formation of first intermediate, “precursor Z (cyclic pyranopterin monophosphate (cPMP)” in MOCO biosynthesis pathway (Fig. 1)[3, 4].

First step is the formation of first intermediate cyclic pyranopterin monophosphate (cPMP) from GTP, which is converted into Molybdopterin (MPT). MPT is adenylated followed by insertion of molybdate into adenylated MPT and release of AMP. Final step involves conversion of MPT-AMP into mature Moco [4].

Molybdenum cofactor (MOCO) containing enzymes are very important as they help in the catalysis of many redox reactions involved in various cycles like carbon, sulfur and nitrogen. Most crucial molybdenum containing enzyme is sulfite oxidase (SO) as it helps in degradation of sulfur containing amino acids [5]. Nitrate reductase (NR) is similar to sulphite oxidase but is abundant in autotrophic organisms. Xanthine oxidoreductase are involved in purine catabolism and leads to apoptosis. Moco depicts an essential role in sulfite oxidase, aldehyde oxidases, xanthine dehydrogenase, and hypotaurine dehydrogenase, its deficiency can cause multiple diseases related to kidney functions [6]. The molybdenum cofactor deficiency may occur due to molybdenum deficiency or disturbed molybdopterin cofactor synthesis.

The first model of MOCO biosynthesis was given by Rajagopalan and co-worker [7]. In eukaryotes like fungi, plants, and humans, six agents have been identified that catalyzes the reactions involved in synthesis of Moco. The naming of these agents in plants is based on cnx nomenclature i.e. cofactor for

nitrate reductase and xanthine dehydrogenase, however, in humans they are named as MOCS (molybdenum cofactor synthesis)[8]. MOCO biosynthesis pathway is divided into four major steps, as shown in Fig. 1. As one can observe, the first step is the conversion of GTP into the first intermediate cyclic pyranopterin monophosphate (cPMP), which is aided by MOCS1A and MOCS1B [9].

MOCS1A is responsible for the cyclization of GTP to (8S)-3',8-cyclo-7,8-dihydroguanosine 5'-triphosphate while MOCS1B is responsible for the subsequent conversion of (8S)-3',8-cyclo-7,8- dihydroguanosine 5'-triphosphate to cPMP [9] (Fig. 3). In *E. Coli*, MoaA [10] and MoaC [11] (homologous proteins of MOCS1A and MOCS1B) are responsible for the generation of cPMP by complicated reactions. MoaA and MoaC proteins purified from *S. Aureus* for the synthesis of cPMP were most recently studied by Hanzelman and Schindelin using an in vitro system [12]. MoaA is involved in the first step but it was the last protein to be characterized structurally after reported crystal structure of *S. Aureus* MoaA by Hanzelman and Schindelin in 2004.

The importance of MOCS1A and its homologous proteins like MoaA and MoaC are described by their structures where two highly conserved cysteine clusters (one in the N-terminal and one in the C-terminal) are proposed to be involved in binding of FeS cluster [13]. Several mutations were identified in molybdenum cofactor deficiency patients these mutations were positioned in the conserved cysteine motifs. This indicates the functional importance for MOCS1/ homologous proteins as their mutations lead to Molybdenum Cofactor deficiency.

Molybdenum cofactor deficiency (MoCD) is a rare human disease responsible for many neurological disorders including brain atrophy which leads to delay in brain development [14, 15]. Few of the MoCD affected individuals shows symptoms of hyperekplexia (startle reaction), microcephaly (small head size) and coarse (abnormal facial features) [16].

MOCO deficiency leads to a loss in activity of Mo-dependent enzymes. There are three types of MoCD, i.e. type A, B, and C caused by the lack/ dysfunctioning of proteins MOCS1, MOCS2, and Gephyrin respectively [17]. Until now, no exact mechanism or reason has been identified that causes the dysregulation in the function of these enzymes. But many studies have suggested that mutation in the sequence and structure of gene may be responsible for this abnormality in MOCO pathway. To determine the cause, one may need to check the structural composition of the protein involved.

To best of our knowledge, in the protein database PDB, the crystallographic 3D structure of MOCS1A/B is not yet reported. Therefore, computational analysis of MOCS1 structure can be of advantage in the prediction of its probable structure. Therefore, we performed computational modeling and simulation in the hope to understand the properties and structure of MOCS1. Later, GTP ligand was docked to MOCS1A after determining binding sites and computational analysis were performed.

2. Materials And Methods

2.1. 3D Structure Prediction

In this study, sequence retrieval was done from Uniprot database for protein MOCS1 with Uniprot ID Q9NZB8. For MOCS1A identifier: Q9NZB8-5 and for MOCS1B identifier: Q9NZB8-1 was retrieved [18]. The MOCS1A/B FASTA sequence functioned as a query sequence against a non-reductant protein database (nr) in BLAST for homology search [19]. SOPMA i.e. Self-optimized prediction method with alignment was used for secondary structure prediction of MOCS1A and MOCS1B proteins [20]. It is a self-optimized prediction method that improves the success rate in the prediction of the secondary structure of protein. The number of conformational states: 4(helix, Sheet, Turn, and Coil), similarity threshold: 8, and window width 17 were set as parameters for SOPMA server [21]. 3D structure prediction was performed using i-TASSER (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) [22], i.e. iterative Threading ASSEmbly Refinement, a platform for protein function and structure prediction from its amino acid sequence by threading [23]. The accuracy of predicted structure is determined by confidence score or C-score, ranges from 0 to 1. The higher value of the C-score, the more reliable is the prediction [24]. The 3D structure evaluation in this study was performed by online quality evaluation tools via Structural analysis and verification server (SAVES) (https://servicesn.mbi.ucla.edu/SAVES/). There are about five servers in SAVES i.e. Verify 3D [24], ERRAT [25], PROCHECK [26], WHATCHECK, and IFCHECK [27]. In our study, we used ERRAT and Verify3D for structure validation. RAMPAGE and PROCHECK servers are used for Ramachandran plot and statistics analysis [26].

2.2. Molecular Dynamics

For protein structure having bad scores in model evaluation, 500,000 steps MD simulations were carried out to relax the bad contacts of proteins. By using GROMACS version 5.1 software and OPLS-AA force field, molecular dynamics simulations were performed to refine models [28]. PyMOL and visual molecular dynamics (VMD) applications were used to visualize molecule structure Post-MD simulations [29]. Grace application in Linux was used to display graphs. Gromacs does not work alone, so all of the above three applications were used for MD simulation analysis[30]. Moreover, MD simulation helps in fixing the bad contacts at the active site of the protein. It also improves bad ERRAT or Verify3D scores [31]. Structure comparison of MOCS1 before and Post-MD simulation was done using Chimera [32]. Ligand binding site for MOCS1 in this study was predicted using DoGSiteScorer [33] and RaptorX [34]. Two software were used for analysis so that we can gain more confidence in the result. The best scores among these were used for further analysis.

2.3. Molecular docking

Molecular docking of ligands into their respective proteins was performed using PyRx [35] which is a Virtual Screening software for Computational Drug Discovery. Autodock, which is a part of PyRx, was used as docking software [36]. While docking runs, the target 3D structure was fixed. To find the best binding modes, the ligand is rotated and moved [35].

3. Results And Discussion

cPMP formation, the first step in molybdenum cofactor biosynthesis, needs two proteins i.e. MOCS1A and MOCS1B [37]. The structure of MOCS1, to best of our knowledge, is not reported in PDB till now. So structure of MOCS1A and MOCS1B was required to be predicted to understand the formation of cPMP. Therefore, to understand the structure and function of MOCS1, structure prediction protocol was followed where computational modeling and simulation techniques were applied to predict the structures of MOCS1A and MOCS1B. Later, binding site of MOCS1A and MOCS1B were also determined.

3.1. 3D Structure Prediction

3.1.1. Sequence Retrieval and Homology Search

The Sequence retrieval from Uniprot database was done to obtain sequence of MOCS1A (identifier: **Q9NZB8-5**) and MOCS1B (identifier: **Q9NZB8-1**) in FASTA format to carry out homology search. MOCS1A helps in catalyzing 5'-GTP conversion to cPMP i.e. cyclic pyranopterin monophosphate, while MOCS1B helps in catalyzing (8S)-3',8-cyclo-7,8-dihydroguanosine 5'-triphosphate conversion to cPMP [38]. MOCS1 protein has eight isoforms known in Uniprot entry out of which Isoforms MOCS1A and MOCS1B most probably form a complex. MOCS1A isoform has 385 amino acids while MOCS1B isoform has 636 amino acid residues, respectively. The first step included to search for homologous templates but it didn't bring fruitful results. As MOCS1A/B percent identity was below 36.45% for both proteins MOCS1A and MOCS1B (Table-1) in BLAST results against protein data bank.

Table 1
Sequences producing significant alignment with MOCS1A and MOCS1B

Protein	Accession	Percent identity	E-value	Query coverage
MOCS1A	1TV7_A	36.45%	2e-60	75%
	2BF2_A	35.45%	2e-57	75%
	6C8V_A	32.47%	1e-08	39%
	6EFN_A	27.11%	3e-09	63%
	5WGG_A	25.00%	0.009	26%
MOCS1B	1TV7_A	36.45%	2e-60	45%
	2BF2_A	35.45%	2e-57	45%
	4PYD_A	48.37%	1e-08	24%
	1EKR_A	48.37%	3e-09	24%
	4PYA_A	47.71%	0.009	24%

For validation of tertiary structure, secondary structure can be used. Secondary structure tells us that how much protein is composed of helix, coil and strands. Therefore, secondary structures for MOCS1A and MOCS1B were also obtained. For MOCS1A, secondary structure components are alpha helix (41%),

extended strand (16.63%), beta turn (6.49%) and random coil (35.8%), while for MOCS1B, secondary structure components are alpha helix (35.38%), extended strand (16.19%), beta turn (5.66%) and random coil (42.77%).

3.1.2. Tertiary Structure Prediction

Tertiary structure prediction was carried out using i-TASSER, SWISS-MODEL, and RaptorX. Out of five best models for MOCS1A and MOCS1B generated by. i-TASSER simulation, model 1 (Figure S1 of Supplementary Material file) with − 1.49 confidence score (Table 2) was selected for MOCS1A, and model 3 (Figure S2 of Supplementary Material file) with C-score − 2.07 was selected for MOCS1B for further study. Swiss model generated two models for MOCS1A and three models for MOCS1B (Figure S3 and S4 of Supplementary Material file). It takes FASTA sequence as input. The best scored model selected based on Q-mean value are SWISS-model for MOCS1A with − 3.20 QMEAN score and SWISS-model for MOCS1B with QMEAN value − 1.67 (Table 2). RaptorX also takes FASTA sequence as input and generate a model (Figure S5 and S6 of Supplementary Material file). The predicted model by RaptorX for MOCS1A was with P-value 2.3e-08 and uGDT value of 223. For MOCS2A it generated three domains (modeling unit i.e. based on template from PDB, query sequence being partitioned into units), domain 1 has P-value of 7.57e-08 and uGDT 211 while domain 2 has P-value 6.18e-09 and uGDT 392 (Table 2).

Table 2
Scores for MOCS1A and MOCS1B generated by i-TASSER

Models	C-score for MOCS1A	C-score for MOCS1B
Model 1	-1.49	-2.98
Model 2	-1.78	-3.71
Model 3	-1.83	-2.07
Model 4	-1.95	-3.87
Model 5	-2.42	-3.82

3.1.3. Model evaluation and validation

To determine the quality of predicted structures, the overall structure's evaluation was carried out by using Ramachandran plot statistics obtained from RAMPAGE for comparing the results generated by i-TASSER, Swiss-Model, and RaptorX. Ramachandran plot statistics for MOCS1A/B protein models obtained by three different methods are presented in Table 3.

Table 3
Ramachandran plot statistics for MOCS1A/B protein models obtained from i-TASSER, Swiss-Model and RaptorX

Software	proteins	Favored region	Allowed Region	Outlier Region
i-TASSER	MOCS1A	279(72.8%)	73(19.1%)	31(8.1%)
	MOCS1B	490(77.3%)	98(15.5%)	46(7.3%)
Swiss-Model	MOCS1A	287(91.4%)	17(5.4%)	10(3.2%)
	MOCS1B	581(92.7%)	28(4.5%)	18(2.9%)
RaptorX	MOCS1A	318(93.5%)	18(5.3%)	4(1.2%)
	MOCS1B	487(94.9%)	21(4.1%)	5(1.0%)

Analysis of the results (Table 3) shows that i-TASSER generated models are with satisfactory results as compared to Swiss-Model and RaptorX, as they both have predicted 3D structure for MOCS1A and MOCS1B partially. Therefore, 3D models for MOCS1A and MOCS1B generated by i-TASSER were considered for evaluation and validation in SAVES server by ERRAT and Verify3D.

ERRAT is a protein structure verification tool that is suitable for evaluating the progress of crystallographic model building and refinement. It works by analyzing the statistics of non-bonded interactions between different atom types and comparing them with databases of high-resolution structures. The region of the model that fall in 90–99% can be rejected. ERRAT gives an overall quality score besides highlighting the error-prone region. Figure 4 shows the overall quality factor of 83.82%.for MOCS1A model by i-TASSER whereas Fig. 5 shows plot for MOCS1B protein structure having 91.06% ERRAT score that is good score showing that model is satisfactory.

The average 3D-1D profile score was obtained using Verify3D tool. Verify3D helps to investigate the compatibility of an atomic model with its own 1D amino acid sequence. It usually shows the 3D-ID score ranges from − 1 (bad score) to + 1 (good score). The Verify3D plot shows that for the protein MOCS1A, 84.68% of residues have averaged 3D-1D score > = 0.2 (Fig. 6) and for protein MOCS1B the average 3D-1D score is 76.26% (Fig. 7).

The structures of MOCS1A and MOCS1B were validated through RC plot. Based on these factors, it seemed that the models were pretty good. The Ramachandran plot of MOCS1A and MOCS1B model proteins obtained from i-TASSER (Figs. 8 and 9) shows that various residues are falling under allowed, favored, and outer regions.

The Ramachandran plot statistics for both model proteins MOCS1A and MOCS 1B are presented in Table 4 which shows that for MOCS1A about 80% residues are in the allowed region and 3.6% residues (only 12 residues) are in disallowed region which is negligible. For MOCS1B, more than 90% residues are in the allowed region and 1.8% residues (only 10 residues) are in disallowed region. It can also be seen from Ramachandran plots that most data lie in favored region.

Table 4
Ramachandran plot for MOCS1A/ MOCS1B by PROCHECK

Plot statistics	MOCS1A		MOCS1B	
Residues in most favoured regions [A,B,L]	228	67.7%	377	67.7%
Residues in additional allowed regions [a,b,l,p]	80	23.7%	150	26.9%
Residues in generously allowed regions [~ a,~b,~l,~p]	17	5.0%	20	3.6%
Residues in disallowed regions	12	3.6%	10	1.8%
Number of non-glycine and non-proline residues	337	100.0%	557	100.0%
Number of end-residues (excl. Gly and Pro)	1		2	
Number of glycine residues (shown as triangles)	28		42	
Number of proline residues	19		35	
Total number of residues	385		636	

3.1.4. Molecular Dynamics Simulation

For removing conflicts and bad contacts between the main chain and side chains of modeled MOCS1 proteins, MD simulations were performed using GROMACS5.1 to refine the models. Analysis of output data was done according to Energy, root-mean-square deviation (RMSD), and Gyration radius. The total energy (sum of the kinetic and potential energy of the molecules) should be constant throughout the simulation process. Kinetic energy is expected to be constant or should follow a declining trend. This is because the increase in the level of kinetic represents confusion of protein structure. While potential energy should follow an increasing trend or should be constant to ensure structure stability.

Distance between the backbones of protein atoms is measured by RMSD. Figures 10 & 11 show plot for RMSD of MOCS1A and MOCS1B respectively. The plot shows that RMSD gets leveled at 0.25nm for MOCS1A and 0.35 nm in the case of MOCS1B, representing protein stability. The structural refinement displays that RMSD plots for predicted models of MOCS1 flattens during 1 ns, which indicates that proteins are stable throughout the simulation.

The compactness (folded or unfolded) of a protein structure was measured using radius of gyration plots. Results indicate that the MOCS1 proteins are very stable in its compact form over the course of 1000ps (1ns) at 300K. Radius of gyration plots for MOCS1A and MOCS1B are presented in the supplementary data file (Figure S7 and S8).

Structure comparison of MOCS1A and MOCS1B Pre and Post-MD simulation was done using chimera. The results showed that predicted structures are constantly stable and it ensured proper folding of the protein. Figures 12 and 13 shows structure comparison of MOCS1A & MOCS1B Pre and Post-MD simulation.

Ramachandran plot analysis was carried out on Post-MD simulation structures using RAMPAGE, showing improvement in statistics of plots (Table 5). For Post-MD MOCS1A, 82.5% residues are in core region and for MOCS1B, 87.2% residues are in the most favoured region.

Table 5
Plot statistics for MOCS1A and MOCS1B obtained from RAMPAGE

Proteins	MD simulation status	Favored region	Allowed region	Outlier region
MOCS1A	Pre-MD	72.8%	19.1%	8.1%
	Post-MD	82.5%	15.7%	1.8%
MOCS1B	Pre-MD	76.7%	16.4%	6.9%
	Post-MD	87.2%	10.7%	2.1%

3.2. Ligand binding site prediction

To study the reaction mechanism associated with MOCS1 proteins, the knowledge of the active site was very important. Therefore, after the structure prediction of MOCS1 proteins, MOCS1A and MOCS1B, identification of binding pockets and the active site was required. There is no binding pocket identified in MOCS1 enzyme for MOCS1A to bind GTP and MOCS1B to bind 3',8-cyclo-7,8-dihydroguanosine 5'-triphosphate in literature. Uniprot only suggested the active site residues located around 317–319 for GTP binding but not the exact active site was reported. DoGSite Scorer and RaptorX were utilized to determine the active site and the important residues present at that site for ligand binding.

Results generated from RaptorX showed that the binding pocket for MOCS1A, with highest multiplicity of 139, contains amino acid residues SER-178 and MET-249 as interacting residues. This pocket also contained additional amino acid residues MET-88, GLU-117, THE-125, GLY-126, GLY-127, GLU-128, THR-154, ASN-156, ARG-192, MET-248, and METH-249 (Fig. 14). Results generated for MOCS1B showed that the binding pocket for GTP contain amino acid residues, ASP-606, LYS-529, and LYS-609 as the interacting residues with a pocket multiplicity of 40 (Fig. 15).

For comparison, binding pockets for MOCS1A and MOCS1B were also identified using DoGSite Scorer. First binding pocket that shows highest drug score of 0.85 is considered as top binding pocket for GTP ligand in MOCS1A protein (Figure S9 presented in Supplementary material file). It contains amino acid residues LEU-50, ALA-51, HIS-53, PHE-60, LEU-61, THR-62, GLN-68, HIS-69, SER-70, VAL-115, GLU-117, GLY-118, ILE-119, ASP-120, LYS-121, GLY-147, GLY-356, VAL-359 and GLY-360. On the other hand, the first binding pocket in MOCS1B for 3',8-cyclo-7,8-dihydroguanosine 5'-triphosphate ligand shows maximum drug score of 0.83 (Figure S10 presented in Supplementary material file) and contain amino acid residues GLN_476, LEU_477, THR_478, SER_479, GLN_481, LEU_482, THR_483, HIS_484, VAL_485, ALA_492, MET_493, VAL_494, ARG_504, ALA_544, LYS_545, THR_547, SER_548, ILE_551, PRO_552, LEU_553, CYS_554, HIS_555, HIS_556, VAL_557, ALA_558, LEU_559, PRO_586, THR_587, GLY_588, VAL_589,

GLU_590, MET_591, GLU_592, LEU_594, ARG-613, ILE_620, LYS_621, LEU_622, SER_624, LYS_625, THR_626, GLY_627, GLY_628, GLN_629, ARG_630, GLY_631, ASP_632, PHE_633, HIS_634, ARG_635.

By comparing the results obtained from DoGSiteScorer and RaptorX, it was observed that there were few amino acid residues common or nearly similar in binding pockets for GTP and 3',8-cyclo-7,8-dihydroguanosine 5'-triphosphate ligand suggested by both tools used for the identification of MOCS1 active site binding pocket search. These amino acids are GLU-117 and MET-249 for MOCS1A and LYS-625, GLY-628 and GLN-629 for MOCS1B. After finding the binding pocket site, it was required to predict the stability of complex which was done using molecular docking.

Molecular docking using PyRx was done to predict the structure of resultant stable complex attained from the known ligand and predicted protein model. Molecular docking of GTP to the predicted binding sites of MOCS1A and of *3',8-cyclo-7,8-dihydroguanosine 5'-triphosphate* to the MOCS1B was done. Interaction of protein residues and ligand showed that GLN-68, ASP-120, LYS-121, GLU-284, GLU-285, SER-286, ASN-316 and ARG-317 are the residues interacting with ligand GTP and SER-398, ILE-399, GLY-465, SER-466, TRP-467, ALA-468, ARG-613, ASP-614 and ILE-615 are the residues of MOCS1B interacting with *3',8-cyclo-7,8-dihydroguanosine 5'-triphosphate* (Molecular docking results of MOCS1A and MOCS1B are presented as Figure S11, Figure S12, Table S1 and Table S2 in Supplementary data file).

4. Conclusion

Computational analysis was done on the MOCS1 proteins, MOCS1A and MOCS1B, involved in the first step, formation of cPMP from GTP, of biosynthesis of molybdenum cofactor in humans. To study its reaction mechanism, the knowledge of the active site was very important. As there was no structure available in PDB for MOCS1A and MOCS1B proteins, therefore, in the first step the 3D structure of MOCS1A and MOCS1B were identified using i-TASSER and verified using Ramachandran plots. After structure prediction the second step was the identification of binding pocket and active sites for the ligands involved in the formation of cPMP. This was done using DoGSite Scorer and RaptorX. After successful prediction of binding pockets, molecular docking was performed. This shows the interaction of GTP and *3',8-cyclo-7,8-dihydroguanosine 5'-triphosphate* with the amino acid residues of MOCS1A and MOCS1B respectively, which thus led us to better understanding of the reaction involved in the formation of cPMP from GTP in molybdenum cofactor biosynthesis.

Declarations

Funding: This research was funded by HEC- National Research Program for Universities Pakistan, grant number 10008/Federal/ NRPU/R&D/HEC/ 2017.

Conflicts of Interest: The authors declare no conflict of interest.

Availability of data and material: Not applicable

Code availability: Not applicable

Author Contributions: Conceptualization, U.H.; methodology, U.H. and R.Z.P.; software, J.B. and M.H.; validation, U.H., R.Z.P., and M.H.; formal analysis, J.B. and M.H.; investigation, J.B. and M.H.; resources, U.H. and R.Z.P.; data curation, U.H., J.B., M.H. and A.Q.; writing—original draft preparation, M.H. and J.B.; writing—review and editing, U.H., R.Z.P., M.H. and A.Q; visualization, U.H., R.Z.P., and M.H.; supervision, U.H.; project administration, U.H.; funding acquisition, U.H.

Ethics approval: Not applicable

Consent to participate: Not applicable

Consent for publication: Not applicable

Acknowledgments: The authors wish to thank Super Computing Research and Education Facility at Research Center for Modeling and Simulation, National University of Sciences and Technology for providing a computational facility to complete this study. The authors are also thankful to HEC-National Research Program for Universities of Pakistan for funding this research.

References

1. T. A. Gray and R. D. Nicholls, “Diverse splicing mechanisms fuse the evolutionarily conserved bicistronic MOCS1A and MOCS1B open reading frames,” *RNA*, vol. 6, no. 7, pp. 928–936, Jul. 2000, doi: 10.1017/S1355838200000182.
2. A. Shalata *et al.*, “Localization of a gene for molybdenum cofactor deficiency, on the short arm of chromosome 6, by homozygosity mapping,” *Am. J. Hum. Genet.*, vol. 63, no. 1, pp. 148–154, 1998, doi: 10.1086/301916.
3. J. Reiss, E. Christensen, G. Kurlemann, M. T. Zabot, and C. Dorche, “Genomic structure and mutational spectrum of the bicistronic MOCS1 gene defective in molybdenum cofactor deficiency type A,” *Hum. Genet.*, vol. 103, no. 6, pp. 639–644, 1998, doi: 10.1007/s004390050884.
4. G. Schwarz, R. R. Mendel, and M. W. Ribbe, “Molybdenum cofactors, enzymes and pathways,” *Nature*, vol. 460, no. 7257. Nature Publishing Group, pp. 839–847, Aug. 13, 2009, doi: 10.1038/nature08302.
5. J. L. Johnson, W. R. Waud, K. V. Rajagopalan, M. Duran, F. A. Beemer, and S. K. Wadman, “Inborn errors of molybdenum metabolism: Combined deficiencies of sulfite oxidase and xanthine dehydrogenase in a patient lacking the molybdenum cofactor,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 77, no. 6 I, pp. 3715–3719, 1980, doi: 10.1073/pnas.77.6.3715.
6. B. K. Burgess, “The iron-molybdenum cofactor of nitrogenase,” *Chem. Rev.*, vol. 90, no. 8, pp. 1377–1406, Dec. 1990, doi: 10.1021/cr00106a002.
7. K. V Rajagopalan, “Biosynthesis and processing of the molybdenum cofactors,” 1997, doi: 10.1042/bst0250757.

8. G. Schwarz and R. R. Mendel, “MOLYBDENUM COFACTOR BIOSYNTHESIS AND MOLYBDENUM ENZYMES,” *Annu. Rev. Plant Biol.*, vol. 57, no. 1, pp. 623–647, Jun. 2006, doi: 10.1146/annurev.arplant.57.032905.105437.
9. J. A. Santamaria-Araujo, V. Wray, and G. Schwarz, “Structure and stability of the molybdenum cofactor intermediate cyclic pyranopterin monophosphate,” *J. Biol. Inorg. Chem.*, vol. 17, no. 1, pp. 113–122, Aug. 2012, doi: 10.1007/s00775-011-0835-2.
10. K. T. Shanmugam *et al.*, “Proposed nomenclature for the genes involved in molybdenum metabolism in Escherichia coli and Saimonella typhimurium,” *Mol. Microbiol.*, vol. 6, no. 22, pp. 3452–3454, 1992, doi: 10.1111/j.1365-2958.1992.tb02215.x.
11. M. M. Wuebbens, M. T. W. Liu, K. V. Rajagopalan, and H. Schindelin, “Insight into molybdenum cofactor deficiency provided by the crystal structure of the molybdenum cofactor biosynthesis protein MoaC,” *Structure*, vol. 8, no. 7, pp. 709–718, Jul. 2000, doi: 10.1016/S0969-2126(00)00157-X.
12. P. Hänzelmann and H. Schindelin, “Crystal structure of the S-adenosylmethionine-dependent enzyme MoaA and its implications for molybdenum cofactor deficiency in humans,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 101, no. 35, pp. 12870–12875, Aug. 2004, doi: 10.1073/pnas.0404624101.
13. Tang et al., “□□□□□NIH Public Access,” *Bone*, vol. 23, no. 1, pp. 1–7, 2008, doi: 10.1016/j.ccr.2010.12.003.The.
14. R. Hille, “Molybdenum and tungsten in biology,” *Trends in Biochemical Sciences*, vol. 27, no. 7. Elsevier Current Trends, pp. 360–367, Jul. 2002, doi: 10.1016/S0968-0004(02)02107-2.
15. A. Aukett, M. J. Bennett, and G. P. Hosking, “Molybdenum Co-factor Deficiency: an Easily Missed Inborn Eror of Metaboism,” *Dev. Med. Child Neurol.*, vol. 30, no. 4, pp. 531–535, Nov. 2008, doi: 10.1111/j.1469-8749.1988.tb04781.x.
16. M. P. Coughlan, “The role of molybdenum in human biology,” *J. Inherit. Metab. Dis.*, vol. 6, no. 1 Supplement, pp. 70–77, Mar. 1983, doi: 10.1007/BF01811327.
17. J. Reiss, “Genetics of molybdenum cofactor deficiency,” *Hum. Genet.*, vol. 106, no. 2, pp. 157–163, Feb. 2000, doi: 10.1007/s004399900223.
18. “UniProt: a hub for protein information,” *Nucleic Acids Res.*, vol. 43, no. D1, pp. D204–D212, Jan. 2015, doi: 10.1093/nar/gku989.
19. S. Shamriz and H. Ofoghi, “Design, structure prediction and molecular dynamics simulation of a fusion construct containing malaria pre-erythrocytic vaccine candidate, PfCelTOS, and human interleukin 2 as adjuvant,” *BMC Bioinformatics*, vol. 17, no. 1, p. 71, Feb. 2016, doi: 10.1186/s12859-016-0918-8.
20. “SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments | Bioinformatics | Oxford Academic.” .
21. B. Hess, C. Kutzner, D. Van Der Spoel, and E. Lindahl, “GRGMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation,” *J. Chem. Theory Comput.*, vol. 4, no. 3, pp. 435–447, Mar. 2008, doi: 10.1021/ct700301q.

22. Y. Zhang, “I-TASSER server for protein 3D structure prediction,” *BMC Bioinformatics*, vol. 9, no. 1, pp. 1–8, Jan. 2008, doi: 10.1186/1471-2105-9-40.
23. J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, and Y. Zhang, “The I-TASSER suite: Protein structure and function prediction,” *Nature Methods*, vol. 12, no. 1. Nature Publishing Group, pp. 7–8, Jan. 2014, doi: 10.1038/nmeth.3213.
24. D. Eisenberg, R. Lüthy, and J. U. Bowie, “VERIFY3D: Assessment of protein models with three-dimensional profiles,” *Methods Enzymol.*, vol. 277, pp. 396–404, Jan. 1997, doi: 10.1016/S0076-6879(97)77022-8.
25. A. Messaoudi, H. Belguith, and J. Ben Hamida, “Homology modeling and virtual screening approaches to identify potent inhibitors of VEB-1 β-lactamase,” *Theor. Biol. Med. Model.*, vol. 10, no. 1, p. 22, 2013, doi: 10.1186/1742-4682-10-22.
26. R. A. Laskowski, M. W. MacArthur, D. S. Moss, and J. M. Thornton, “PROCHECK: a program to check the stereochemical quality of protein structures,” *J. Appl. Crystallogr.*, vol. 26, no. 2, pp. 283–291, Apr. 1993, doi: 10.1107/s0021889892009944.
27. J. Pontius, J. Richelle, and S. J. Wodak, “Deviations from standard atomic volumes as a quality measure for protein crystal structures,” *J. Mol. Biol.*, vol. 264, no. 1, pp. 121–136, 1996, doi: 10.1006/jmbi.1996.0628.
28. D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen, “GROMACS: Fast, flexible, and free,” *Journal of Computational Chemistry*, vol. 26, no. 16. John Wiley & Sons, Ltd, pp. 1701–1718, Dec. 2005, doi: 10.1002/jcc.20291.
29. “VMD.pdf.” .
30. S. Yuan, H. C. S. Chan, S. Filipek, and H. Vogel, “Letter PyMOL and Inkscape Bridge the Data and the Data Visualization,” *Struct. Des.*, vol. 24, no. 12, pp. 2041–2042, 2016, doi: 10.1016/j.str.2016.11.012.
31. Y. Y. Ke *et al.*, “Homology modeling of DFG-in FMS-like tyrosine kinase 3 (FLT3) and structure-based virtual screening for inhibitor identification,” *Sci. Rep.*, vol. 5, no. 1, pp. 1–12, Jun. 2015, doi: 10.1038/srep11702.
32. E. F. Pettersen *et al.*, “UCSF Chimera - A visualization system for exploratory research and analysis,” *J. Comput. Chem.*, vol. 25, no. 13, pp. 1605–1612, Oct. 2004, doi: 10.1002/jcc.20084.
33. A. Volkamer, D. Kuhn, F. Rippmann, and M. Rarey, “Dogsitescorer: A web server for automatic binding site prediction, analysis and druggability assessment,” *Bioinformatics*, vol. 28, no. 15, pp. 2074–2075, 2012, doi: 10.1093/bioinformatics/bts310.
34. M. Källberg, G. Margaryan, S. Wang, J. Ma, and J. Xu, “Raptorx server: A resource for template-based protein structure modeling,” *Methods Mol. Biol.*, vol. 1137, pp. 17–27, 2014, doi: 10.1007/978-1-4939-0366-5_2.
35. S. Dallakyan and A. J. Olson, “Small-molecule library screening by docking with PyRx,” *Methods Mol. Biol.*, vol. 1263, pp. 243–250, 2015, doi: 10.1007/978-1-4939-2269-7_19.
36. O. Trott and A. J. Olson, “Software news and update AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading,” *J.*

Comput. Chem., vol. 31, no. 2, pp. 455–461, Jan. 2010, doi: 10.1002/jcc.21334.

37. J. Reiss and R. Hahnewald, “Molybdenum cofactor deficiency: Mutations in GPHN, MOCS1, and MOCS2,” *Hum. Mutat.*, vol. 32, no. 1, pp. 10–18, 2011, doi: 10.1002/humu.21390.
38. E. Yuriev, J. Holien, and P. A. Ramsland, “Improvements, trends, and new ideas in molecular docking: 2012-2013 in review,” *Journal of Molecular Recognition*, vol. 28, no. 10. John Wiley and Sons Ltd, pp. 581–604, Oct. 2015, doi: 10.1002/jmr.2471.

Figures

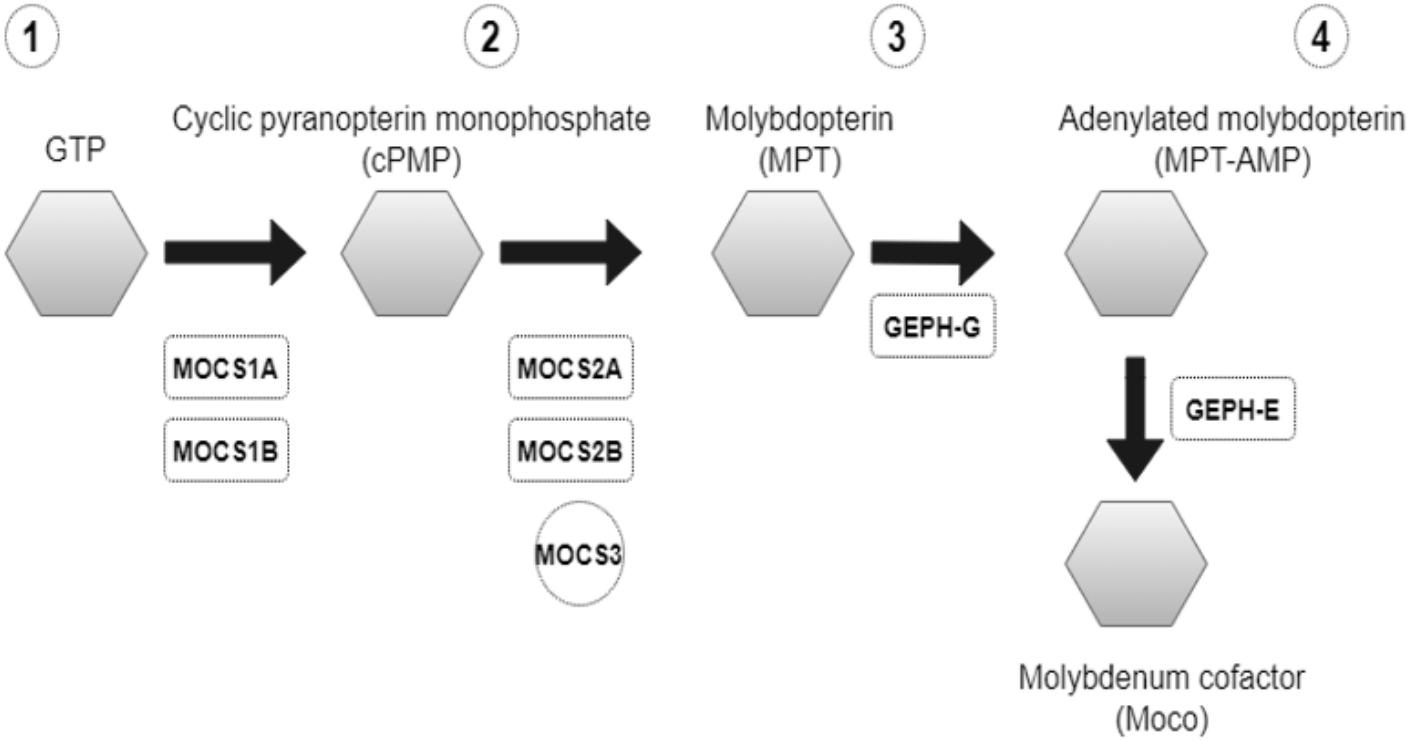

Figure 1

Biosynthesis of Molybdenum Co-factor (Moco).

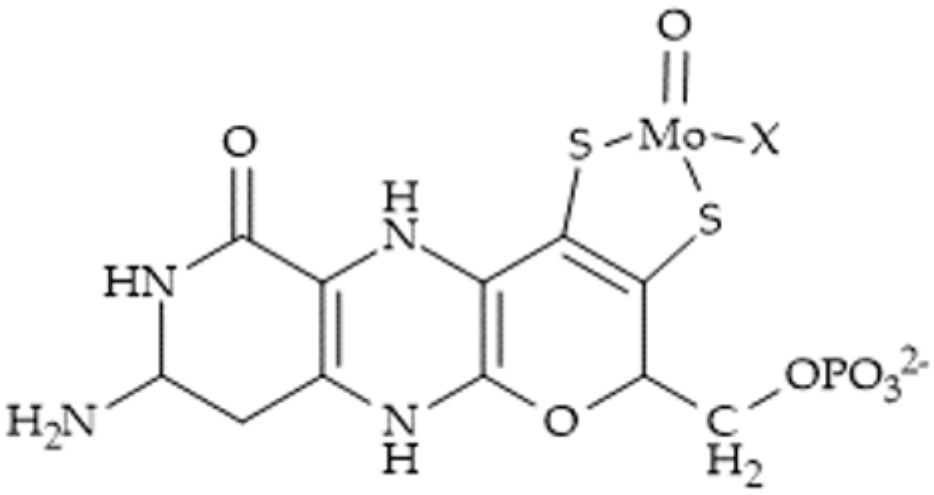

Figure 2

Molybdenum Cofactor

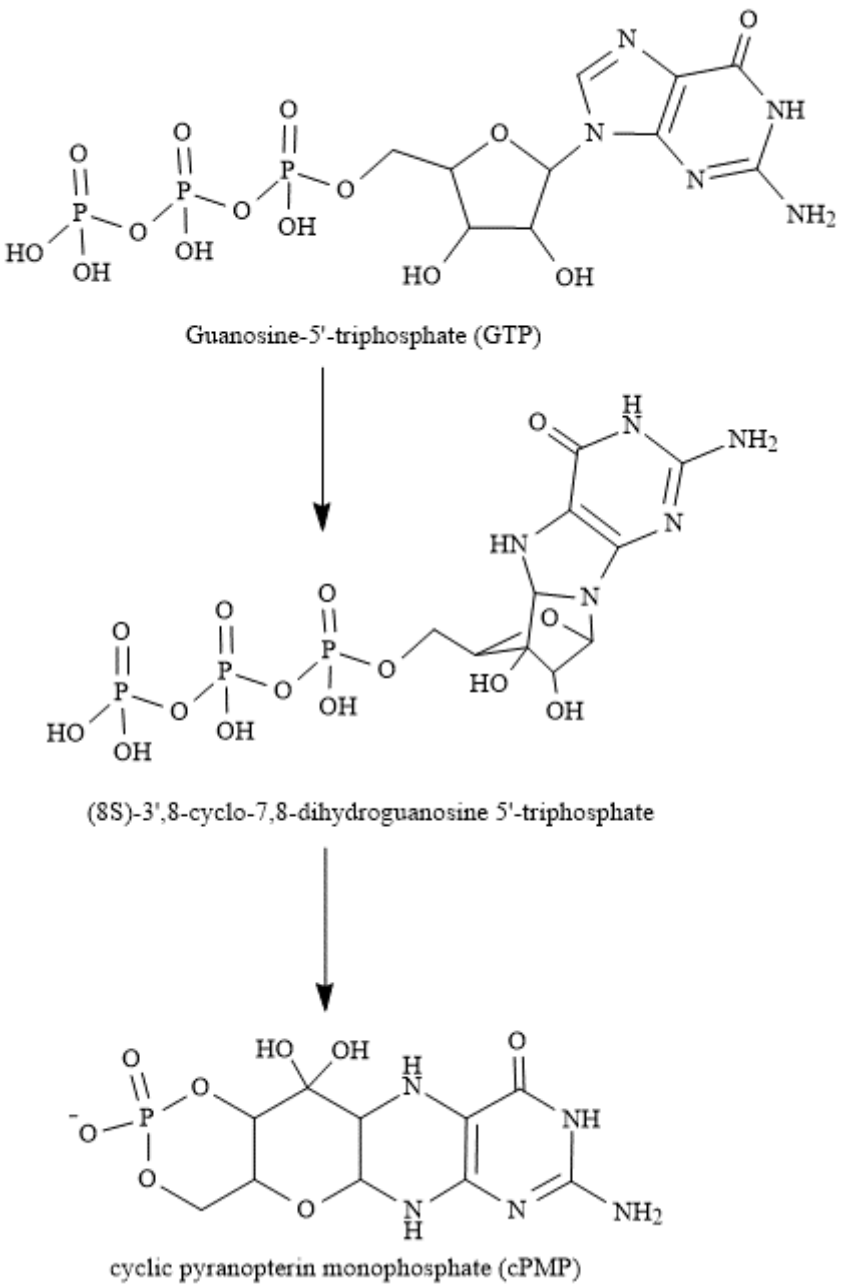

Figure 3

Formation of cPMP from GTP

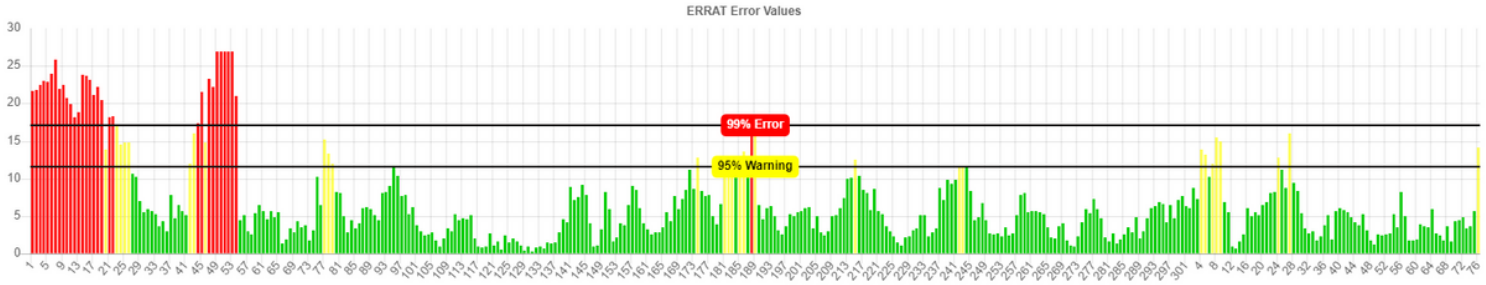

Figure 4

ERRAT plot for MOCS1A

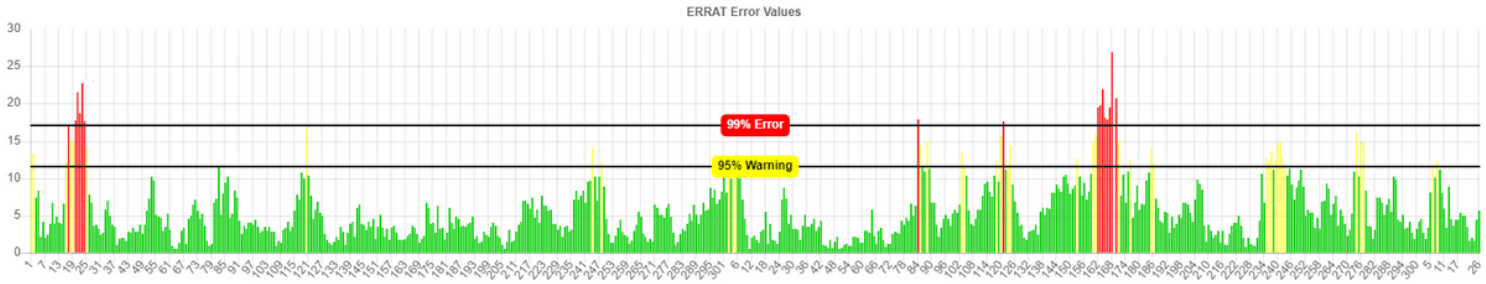

Figure 5

ERRAT plot for MOCS1B

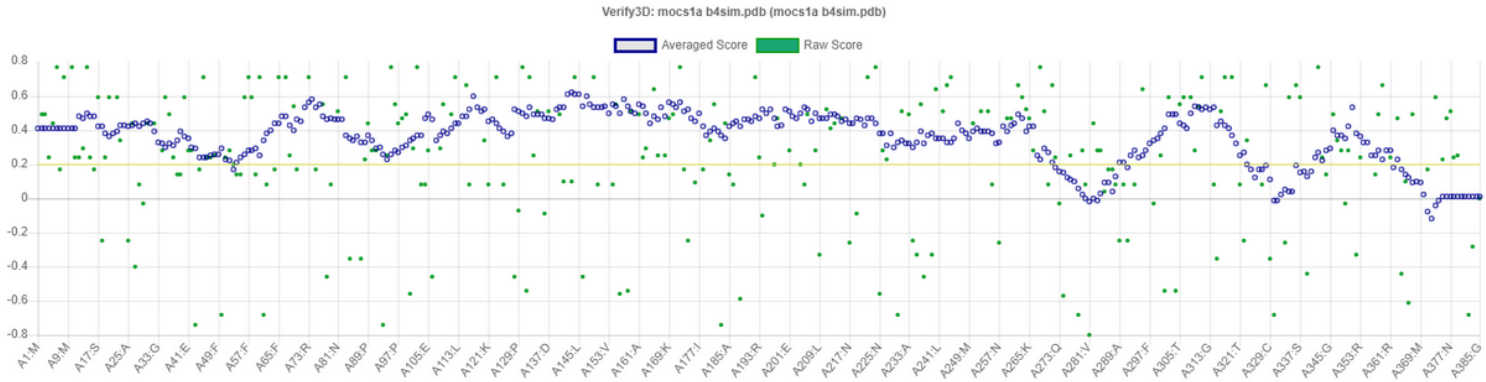

Figure 6

Verify3D score for MOCS1A

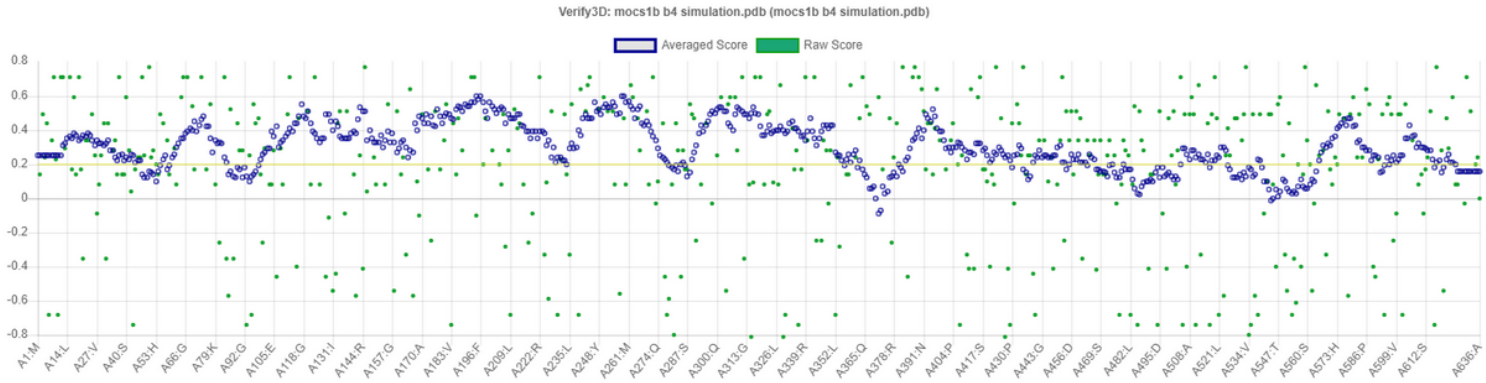

Figure 7

Verify3D score for MOCS1B

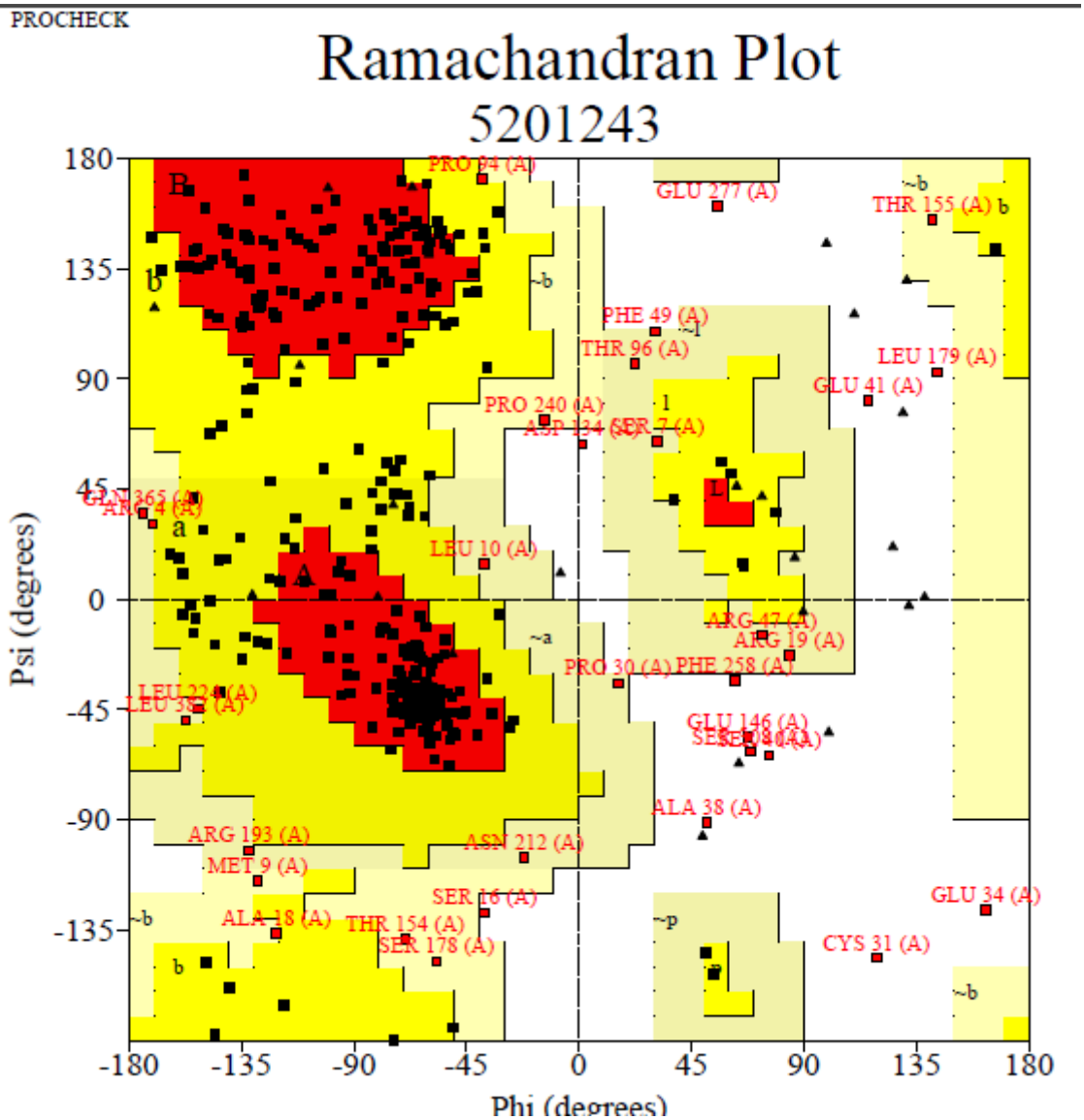

Figure 8

Ramachandran plot for MOCS1A

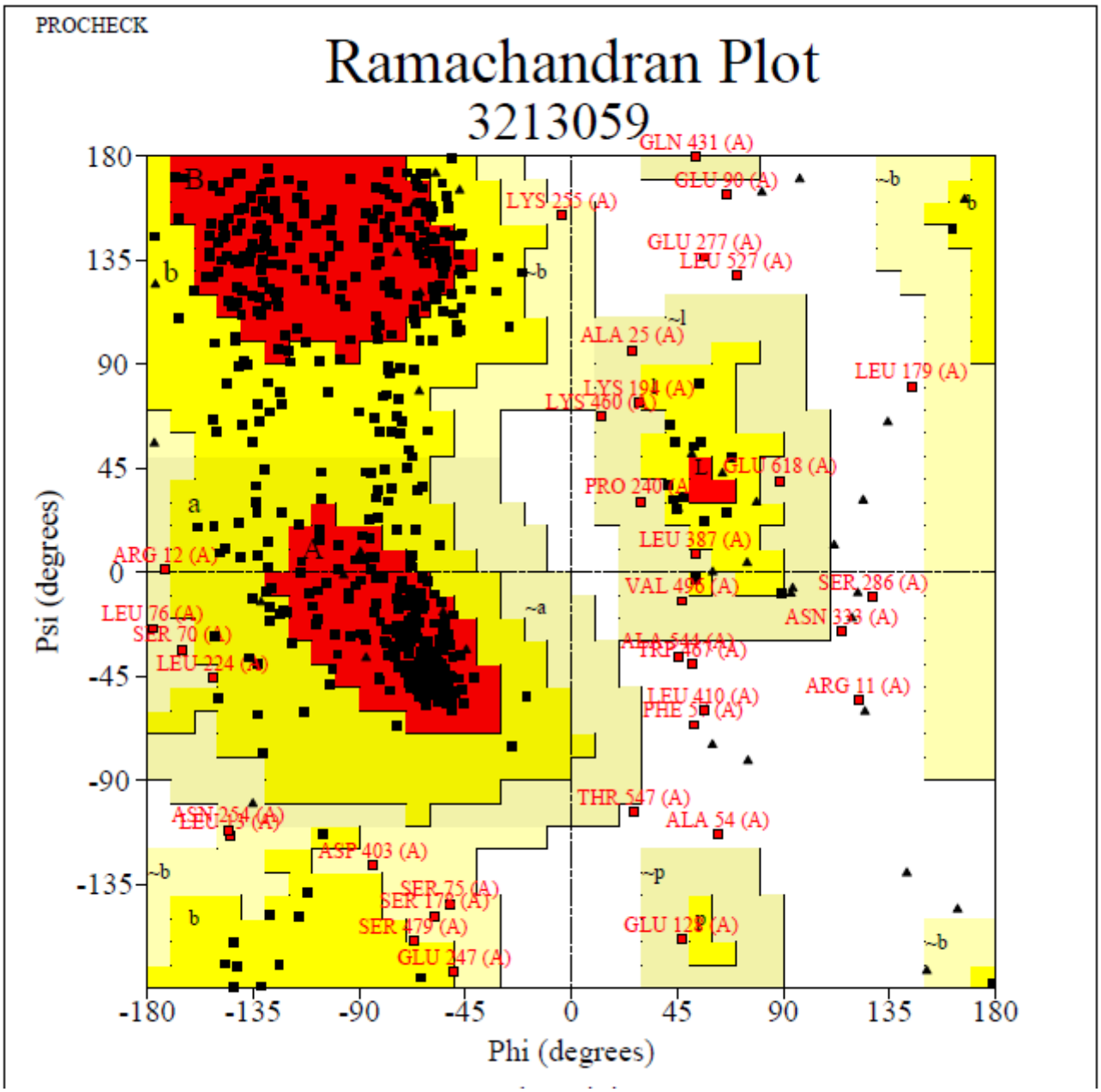

Figure 9

Ramachandran plot for MOCS1B

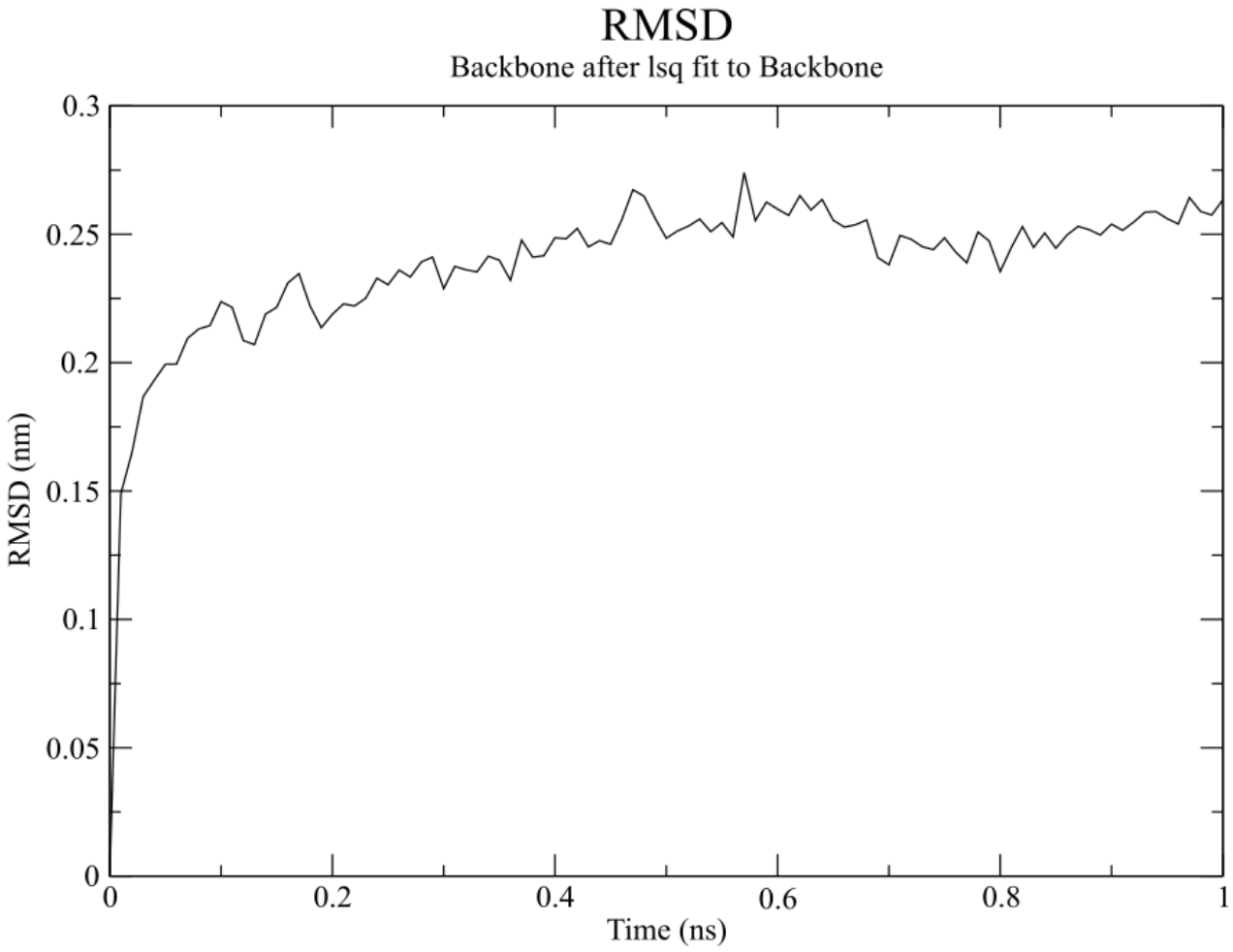

Figure 10

RMSD plot for MOCS1A

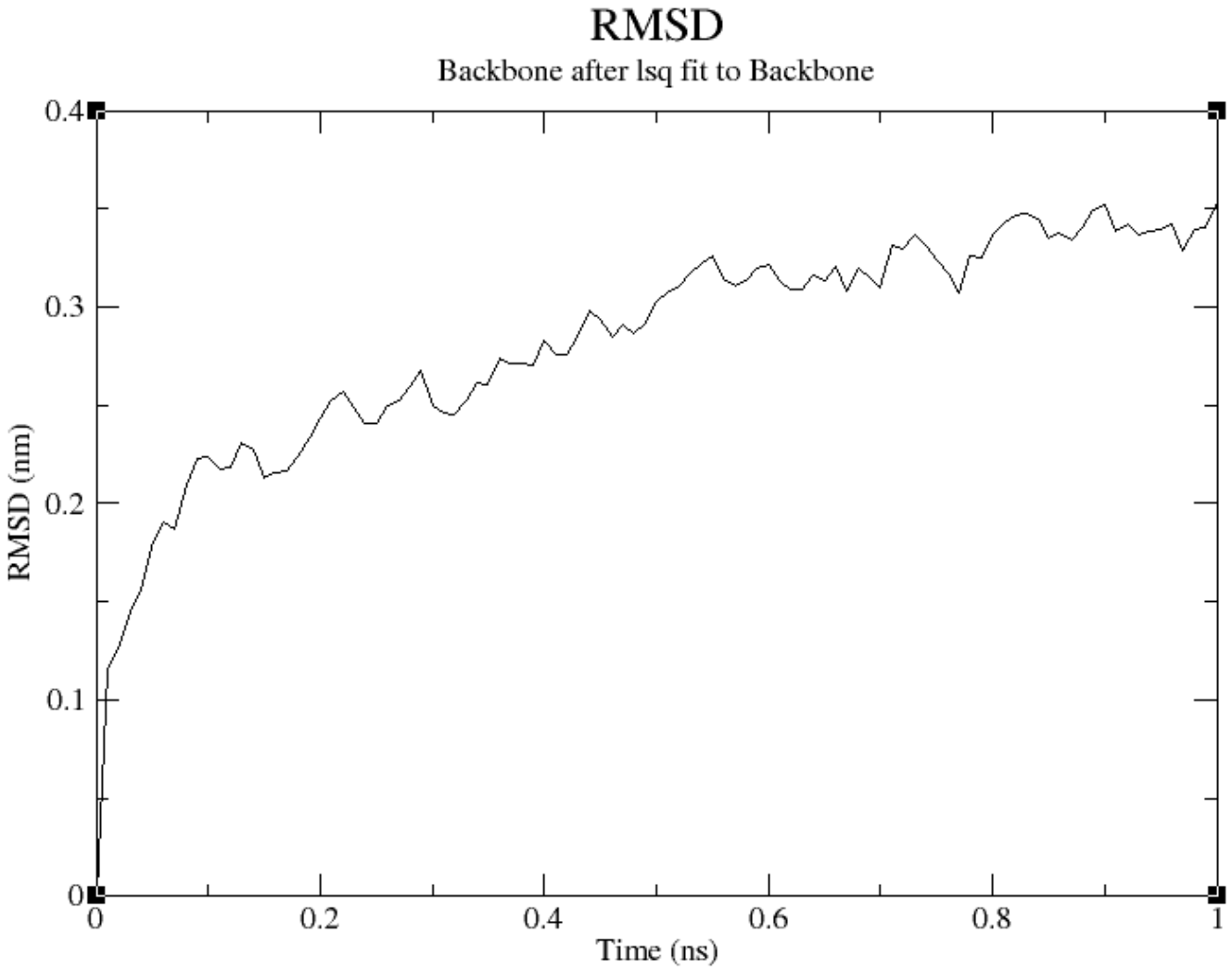

Figure 11

RMSD plot for MOCS1B

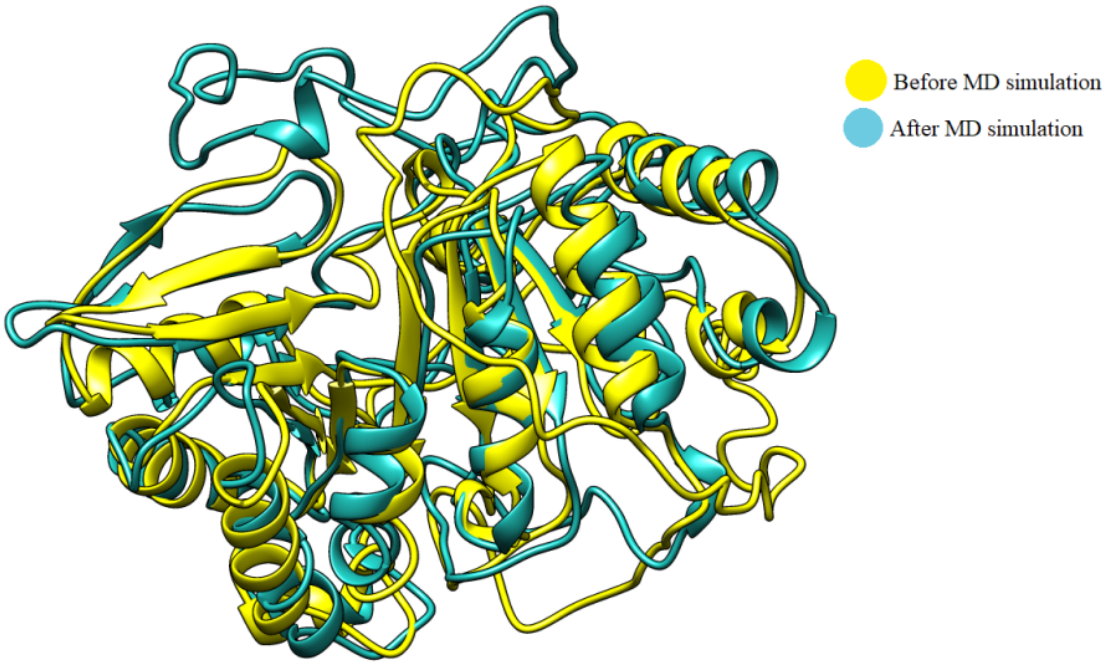

Figure 12

Structure comparison of MOCS1A Pre and Post-MD simulation with chimera

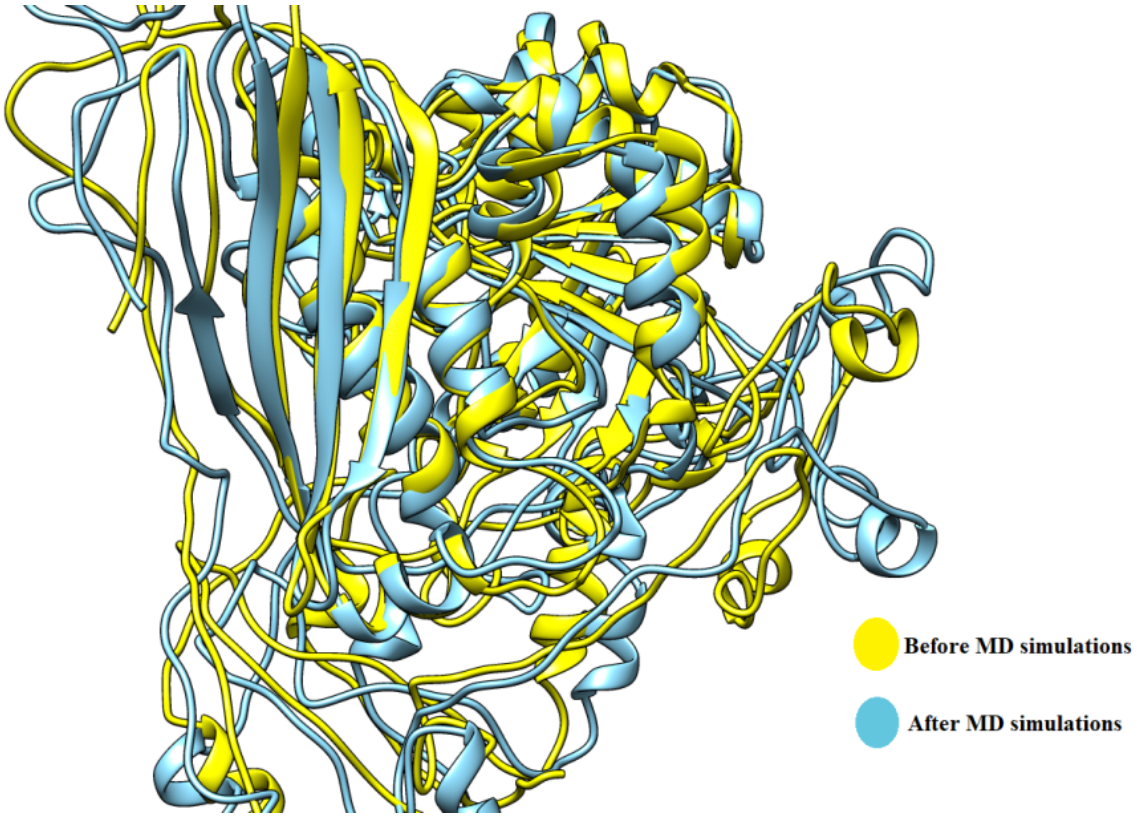

Figure 13

Structure comparison of MOCS1B Pre and Post-MD simulation with chimera

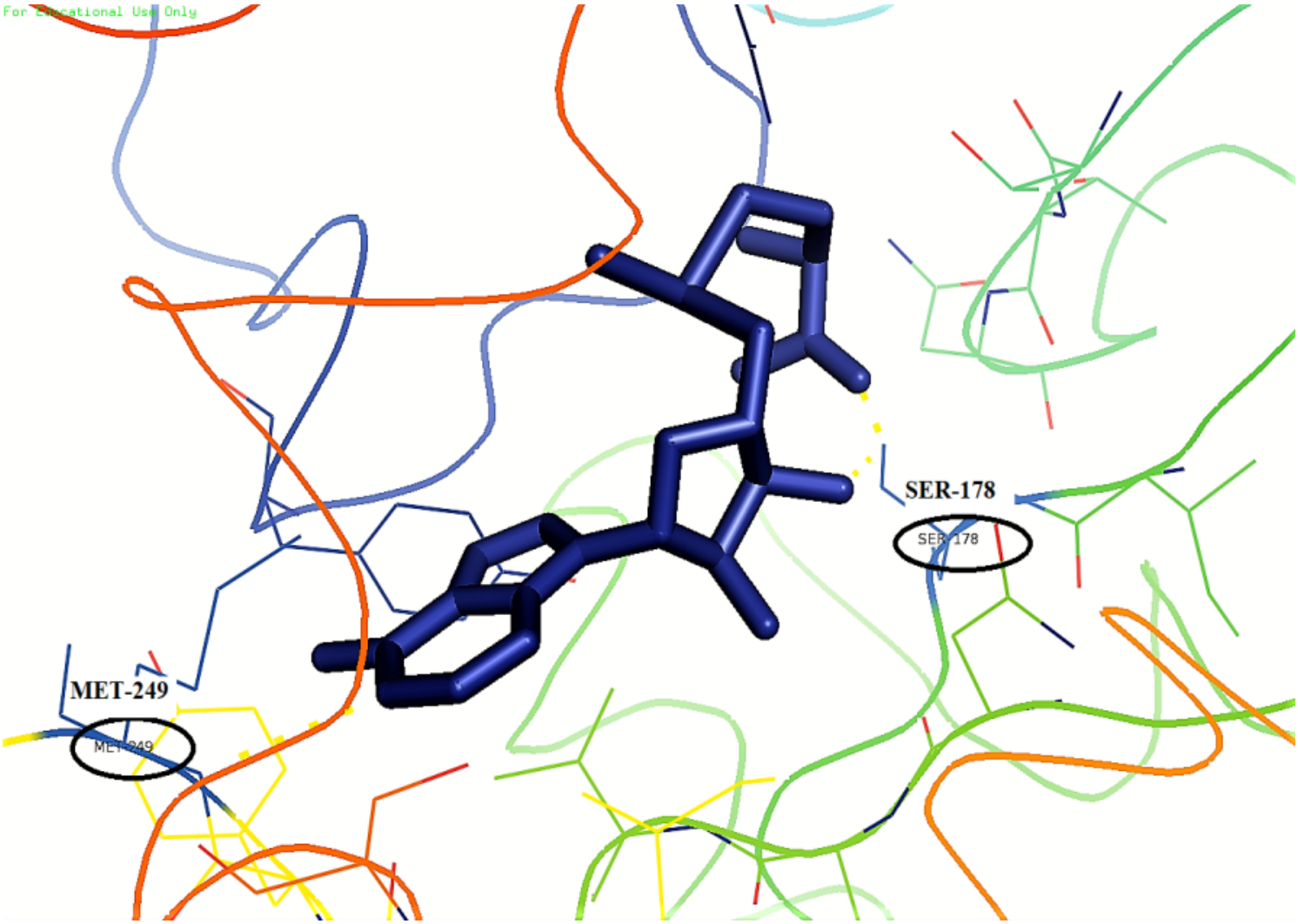

Figure 14

PyMol view of First Binding pocket for GTP obtained from RaptorX

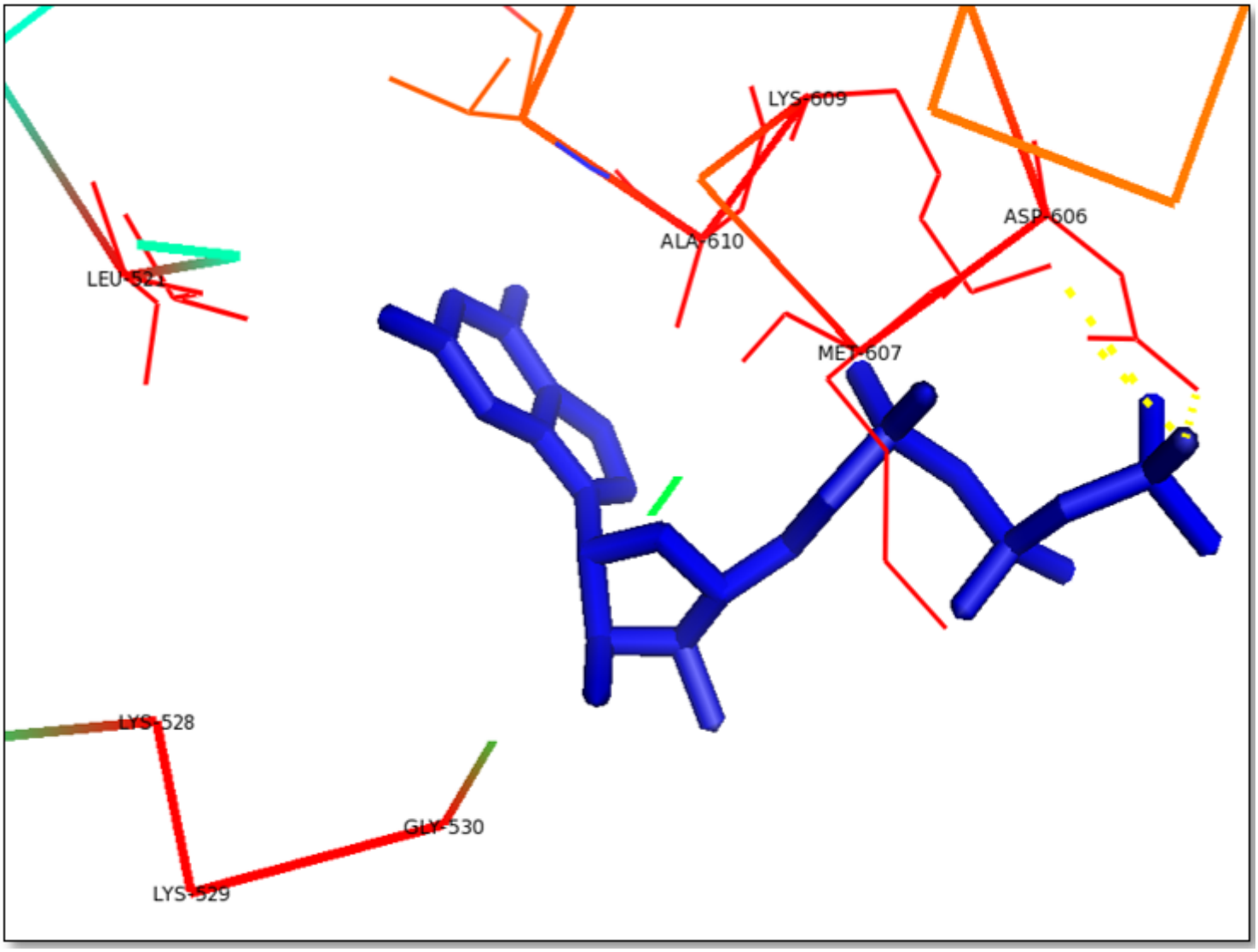

Figure 15

PyMol view of First Binding pocket for 3',8-cyclo-7,8-dihydroguanosine 5'-triphosphate obtained from RaptorX

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterialMOCS1.doc