International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 06 | June - 2022 Impact Factor: 7.185 ISSN: 2582-3930

STUDIES OF PUTATIVE INSULINASE OF *PLASMODIUM FALCIPARUM (3D7)* AS DRUG TARGET FOR *ARTEMISININ*

B. Bagyalakshmi[1], Puniethaa Prabhu[2] , Joram Yam[1*]

Department of Biotechnology Technology, K. S. Rangasamy College of Technology, Tiruchengode – 637 215, Tamil Nadu, India

Abstract

Malaria is one of the most common infectious diseases in tropical and subtropical regions and more than 240 million causes malaria and 627,000 deaths were reported globally in 2022. Among all the malarial parasites, *plasmodium falciparum* is the most virulent and associated with more complications. Now these dangerous parasites have developed resistance against most of its drugs. This review focused on the putative insulinase of PF3D7_1118300 genome from *plasmodium falciparum* as a new target site for *artemisinins*. Putative insulinase 3D structure was downloaded from UniProt. GROMACS 2022.2 package was used to perform MD simulation in water and all atom OPLS forcefield was used. *Artesiminins* was downloaded from PubChem. Active site of protein-ligand was predicted by PyMol and their interaction was performed using AutoDock to determine the binding affinities and interaction between phytochemicals and the putative insulinase of PF3D7_1118300 and BIOVIA discovery studio was used to visualized the binding site. Finally, protein-ligand MD simulation was carried out using the GROMACS 2022.2 package. GROMOS96 forcefield was used for protein to generate the topology and PRODRG 2.5 package was used to generate the ligand topology. Simple point charge water model [SPC216] was used to solvate the protein under cubic box. The system was then minimized energy and equilibrated for 100ps under NVT and NPT using 50,000 steps and then simulation production was done. At the end, results were analyzed by root-mean-square deviation, root-mean-square fluctuation and the radius of gyration and solvent-accessible area.

Keywords: *plasmodium falciparum*, molecular docking, AutoDock, PyMol, GROMACS

1. Introduction

Malaria is one of the most common infectious diseases in tropical and subtropical regions and more than 420 million people are infected and about 627,000 cause deaths globally in 2022 as per the WHO. Among all the malarial parasite, *plasmodium falciparum* is more virulent and associated with more complication [5]. The causes or the mechanism behind these clinical manifestations are still unclear and *plasmodium falciparum* has developed a significant drug resistance against its drug due to which more health problem has been arising [5]. The complete sequenced annotation of the *Plasmodium falciparum* 3D7 genome was done in early 2000[1] and most of the protease has not been characterized. So, this putative insulinase can be an effective target for the *artemisinin*. This drug has shown good results for uncomplicated malaria disease of *P.falciparum* [3].

First target protein was done MD Simulations with water in cubic unit box to activate and to determine the structural stability, functions, and its flexibility [6]. Active binding site between the target protein and small molecules was predicted by PyMol [7] and once site was predicted, molecular docking was done by using AutoDock [8]. Then the output i.e., dlg file was visualized in AutoDock and best conformation was chosen and complex pdbqt file of protein and chosen conformation was created. Discovery studio is a

visualizer tool used for docking, MD simulation etc. here complex pdbqt was visualized using discovery studio to find the active site [9]. Protein-Ligand simulation required separate topology file. GROMOS96 all atom was used to generate protein topology [13], and PRODRG was used for ligand topology [14] and then system was equilibrated for better stability [11] and ligand restrained for further equilibrium under NPT and NVT phase at 300K thus data and trajectories were recorded and visualized.
From the results recorded and visualization we can conclude that this putative insulinase of *plasmodium falciparum* (3D7) can be used as new target for the *artemisinin* phytochemicals.

2.0 Material and Methodology

2.1 Protein selection

Tertiary structure (3D images) of putative insulinase PF3D7_1118300 were obtained from UniProt bearing UniProtKb ID number of Q6III5. This 3D protein structure was predicted by AlphaFold method. Its sequence length is 1488.

2.2 Molecular Dynamic Simulations in water

MD Simulations was done by GROMACS 2022.2. 3D files from UniProt were verified and topology file were generated; topology file contains all the necessary information during whole simulation. It includes all the atom types, charge, bonds, angles, dihedrals etc., A cubic box was constructed and using solvate module, water was filled in it. Then ions were added according to its needed using genion module and energy minimization was done to stabilize the protein for simulation. Then protein equilibration phase was conducted under NVT and NPT with 50,000 steps at 300K temperature in 100-ps. Once the system was well-equilibrated MD simulations were run at 1-ns of 5000000 steps, trajectories image was recorded, and Root Mean Standard Fluctuations (RMSF) was calculated.

2.3 Active Site Prediction by PyMol

Artemisinin was obtained from PubChem which was in sdf file. Then sdf file was converted to pdbqt file in OpenBabel software. The protein and the ligand were uploaded in PyMol tool and by clicking in the bonds between the protein-ligand active site was predicted and the amino acids was visualized.

2.4 Molecular Docking

MGL Tools, AutoDock was downloaded from their respective website and installed. Target protein was prepared in pdbqt file and was read in AutoDock and then edited by adding hydrogens in polar and Kollman charges and saved by giving new name. Then prepared ligand was uploaded and torsion tree was chosen to detect the root and to set the torsion numbers and saved it using new file. Once all the protein and ligand were prepared for AutoDock, grid was set according to the requirements and saved the output as gpf file by clicking GPF. Then docking parameter file dpf by selecting macromolecules as rigid. Using both the gpf and dpf file docking was run using genetics algorithm and LamarkianGA.

2.5 Protein-Ligand Simulations

Protein-Ligand Simulations was done by using GROMOS96 forcefield and protein topology was generated using pdb2gmx module. Ligand topology was generated externally using the PRODRG in which itp and gro file was included. Then cubic unit cell was constructed and solvated using solvate module and required ions was added to neutralized it and topology was always updated. Energy minimization was done and before equilibration ligand was restraint by creating index group in which all are non-hydrogen atoms and then executed the index group to get the porse.itp file of ligand. Added this itp file in the topology file for further equilibration under NPT and NVT phase for 50.000 steps at 300K temperature. Well-equilibrated system was then done MD Simulations for data collection at 1-ns and 5000000 steps. Then analyzed the energy, RMSD, RMSF, and radius of gyration.

3.0 Results

3.1 MD Simulations in water

OPLS-AA/L all atom forcefield was used to construct the topology file for protein and it was placed into cubic unit cell center with 1.0nm away from the box edge (Fig:1) and the system were added 127188 water molecules using solvate module and 47 sodium ions were added to the system to neutralize the net charge (Fig:1) and reached its some physiological concentration. Then system was energy minimized and equilibration phase was done under NVT and NPT at maximum temperature of 300K in 100-ps (Fig:2). Once the temperature and pressure of the system is stabilized, we run the MD production at 1000-ps with 5000,000 steps to collect the data and trajectories results were recorded. Root mean standard deviations (RMSD) and root mean square fluctuation (RMSF) were observed (Fig:3).

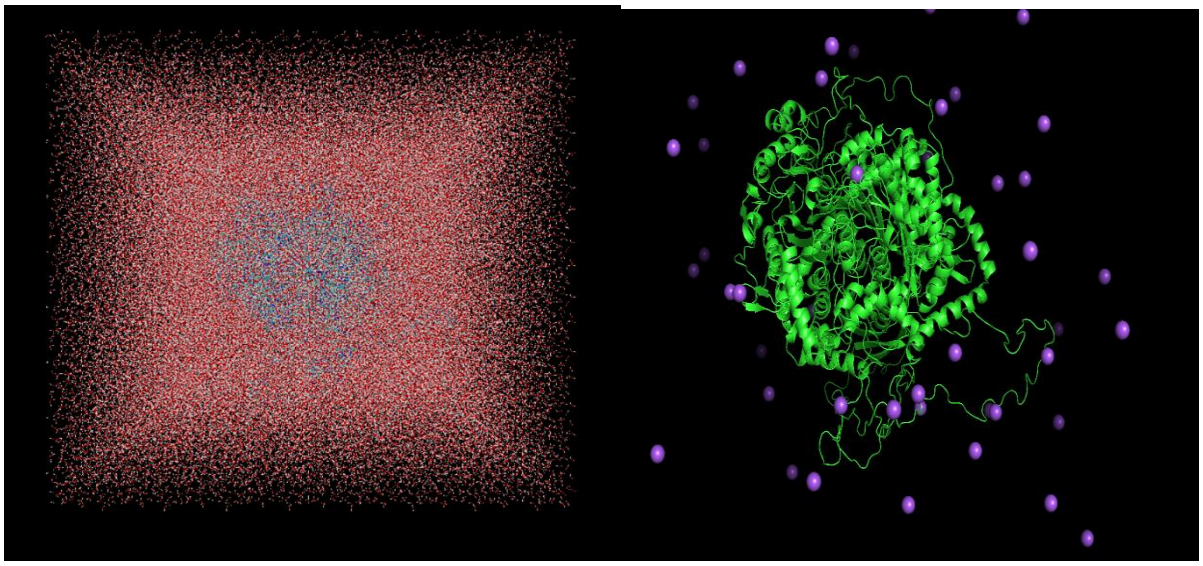

Fig:1. *Protein solvated in cubic box (left image), required ions is added (right image)*

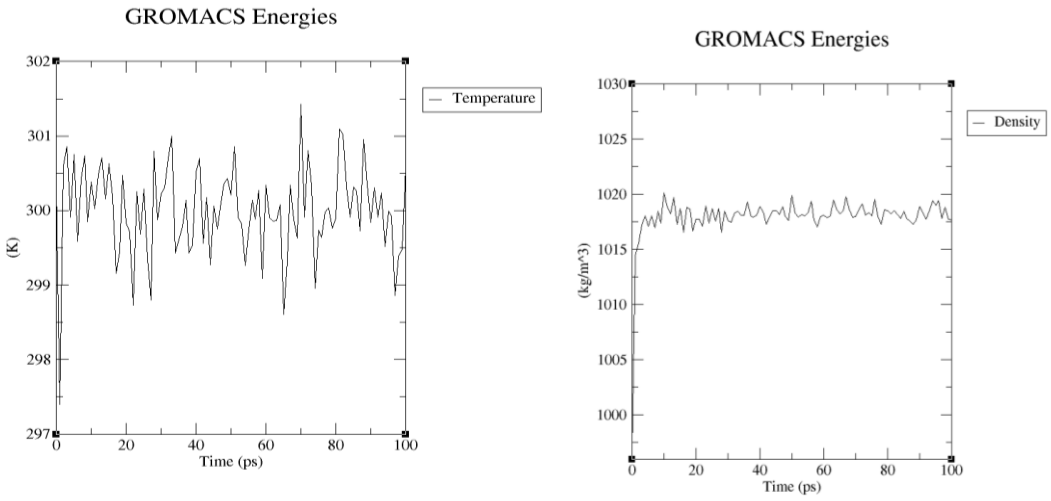

Fig:2. *System temperature and density as a function of time during equilibration*

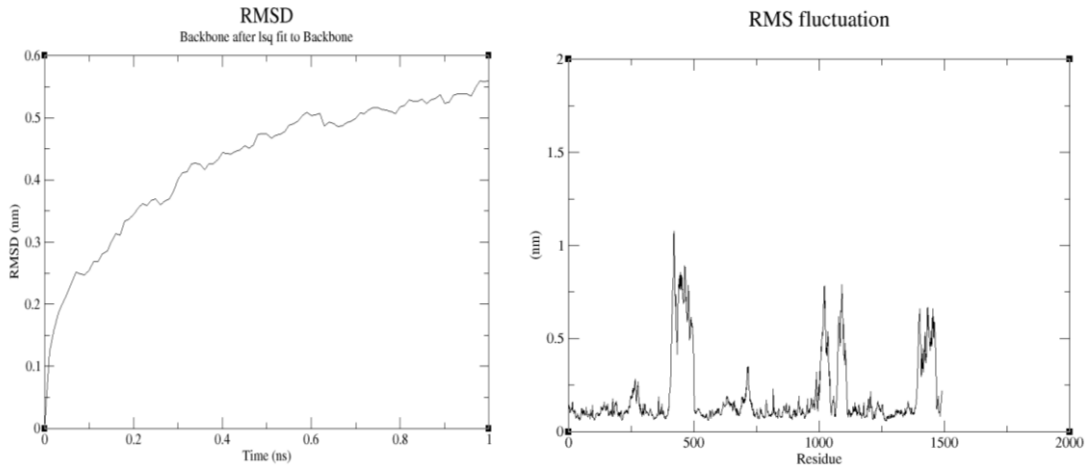

Fig:3. *RMSD and RMSF averaged over 100-ps simulation*

3.2 Active binding site prediction

Both the protein and ligand were uploaded in the PyMol tool and allowed them to interact, then focused on the ligand by clicking present and choose the place in protein where the ligand is bind and later predicted the active binding site (Fig:4). Active binding site are 1053, 1054, 1056, 1058, 1476, 1482, 1486, 1485, 1487, and 1488. After predicting the active binding site, it was visualized in 3D in PyMol (Fig:5).

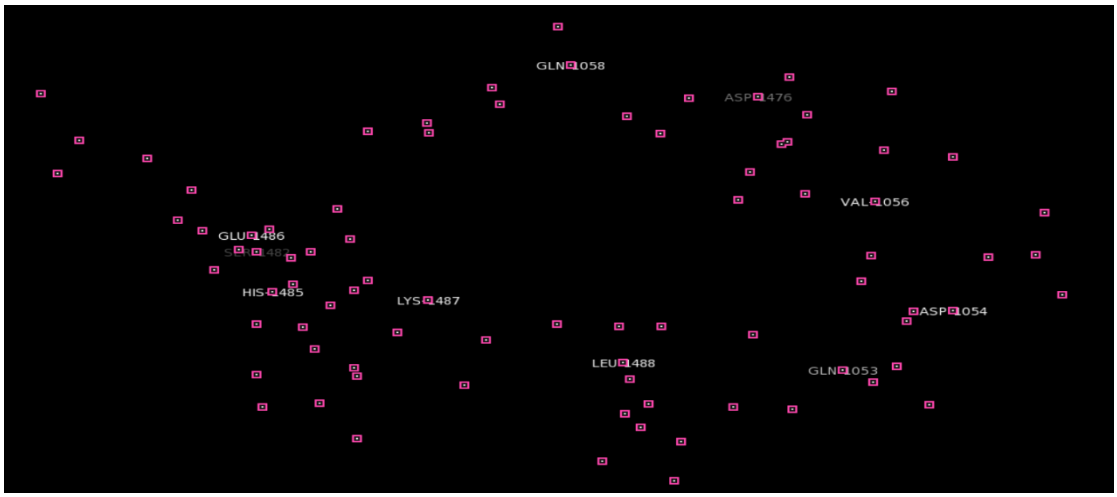

Fig:4. *Active site predicted by PyMol*

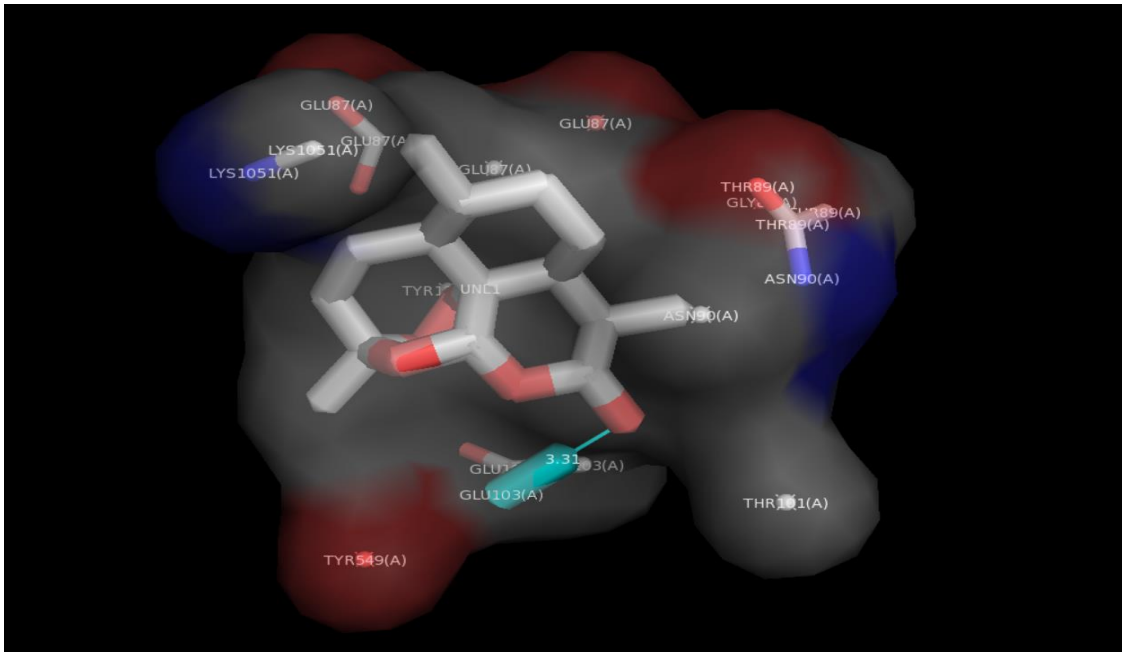

Fig:5. *Showing active site in 3D in PyMol*

3.3 Molecular Docking

Protein-Ligand docking was done in AutoDock adding kollman charge to the protein and creating dlg and glg file with Autogrid4 and Autodock4: dlg file contains all the information like RMSD table, binding affinities, all the binding conformations, ligand charges etc. From the file we came to know that there were 10 binding conformations, and we choose the best conformation with -9.09 binding energy and 0.00 RMSD with 2.00 Argstrom, then this conformation was interacted with protein and saved in pdbqt format for visualization in BIOVIA Discovery studio. Then we visualized the binding site (Fig:6) and found that 1052, 1053, 1054. 1056, 1474, 1475, 1476, and 1488, out of which 1053, 1054, 1056, 1476, and 1488 are the main binding site when compare with the predicted binding site. From the Fig:6 and Fig:7, we have found out the hydrogen bonds between the molecules. The conventional are the hydrogen bond with green color and purple color are the alkyl group as shown in the figure.

We also visualized the measurements between the ligand and the amino acids group.

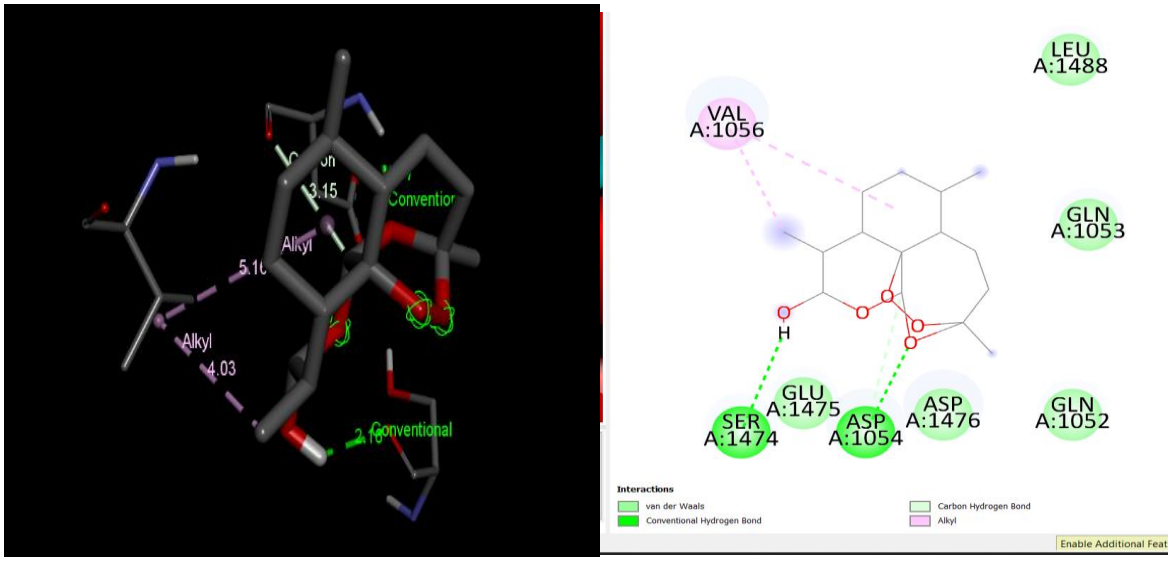

Fig:6. *Protein-Ligand molecular interaction site in 2D*

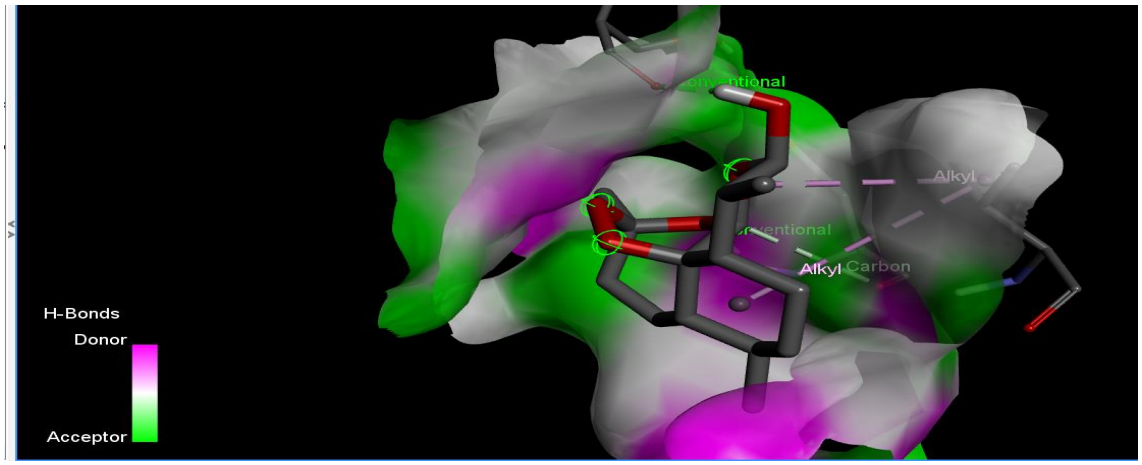

Fig:7. *Visualized at Discovery studio.*

3.4 Protein-Ligand Simulations

Here GROMOS96 all atom force field was used to generate the topology file for protein and PRODRG was used to generate the ligand topology. Once all the required process was done, data collection and trajectories were recorded. **Fig:8** includes all the graph of energy minimization, and you can see that the system has minimized well. It also includes the RMSD and RMSF graph. We can observe that in RMSD, at initial it increases but after 0.7 it stabilizes (**Fig:8:** 3rd graph).

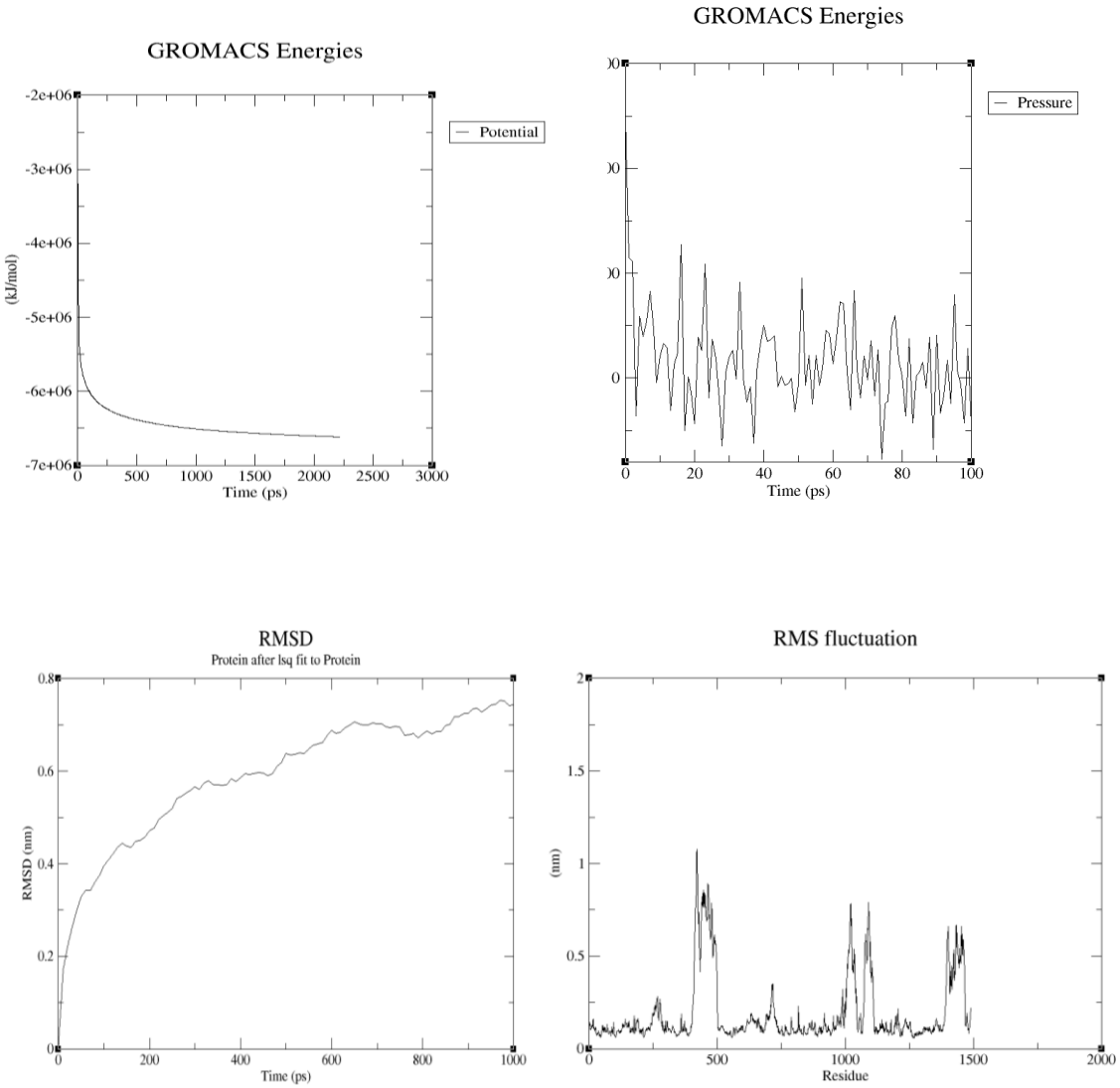

Fig:8. *Trajectories analyzation of Protein-Ligand Interaction.*

4.0 Discussions

From the molecular docking, we found that 1053. 1054, 1056, 147, and 1488 are the active binding site of between the protein and small molecules. MD simulation is a very sensitive process, so we need to do it carefully. Here, we have done protein simulations with water and then with the ligand. We found out that the root mean standard deviation (RMSD) of both increases initially but after some time it got stabilizes. RMSD for protein in water stabilizes in 0.5 and for protein-ligand, it was 0.7. Since this process is very sensitive, your system should be of good like high RAM because everything depends upon the system. Results also stated that the system was well minimized and equilibrated, so which is good. So, we can conclude that the target protein can be of good source for *artemisinin* to treat the malarial parasite *plasmodium falciparum*.

5.0 Conclusion

The aim of this study was to find the active site between the target protein and selected phytochemicals and to study if the protein is potential target site for artemisinin to act on it. Since we have found the active site between them and checked the dynamic properties of them, we can conclude that it would be good try if we used putative insulinase as a new target site for artemisinin drug and we can go for the drug fabrication and then clinical trial.

6.0 Acknowledgement

We like to express my sincere gratitude to Department of Biotechnology, K.S.Rangasamy College of Technology, Tiruchengode Tamil Nadu for providing facilities, interpretation and analysis of data. We would also like to acknowledge the knowledge and support provided by the online tools, servers, and software for successfully completing our work.

References

1. Gardner, Malcolm J., Neil Hall, Eula Fung, Owen White, Matthew Berriman, Richard W. Hyman, Jane M. Carlton et al. "Genome sequence of the human malaria parasite Plasmodium falciparum." *Nature* 419, no. 6906 (2002): 498-511.

2. Miao, J., Lawrence, M., Jeffers, V., Zhao, F., Parker, D., Ge, Y., ... & Cui, L. (2013). Extensive lysine acetylation occurs in evolutionarily conserved metabolic pathways and parasite-specific functions during P lasmodium falciparum intraerythrocytic development. *Molecular microbiology*, *89*(4), 660-675.

3. White, N. J. (2008). Qinghaosu (artemisinin): the price of success. *Science*.

4. Lyu, H. N., Ma, N., Meng, Y., Zhang, X., Wong, Y. K., Xu, C., ... & Wang, J. (2021). Study towards improving artemisinin-based combination therapies. *Natural Product Reports*, *38*(7), 1243-1250.

5. Snounou, G., Viriyakosol, S., Jarra, W., Thaithong, S., & Brown, K. N. (1993). Identification of the four human malaria parasite species in field samples by the polymerase chain reaction and detection of a high prevalence of mixed infections. *Molecular and biochemical parasitology*, *58*(2), 283-292.

6. Zhong, D., Pal, S. K., & Zewail, A. H. (2011). Biological water: A critique. *Chemical Physics Letters*, *503*(1-3), 1-11.

7. Seeliger, D., & de Groot, B. L. (2010). Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *Journal of computer-aided molecular design*, *24*(5), 417-422.

8. Morris, G. M., Huey, R., & Olson, A. J. (2008). Using autodock for ligand-receptor docking. *Current protocols in bioinformatics*, *24*(1), 8-14.

9. Studio, D. (2008). Discovery Studio. *Accelrys [2.1]*.

10. Wang, W., Donini, O., Reyes, C. M., & Kollman, P. A. (2001). Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein-ligand, protein-protein, and protein-nucleic acid noncovalent interactions. *Annual review of biophysics and biomolecular structure*, *30*, 211.

11. Shukla, R., & Tripathi, T. (2020). Molecular dynamics simulation of protein and protein–ligand complexes. In *Computer-Aided Drug Design* (pp. 133-161). Springer, Singapore.

12. Clark, A. J., Tiwary, P., Borrelli, K., Feng, S., Miller, E. B., Abel, R., ... & Berne, B. J. (2016). Prediction of protein–ligand binding poses via a combination of induced fit docking and metadynamics simulations. *Journal of chemical theory and computation*, *12*(6), 2990-2998.

13. Bonvin, A. M., Mark, A. E., & van Gunsteren, W. F. (2000). The GROMOS96 benchmarks for molecular simulation. *Computer physics communications*, *128*(3), 550-557.

14. Schüttelkopf, A. W., & Van Aalten, D. M. (2004). PRODRG: a tool for high-throughput crystallography of protein–ligand complexes. *Acta Crystallographica Section D: Biological Crystallography*, *60*(8), 1355-1363.