BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 10(1)

In silico analysis of Myoglobin in *Channa striata*

Farzana Parveen†* & Vineet Kumar Mishra†

Department of Biotechnology, Faculty of Science, Jamia Hamdard, New Delhi, India; Farzana Parveen - Email: fparveen.jh@gmail.com; Phone: +919560171436; *Corresponding author
†- Authors equally contributed

Received January 06, 2014; Accepted January 17, 2014; Published January 29, 2014

Abstract:
Myoglobin is a cytoplasmic hemoprotein, expressed solely in cardiac myocytes and oxidative skeletal muscle fibers, that reversibly binds O_2 by its heme residue. Myoglobin is an essential oxygen-storage hemoprotein capable of facilitating oxygen transport and modulating nitric oxide homeostasis within cardiac and skeletal myocytes. Functionally, myoglobin is well accepted as an O_2-storage protein in muscle, capable of releasing O_2 during periods of hypoxia or anoxia. There is no evidence available regarding active sites, ligand binding sites, antigenic determinants and the ASA value of myoglobin in *Channa striata*. We further document the predicted active sites in the structural model with solvent exposed ASA residues. During this study, the model was built by CPH program and validated through PROCHECK, Verify 3D, ERRAT and ProSA for reliability. The active sites were predicted in the model with further ASA analysis of active site residues. The discussed information thus provides the predicted active sites, ligand binding sites, antigenic determinants and ASA values of myoglobin model in *Channa striata.*

Keywords: Validation, Active Sites Prediction, Ligand Binding Sites, Antigenic Determinants and ASA analysis.

Background:
Myoglobin is a cytoplasmic hemoprotein expressed solely in cardiac myocytes and oxidative skeletal muscle fibers **[1].** Hemoproteins are widely distributed among prokaryotes, unicellular eukaryotes, plants and animals **[2].** Myoglobin, a mobile carrier of oxygen, is developed in red muscle in response to mitochondrial demand for oxygen and transports oxygen from the sarcolemma to the mitochondria of vertebrate heart and red muscle cells **[3].** Myoglobin was so named because of its functional and structural similarity to hemoglobin **[4].** Myoglobin facilitates oxygen transport from the erythrocyte to the mitochondria in cardiomyocytes and oxidative skeletal myofibers **[2].** Under some circumstances it also may provide a reserve of oxygen after interruption in oxygen delivery **[5].** Skeletal muscle myoglobin concentration is positively and significantly correlated with dive duration in some species **[6].** Myoglobin concentration in skeletal muscle is also increased in humans and other species living at high altitude **[1].** Calculations of oxygen flux by simple diffusion suggest that, in the absence of myoglobin, oxygen pressure at the centre of myocytes working in normal oxygen (normoxic) levels would fall to levels insufficient to maintain mitochondrial oxidative phosphorylation **[4].** In support of this prediction, contractile function of cardiac or skeletal muscles is sensitive to chemical agents (nitrite, hydroxylamine or hydroperoxides) or carbon monoxide administered to reduce effective concentrations of oxymyoglobin **[7].**

These studies utilized chemical agents such as carbon monoxide or sodium nitrite to inactivate myoglobin. These agents function either in the binding of the heme domain of myoglobin (i.e. carbon monoxide) or in the oxidation of myoglobin (i.e. sodium nitrite) that inhibits the binding to oxygen. Use of these chemical agents to inhibit myoglobin function results in a reduction of oxygen consumption, adenosine triphosphate synthesis and muscle contractility **[2].** Although myoglobin does not cross the mitochondrial membrane, a process termed myoglobin-mediated oxidative phosphorylation have proposed **[8].** This is distinguished from a role for myoglobin bulk transport of oxygen or as a reservoir

for oxygen on the premise that reducing or oxidizing equivalents transduce the effect of oxymyoglobin across the mitochondrial membrane and enhance mitochondrial oxidative phosphorylation. Myoglobin may also function as a nitric oxide (NO) scavenger **[9].** In beating heart and exercising skeletal muscle, myoglobin acts as a short-term oxygen store (i.e. an oxygen buffer). The high concentration of skeletal myoglobin in aquatic mammals was thought to enable a long-term oxygen supply during diving **[10].** Essentially all of the oxygen consumed by skeletal muscle and heart is taken up by cytochrome oxidase. Myoglobin is developed in skeletal muscle more or less in proportion to the cytochrome oxidase content of the muscle **[3].** Since myoglobin cannot cross the mitochondrial outer membrane, cytochrome oxidase, located in the inner mitochondrial membrane and cristae, must be supplied by dissolved oxygen diffusing from the sarcoplasm. The exceedingly thin mitochondrial outer membrane will scarcely impede oxygen diffusion **[11].** Fully reduced cytochrome oxidase combines very rapidly with oxygen to form an oxygenated intermediate **[12].**

Cytochrome oxidase is the major route by which NO is removed from mitochondria-rich cells **[6].** Related to its role as a tissue reservoir of O_2, myoglobin has been proposed to also serve as a buffer of intracellular PO- in a number of species including the human, rodent and bovine models. Similar to the role of creatine phosphokinase, which functions to buffer ATP concentrations when muscle activity increases, myoglobin functions to buffer O_2 concentrations under similar conditions. Myoglobin desaturates rapidly at the onset of muscle activity, increasing the PO• gradient from capillary blood to cytoplasm. Furthermore, it has been proposed that desaturated myoglobin close to the cell membrane then binds O_2 and diffuses to the mitochondria, providing a parallel path that supplements simple diffusion of dissolved O_2. NO has been shown to inhibit cytochrome oxidase and thus impair mitochondrial respiration **[1].** Since, there is no adequate information related to the active site, ligand binding sites, antigenic determinants and the nature about myoglobin in *Channa striata* (snakehead murrel). So, this study deals with the identification of active site, ligand binding sites, antigenic determinants and accessible surface area analysis in protein model of *Channa striata*.

Methodology:

Template selection from RCSB PDB and isolation of myoglobin

The Protein- Hemoglobin was isolated from RCSB PDB **[13]** as template and myoglobin was isolated by comparing against a complete Non-Redundant Protein Database using The NCBI PSI-BLAST with 27% similarity of the template ensuring that those protein in the first step truly had no myoglobin in *Channa striata*.

Visualization and Model Validation

The protein structure was predicted by homology modeling with different server CPH model and SAVES (http://www.cbs.dtu.dk /services/CPHmodels, http:// nihserver.mbi.ucla.edu/SAVES/). The program CPH models 3.0 server **[14]** was used to build the model as PDB file of myoglobin according to the homology modeling method. The PDB of myoglobin was visualized by using PyMOL version 1.3 (http://pymol.sourceforge.net/) **[15].** The structure which was obtained from homology modeling was validated by SAVES. The steriochemical quality of the model was verified with program PROCHECK **[16]** in order to select best model. 3D-profiling of the residue was done by VERIFY 3D Structure evaluation Server **[17].** ERRAT was used for verifying protein structure for evaluating the progress of crystallographic model building and refinement **[18].** The model was submitted to the ProSA-WEB **[19]** to obtain the Z-Score.

Prediction of Protein Structure

Secondary structure of the protein was predicted through SOPMA program (Self-Optimized Prediction Method), the program determined individual role of amino acid for building the secondary structure with their positions (http:// npsapbil.ibcp.fr /cgibin/npsa_automat.pl?page =/NPSA /npsa_sopma.html) **[20].**

Prediction of active site

One of the keenest areas in the bioinformatics is the active sites of the protein and its residual identification. The structure was used to find out the ligand binding efficiency with the predicted model and to detect whether the protein was better or weak with other reported sources. "CASTp" was used for identification of active site from the protein structure **[21].**

Predicting of Antigenic Peptides

This program predicts those segments from within a protein sequence that are likely to be antigenic by eliciting an antibody response. Antigenic peptides are determined using the method of Kolaskar and Tongaonkar. Predictions are based on a table that reflects the occurrence of amino acid residues in experimentally known segmental epitopes (http:/ /imed.med.ucm.es/Tools/antigenic.pl) **[22].**

Prediction of ligand site

3DLigandSite is an automated method for the prediction of ligand binding sites. The structure is used to search a structural library to identify homologous structures with bound ligands. These ligands are superimposed onto the protein structure to predict a ligand binding site. "3DLigandSite" was used for identification of ligand binding site from the protein structure **[23].**

Discussion:

Template selection from RSCP PDB and isolation of absent myoglobin

The Hemoglobin from the *Gasterophilus intestinalis* (PDB ID: 2C0K) was selected from RCSB PDB and was compared against complete Non-Redundant Protein Database using The NCBI PSI-BLAST and myoglobin in *Channa striata* with 27% similarity of template which is absent in PDB, is used as subject for study (Accession: AGG38019.1)

Visualization and Model Validation

The three dimensional structure of myoglobin was generated by using PyMOL program **(Figure 1).** To verify the predicted structure, validation was carried out with PROCHECK program. Ramachandran plot of non-glycine and non-proline residue in the structure showed that 94.5% of the total amino acids were presented in most favored regions and the other 5.5% of amino acids were presented in allowed regions including disallowed region with 0.0%. VERIFY_3D shows

99.32% of the residues had an averaged 3D-1D score greater than 0.2 indicates that the environment profile of the model is good. ERRAT2 shows 99.275% overall quality factors indicating good resolution structure. Moreover, quality of the model can be compared to reference structure of high resolution obtained from X-Ray crystallography analysis through Z score and "0" is the average Z score for good model. The Z score of myoglobin is -7.8 showing the possibility to be a better model.

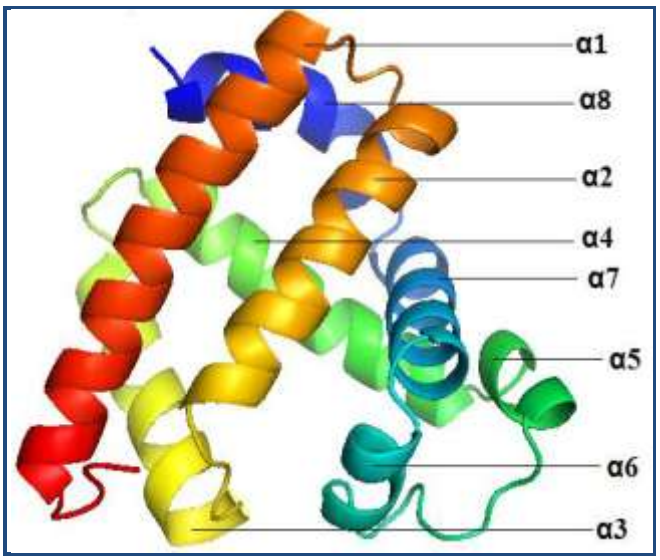

Figure 1: 3-D structure of the predicted model by PyMOL.

Prediction of Protein Structure

The secondary analysis indicates whether a given amino acid is located in helix strand or coil. The result obtained from SOPMA described that about 70.75% of amino acids presented in alpha helix, 6.12% of amino acids in beta turn, and 23.13% of amino acids in random coil.

Prediction of Active Site

A total of 15 active sites were evaluated in the structure through CASTp software with ideal parameters. All 15 pockets were characterized to find out its residues around probe radius of 1.4Å and among them, largest active site has an area of 803.6 Å and volume of 905.3 Å. The green color **(Figure 2)** shows the largest active site position in the build protein which lies between amino acid 1 and 147.

Predicting of Antigenic Determinants

There are total 8 antigenic determinants in the sequence. These antigenic determinants are: PHE4-GLY12, ARG20-THR31, GLU35-ASP50, THR53-GLU70, GLY76-HIS89, HIS93-PHE100, LEU102-GLU109 and ALA124-THR134.

Prediction of Heme Binding Sites

There are total 17 ligand binding sites present in the structure. These sites are THR36, LEU39, PHE40, PRO41, LYS42, HIS60, THR63, VAL64, LYS67, LEU85, SER88, HIS89, HIS93, ILE95, ASN99, PHE100 and ILE103 which can bind to one cluster of 25 ligands (heme). Accessible surface area (ASA) analysis of the predicted model showed the amino acids with low ASA value are buried inside the catalytic cleft and with high ASA values are on the surface of the cleft. Some of the residues were found to have high ASA values (GLU15, PRO41, LYS42, LYS83, LYS92, HIS93) and some others were found to have low ASA values (VAL7, VAL14, GLY22, LEU26, ILE51, VAL57, GLY61, LEU72, LEU82, ALA90, ALA108, LEU111, GLY121, LEU125, VAL128, TYR140). The active site amino acids, which are hydrophobic, are VAL7, VAL14, GLY22, LEU72, LEU82, LEU111, VAL128 with low ASA values. While those active site amino acids, which is hydrophilic are LYS42, LYS92, HIS93 with high values. The protein has 22% amino acids are hydrophilic, 50% are hydrophobic and 27% others. Molecular weight of the protein is 15806.42 g/mol, isoelectric point is pH=7.51. The protein has poor water solubility.

Figure 2: Active site in predicted model computed with CASTp. Green color represents active site with largest area and volume and other colors represent the remaining active site with different areas and volumes.

Conclusion:

In this study, we proposed a valid and stable 3D model of Myoglobin in Channa striata whose structure is not present in PDB (Protein Data Bank). Further analysis provides information about its active sites, ligand binding sites, antigenic determinants and their ASA value analysis in the predicted model. On the basis of the findings, it could be concluded that further characterization of Myoglobin in Channa striata will be important as the nitric oxide (NO) scavenger and myoglobin-mediated oxidative phosphorylation. This study can be used in broad screening on inhibitors of the protein and can be effectively used to raise monoclonal antibodies against its antigenic determinants sites that can be used to support the drug design project and can be further implemented in future studies.

References:

[1] Ordway GA & Garry DJ, *J Exp Biol*. 2004 **207**: 3441 [PMID: 15339940]
[2] Garry DJ *et al. Cell Mol Life Sci*. 2000 **57**: 896 [PMID: 10950305]
[3] Wittenberg JB & Wittenberg BA, *J Exp Biol*. 2003 **206**: 2011 [PMID: 12756283]
[4] Wittenberg BA & Wittenberg JB, *Annu Rev Physiol*. 1989 **51**: 857 [PMID: 2653210]
[5] Moore LG *et al. High Alt Med Biol.* 2002 **3**: 39 [PMID: 12006163]
[6] Noren SR & Williams TM, *Comp Biochem Physiol A mol Integr Physiol.* 2000 **126**: 181 [PMID: 10936758]
[7] Garry DJ *et al. Nature* 1998 **395**: 905 [PMID:9804424]
[8] Doeller JE & Wittenberg BA, *Am J Physiol*. 1991 **261**: H53 [PMID: 1858930]
[9] Poderoso JJ *et al. Am J Physiol*. 1998 **274**: C112 [PMID: 9458719]
[10] Suzuki T & Imai K, *Cell Mol Life Sci*. 1998 **54**: 979 [PMID: 9791540]
[11] Vanderkooi JM *et al. Biochemistry* 1990 **29**: 5332 [PMID: 2383550]
[12] Verkhovsky MI *et al. Nature* 1996 **380**: 268 [PMID: 8637579]
[13] Berman HM *et al. Nucleic Acids Res.* 2000 **28**: 235 [PMID: 10592235]

[14] Nielsen M *et al. Nucleic Acids Res*. 2010 **38**: W576 [PMID: 20542909]
[15] Lorenz IC *et al. Viruses* 2010 **2**: 1635 [PMID: 21994698]
[16] Laskowski RA *et al. J Biomol NMR*. 1996 **8**: 477 [PMID: 9008363]
[17] Bowie JU *et al. Science* 1991 **253**: 164 [PMID: 1853201]
[18] Colovos C & Yeates TO, *Protein Sci*. 1993 **2**: 1511 [PMID: 8401235]
[19] Wiederstein M & Sippl MJ, *Nucleic Acids Res*. 2007 **35**: W407 [PMID: 17517781]
[20] Geourjon C & Deléage G, *Comput Appl Biosci*. 1995 **11**: 681 [PMID: 8808585]
[21] Dundas J *et al. Nucleic Acids Res*. 2006 **34:** W116 [PMID: 16844972]
[22] Kolaskar AS & Tongaonkar PC, *FEBS Lett*. 1990 **276**: 172 [PMID: 1702393]
[23] Wass MN *et al. Nucleic Acids Res*. 2010 **38:** W469 [PMID: 20513649]

Edited by P Kangueane
Citation: **Parveen & Mishra,** Bioinformation 10(1): 019-022 (2014)