



UNIVERSAL WISER
PUBLISHER

Molecular Structure Properties of Heme Group in the Oxymyoglobin Protein (1mbo.pdb) Using PyMOL & UCSF Chimera

Chetanath Neupane^{1*}, Seeram Ramakrishna², Jeevan Jyoti Nakarmi¹, Binod Adhikari¹

¹St. Xavier's College, Kathmandu, Nepal

²Centre of Nanofibers & Nanotechnology, National University of Singapore, Singapore

Email: chetanath.neupane@gmail.com

Abstract: PyMOL and UCSF Chimera are open source multifunctional molecular visualization system developed for the use in structural biology. Both of these tools are well designed with various visualization options available for the user; with the help of such options, we have observed that the iron core binds oxygen in the heme group of oxymyoglobin. Also, various molecular surfaces, internal structures as well as bond lengths between the iron core & molecular oxygen in the heme group and that to the histidine-93 of the oxymyoglobin protein have been calculated. Implementing structural analysis feature of UCSF Chimera, the distance between iron core of heme group and molecular oxygen and that between iron core and histidine 93(H93) have been estimated and found to be 1.827 Å and 2.065 Å respectively which are in good agreement (within 0.5% error) with the bond lengths calculated by x-ray diffraction method^[4]. The small value of bond length between Fe and O₂ suggests the higher stability of oxygen which is important for transportation, activation and storage of enzymes to support life of living organisms on the Earth. So, it provides an insight in enzymology like oxygenase. Interestingly, oxymyoglobin protein or its constituent amino acid chain may have important role in the origin of early life on the Earth.

Keywords: oxymyoglobin protein, PyMOL, UCSF Chimera, visualizing tools, molecular dynamic simulations, molecular structure

1. Introduction

Myoglobin is a protein of mass 17.8 kDa formed by a single polypeptide chain of 153 residues which is found in the muscles of vertebrate animals^[1]. This protein plays a vital role in picking up oxygen from the blood and supplies it to the mitochondria of the cell. It is registered as the first protein having 3D structure determined by x-ray crystallographic method^[2] which shows that the structure is formed by the peptide chain folded into eight helical regions viz; A, B, C, D, E, F, G & H within which heme group is aligned between E and F helices^[3]. Myoglobin can exist in different forms like oxymyoglobin (MbO₂), carboxymyoglobin (MbCO) and metmyoglobin (met-Mb). Oxymyoglobin consists porphyrin ring containing iron core at the centre to which proximal histidine group (His-93) is attached at one face whereas distal histidine group (His-64) to the opposite face^[4]. Among these two histidine group, His-64 is not bonded to the iron core but interacts with oxygen substrate (O₂) which encourages the binding of O₂ due to which the iron core slightly changes its structure by minimizing its radius^[4].

1.1 Protein conformation and molecular dynamics

Proteins perform remarkably large varieties of functions like regulations and control of compounds; from the cellular level to the transportation process by enzymes in animals. Every internal constituent of biomolecules possesses fluctuations which are due to the atomic level conformations hence introducing the biomolecular dynamics called protein dynamics. Protein dynamics defines the time dependent changes in such atomic positions due to both equilibrium and non-equilibrium fluctuations^[5]. Molecular dynamic (MD) simulation is one of the prominent techniques of molecular position tracking in which time evolution of large set of interacting atoms is followed by tracing their trajectory so that their corresponding equation of motion can be understood^[6]. MD provides a framework to track every atoms of the complex system providing information of conformational changes, bond angles, thermodynamic properties and dynamical behavior of a molecule by solving the Newton's equation of motion. Understanding the thermodynamics of the system one can develop various dynamic simulations using variety of algorithms in computer^[7]. Previously, the accuracy and length of the simulations were

Copyright ©2020 Chetanath Neupane, et al.

DOI: <https://doi.org/10.37256/nat.122020425>

This is an open-access article distributed under a CC BY license
(Creative Commons Attribution 4.0 International License)

<https://creativecommons.org/licenses/by/4.0/>

restricted by the computing resources. However, those shortcomings are solved significantly with the developments in computational infrastructures such as parallel computing platforms. Additionally, graphical processing units (GPUs), cell processors and other specialized computer hardware for MD have been designed. To address the compatibility, many popular MD codes are modified to use GPUs to accelerate certain computations. For example, NAMD^[8] has been modified to run on GPU for accelerating explicit solvent and non-bonded interaction computations. The first MD simulation of a protein called Bovine pancreatic trypsin inhibitor (BPTI) with 987 atoms was performed only for 9.2ps^[9] however present simulations of proteins have reached a timescale of up to 1microsecond.

2. Computational details

Structure visualization systems provide us the better understanding of the molecular dynamics, movement of the atoms, their interactions, the trajectories of the atoms and patterns from the given simulation results by relying on specific algorithm. Such tools have their own flexibility level and performance scheme during the computation process and works on the basis of information contained in a PDB file which then translated into the position of atoms in 3D space. Visualization is important for validating new models because it is difficult to test the model through experiments due to their large structure and wide range of interactions involved. Such inadequacy can be minimized by using corresponding features available on visualization systems like PyMOL and Chimera. These two systems have been implemented for the study of heme structures of oxymyoglobin protein.

2.1 PyMOL

PyMOL is an open source, python based and cross-platform system originally created by *Warren Lyford DeLano* which is then largely implemented for the visualization of large and complex macromolecules^[10]. It was designed primarily to visualize dynamically a number of conformations of single structure of molecules by providing graphical representations. It has an in-built feature called click-and-drag which is now reproduced in almost all molecular visualization systems^[11]. PyMOL is capable to display macromolecular electrostatics calculations by APBS^[12] feature that automatically sketches the qualitative electrostatic representation of protein's contact potential by generating a charge-smoothed surface^[12]. PyMOL has change the arena of molecular biology by solving previous frustration mounted with other modeling systems. It has various applications among them most significant are listed below;

- To visualize multiple conformations of a single structure,
- To establish interface with external programs, provide professional strength graphics (under both Windows and Unix),
- To prepare publication quality images.

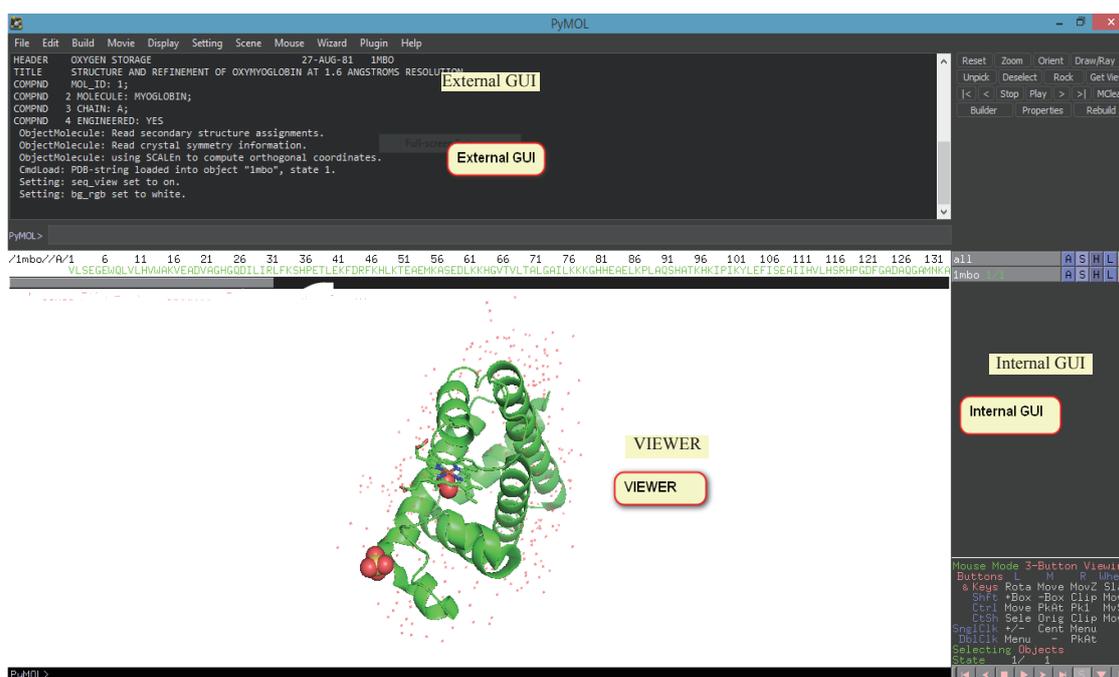


Figure 1. PyMol 2.3.0.0 user interface with loading 1MBO.pdb. GUI is abbreviated form of Graphical User Interface that consists of menus, buttons, text boxes and other options. By default, it has two graphical user interfaces; (1) an "Internal" GUI which used to show molecules (Viewer Window), and (2) an "External" GUI which appears inside of its own window

2.2 UCSF Chimera

Chimera is developed by the *Resources for Biocomputing, visualization, and informatics (RBVI) at the University of California, San Francisco, USA* and available for noncommercial use. Large molecule contains atoms of order of several millions and therefore simple visualization tool cannot display it properly^[13]. It is due to the reason that normal representations are so detailed that they require large computer memory leading ultimately to the insufficient graphics rendering speed^[14].

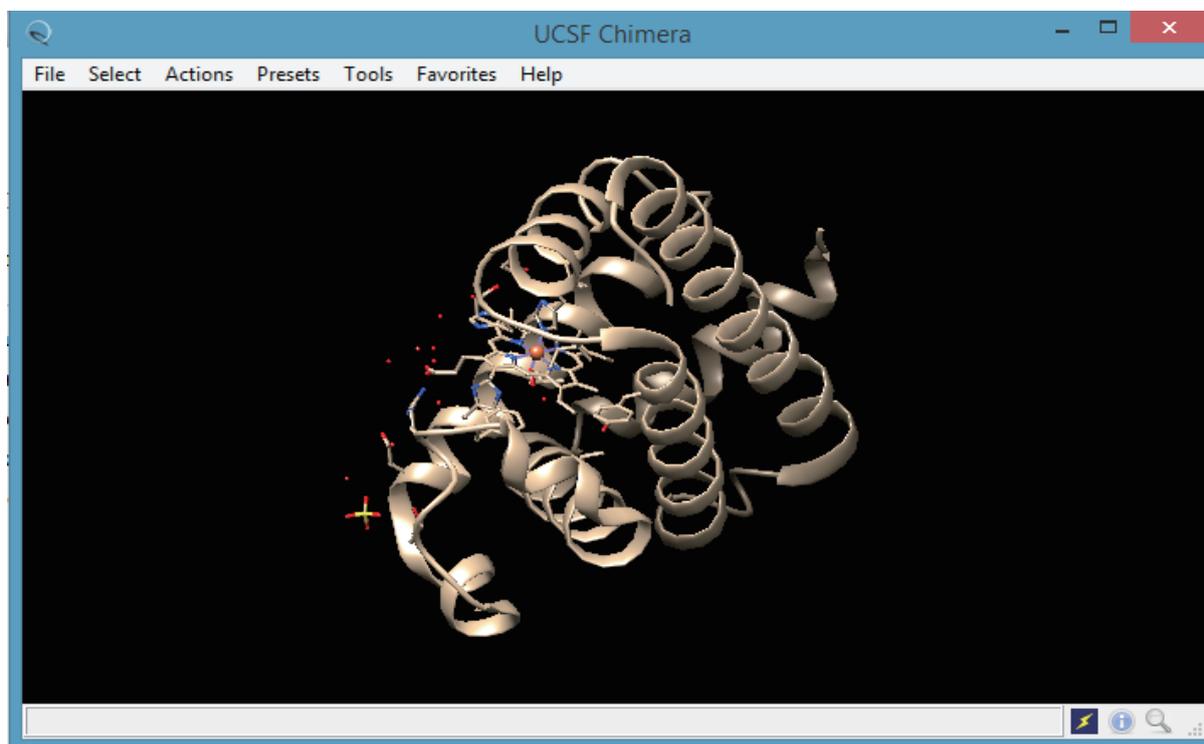


Figure 2. Graphical user interface of UCSF Chimera

GUI and command-line interfaces of chimera provides rich and overlapping sets of functionality. It is available for Windows, Mac OSX and Linux or other UNIX based systems. It generates low-resolution assembly components with high resolution depiction of regions of interest. It offers interactive visualization to interactively explore such large complexes important to have the ability to build multimeric forms with low-resolution representations.

The overall architecture of Chimera contains two parts viz; core and extensions as discussed below:

Core: It is based on C++ programming for basic services and state-of-the-art visualization. Due to the foundation of object oriented programming C++, core allow file input/ output, graphical representations including, wire-frame, ball-and-stick, ribbon, and sphere also for generation of molecular surface using mass spectroscopy algorithm, select parts of structures, control transparency, near and far clipping planes, and lenses.

Extensions: It is designed for extra higher level functionality^[15] to visualize large-scale molecular assemblies to share remotely in real time. Extensions also used for multiple sequence alignments together with associated structures; for displaying and saving MD trajectories as a video etc. It is now started to integrate MODELLER (it is used for homology or comparative modelling of three dimensional structures) and other modeling tools for structural modeling of multiprotein complexes from sequence to 3D structure^[16].

2.3 Comparison between PyMOL and UCSF Chimera

Both Chimera and PyMOL functions are quite equivalent when it comes to adding or removing water, heteroatoms, zooming in on active site, alignment of two structures etc. However, Chimera is easier in script utilization on the one side and PyMOL has a lot of add on features to the other side. Selecting parts of the protein chain seems easier in PyMOL by showing sequence. PyMOL lacks features such as a general “undo” capacity as of now, however, it now has many useful capabilities for the practicing research scientist. In case of compatibility, Chimera reads any file format and offers routes to so many other applications, such as Auto Dock. As to conclude, researcher can probably be comfortable with one over the other based on the research task.

In order to investigate the structure of oxymyoglobin; first we have downloaded its PDB file^[4] from RCSB PDB (<http://www.rcsb.org/>) having following details:

PDB ID: 1MBO.pdb (Structure and refinement of oxymyoglobin at 1.6 angstroms resolution)

DOI: 10.2210/pdb1MBO/pdb

Organism(s): *Physeter cotodon*

Deposition Author(s): Phillips, S.E.V.

3. Results and discussions

a. Using the feature of PyMOL, it is found that the iron atom is connected with the heme group's plane binding with molecular oxygen. Also, it is observed that the iron core of the heme (HEM) group in oxymyoglobin combine reversibly with oxygen as show in figure 3 below.

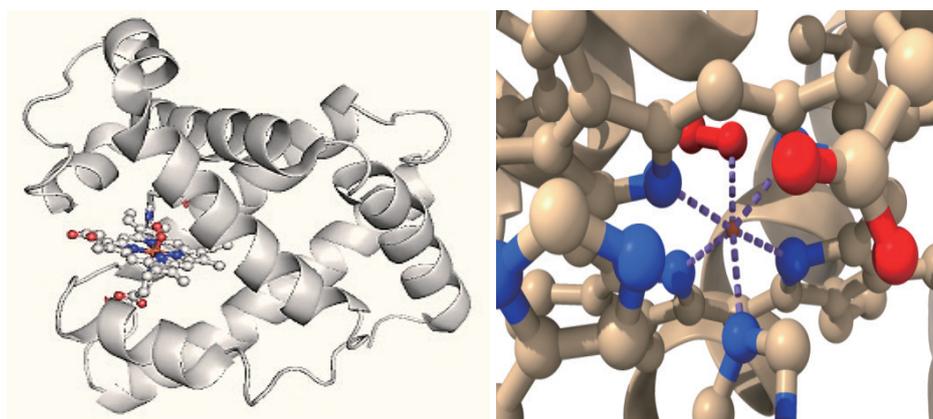


Figure 3. Oxygen binding site in heme of oxymyoglobin observed using PyMol

For the study of coordination environment in the heme group we first turned on the sequence mode of PyMOL which allowed us to select individual amino acids and then selected histidine 64 (H64) & histidine 93 (H93) represented in sticks mode that allowed us to observe heme (HEM) group which is treated as a residue in PyMOL as shown in the 3D graphics in the figure 4 below.

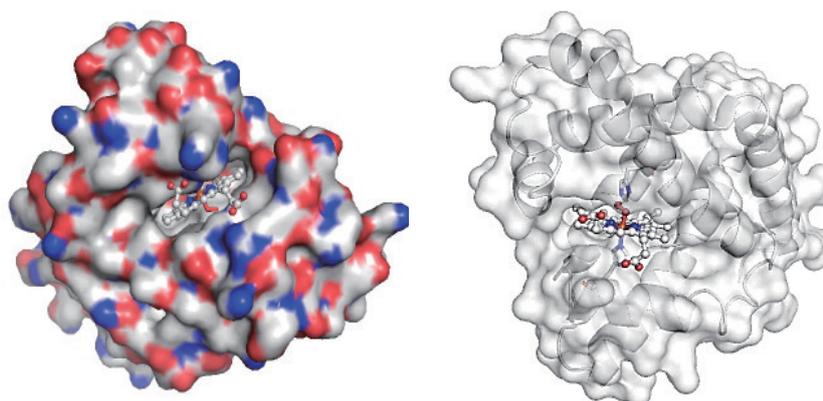


Figure 4. Molecular surfaces of oxymyoglobin drawn by PyMol showing heme group

b. We have implemented UCSF Chimera system for the calculations of bond length between H93 and oxygen of heme group as shown in figure 5 below.

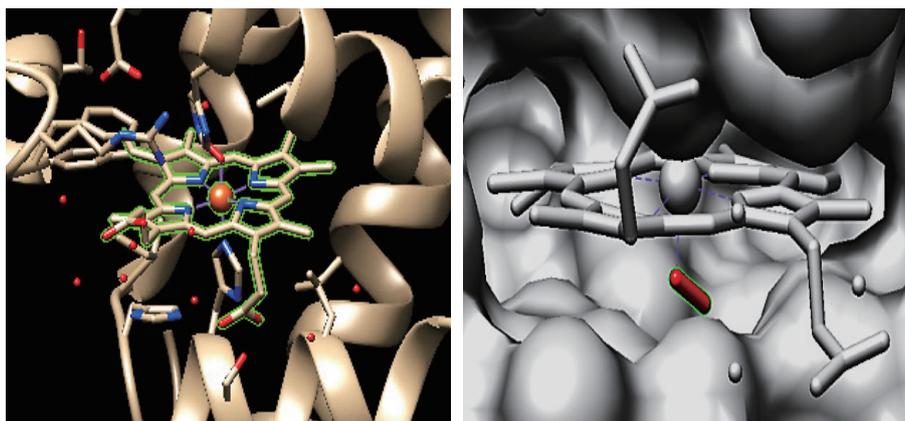


Figure 5. Surfaces of oxymyoglobin highlighting heme group and molecular oxygen using Chimera

c. The distance between iron (in the core of heme) to molecular oxygen and that to the histidine 93(H93) are found to be 1.827 Å and 2.065 Å respectively when calculated by Chimera as shown in figure 6 below.

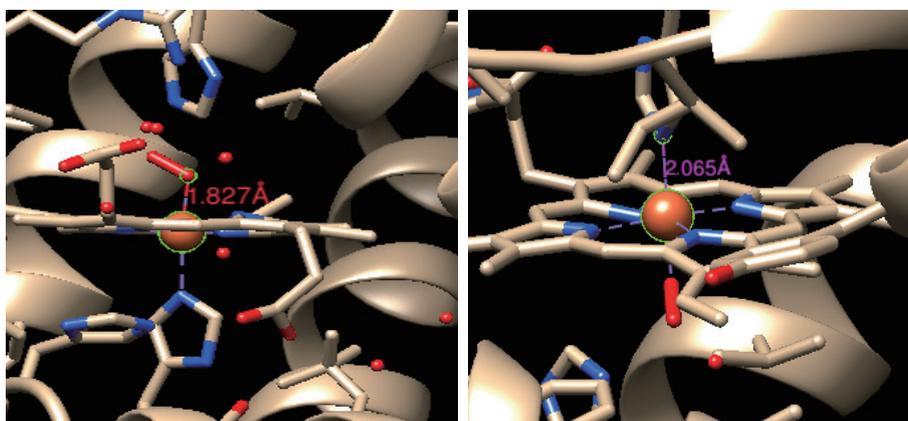


Figure 6. Distance between iron to O_2 and that to the H93 respectively using Chimera

Above numerical values of bond lengths between $Fe - O_2$ and $Fe - H93$ estimated by using Chimera system are in good agreement with that calculated (as $Fe - O_2 = 1.83(6)$ Å & $Fe - H93 = 2.07(6)$ Å) by Phillips et al.^[4] by X-ray diffraction. So, bond length values obtained by structural analysis feature of Chimera are found to be within 0.5% error as that calculated by x-ray diffraction.

4. Conclusions

The small value of bond length between Fe and O_2 suggests the higher stability of oxygen which is important for transportation, activation and storage to support life of living organisms on the Earth. So, it provides an insight in enzymology like oxygenase. Interestingly, oxymyoglobin protein or its constituent amino acid chain may have important role in the origination of early life on earth.

As a concluding remark on visualizing systems, both systems are handful to the researcher interested to work on biophysical domain.

5. Acknowledgements

I thank Prof. Ramakrishna, Prof. Nakarmi and Dr. Adhikari for continuous encouragement and valuable advices.

References

- [1] Lehninger LA, Nelson MM. *Lehninger Principles of Biochemistry*. New York: W. H. Freeman; 2008.
- [2] Kendrew G, Bodo H, Dintzis M, Parrish RG, Wyckoff H, Phillips DC. *Nature*. 1958; 181: 662-666.
- [3] Pettersen EF, Goddard TD, Huang CC. UCSF chimera: a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*. 2004; 25(13): 1605-1612.
- [4] Phillips SE. Structure and refinement of oxymyoglobin at 1.6 Å resolution. *Journal of Molecular Biology*. 1980; 142(4): 531-554.
- [5] Voet D, Voet JG. *Biochemistry*. 4th ed. Hoboken, NJ: John Wiley and Sons; 2011.
- [6] Christen M. The GROMOS software for biomolecular simulation: GROMOS05. *Journal of Computational Chemistry*. 2005; 26(16): 1719-1751.
- [7] Spoel VD, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ. GROMACS: fast, flexible, and free. *Journal of Computational Chemistry*. 2005; 26(16): 1701-1718.
- [8] Phillips JC. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*. 2005; 26(16): 1781-1802.
- [9] McCammon JA, Gelin BR, Karplus M. Dynamics of folded proteins. *Nature*. 1977; 267(5612): 585-590.
- [10] DeLano WL. PyMOL: an open-source molecular graphics tool. *CCP4 Newsletter on Protein Crystallography*. 2002; 40: 82-92.
- [11] Brunger AT, Wells JA, Warren L, Obituary. *Nature Structural & Molecular Biology*. 2009; 16(1): 1202-1203.
- [12] Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences of the USA*. 2001; 98(1): 10037-10041.
- [13] Goddard TD, Huang CC, Ferrin TE. Visualizing density maps with UCSF Chimera. *Journal of Structural Biology*. 2007; 157(1): 281-287.
- [14] Goddard TD, Huang CC, Ferrin TE. Software extensions to UCSF chimera for interactive visualization of large molecular assemblies. *Structure*. 2005; 13(3): 473-482.
- [15] Pettersen EF, Goddard TD, Huang CC. UCSF Chimera-a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*. 2004; 25: 1605-1612.
- [16] Yang Z, Lasker K. UCSF Chimera, MODELLER, and IMP: An integrated modeling system. *Journal of Structural Biology*. 2012; 179: 269-278.