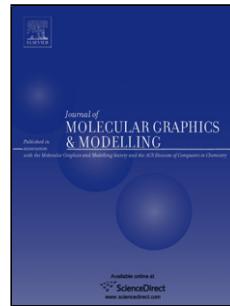


# Accepted Manuscript

Title: Structural insights into the interaction between molluscan hemocyanins and phenolic substrates: an *in silico* study using docking and molecular dynamics.

Author: K.N. Naresh Arun Sreekumar S.S. Rajan



PII: S1093-3263(15)30029-2

DOI: <http://dx.doi.org/doi:10.1016/j.jmgm.2015.07.006>

Reference: JMG 6578

To appear in: *Journal of Molecular Graphics and Modelling*

Received date: 14-4-2015

Revised date: 2-7-2015

Accepted date: 18-7-2015

Please cite this article as: K.N.Naresh, Arun Sreekumar, S.S.Rajan, Structural insights into the interaction between molluscan hemocyanins and phenolic substrates: an *in silico* study using docking and molecular dynamics., Journal of Molecular Graphics and Modelling <http://dx.doi.org/10.1016/j.jmgm.2015.07.006>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Structural insights into the interaction between molluscan hemocyanins and phenolic substrates: an *in silico* study using docking and molecular dynamics.**

K.N.Naresh<sup>1</sup>, Arun Sreekumar<sup>2</sup> and S.S. Rajan\*<sup>1</sup>

1. Department of Biosciences, Sri Sathya Sai Institute of Higher Learning,  
Prasanthi Nilayam - 515134, Anantapur,  
Andhra Pradesh, India.

2. Department of Molecular and Cell Biology,  
Alkek Center for Molecular Discovery,  
Baylor College of Medicine,  
Houston, TX, USA.

\*Corresponding Author:

S.S. Rajan, Ph.D.  
Professor  
Sri Sathya Sai Institute of Higher Learning  
Vidyagiri, Prasanthi Nilayam  
Puttaparthi  
Anantapur District  
Andhra Pradesh  
India  
PIN-515134  
Email: ssrajan@sssihl.edu.in  
Phone: 011-91-950290014  
Fax: 011-91-08555-286919

Highlights

- Molluscan hemocyanin, an oxygen carrier protein exhibits phenoloxidase activity.

- Docking studies reveal that the c-terminal  $\beta$ -domain blocks access to active site.
- Enzyme kinetics show SDS-induced enhancement in the enzyme activity of hemocyanin.
- MD studies show that SDS improves active site access by displacing the  $\beta$ -domain.
- Binding of phenolic substrates to hemocyanin improves after exposure to SDS.

## Abstract

Hemocyanin is a multimeric type-3 copper containing oxygen carrier protein that exhibits phenoloxidase-like activity and is found in selected species of arthropoda and mollusca. The phenoloxidase activity in the molluscan hemocyanins can be triggered by the proteolytic removal of the C-terminal  $\beta$ -rich sandwich domain of the protein or by the treatment with chemical agents like SDS, both of which enable active site access to the phenolic substrates. The mechanism by which SDS treatment enhances active site access to the substrates is however not well understood in molluscan hemocyanins. Here, using a combination of *in silico* molecular dynamics (MD) and docking studies on the crystal structure of *Octopus dofleini* hemocyanin (PDB code:1JS8), we demonstrate that the C-terminal  $\beta$ -domain of the protein plays a crucial role in regulating active site access to bulky phenolic substrates. Furthermore, MD simulation of hemocyanin in SDS revealed

displacement of  $\beta$ -domain, enhanced active site access and a resulting increase in binding affinity for substrates. These observations were further validated by enzyme kinetics experiments.

**Keywords:**

Molluscan hemocyanin,  
SDS-induced activation,  
Molecular docking,  
Molecular dynamics,  
Phenoloxidase activity

## **Structural insights into the interaction between molluscan hemocyanins and phenolic substrates: an *in silico* study using docking and molecular dynamics.**

### Abstract

Hemocyanin is a multimeric type-3 copper containing oxygen carrier protein that exhibits phenoloxidase-like activity and is found in selected species of arthropoda and mollusca. The phenoloxidase activity in the molluscan hemocyanins can be triggered by the proteolytic removal of the C-terminal  $\beta$ -rich sandwich domain of the protein or by the treatment with chemical agents like SDS, both of which enable active site access to the phenolic substrates. The mechanism by which SDS treatment enhances active site access to the substrates is however not well understood in molluscan hemocyanins. Here, using a combination of *in silico* molecular dynamics (MD) and docking studies on the crystal structure of *Octopus dofleini* hemocyanin (PDB code:1JS8), we demonstrate that the C-terminal  $\beta$ -domain of the protein plays a crucial role in regulating active site access to bulky phenolic substrates. Furthermore, MD simulation of hemocyanin in SDS revealed

displacement of  $\beta$ -domain, enhanced active site access and a resulting increase in binding affinity for substrates. These observations were further validated by enzyme kinetics experiments.

## Introduction:

Hemocyanins are multimeric proteins found freely dissolved in the hemolymph of select species of arthropoda and mollusca [1, 2]. Between the phyla, the protein exhibits similar dicopper active site but differs in overall amino acid sequence and structure. However, within each phylum the protein remains conserved [3]. Historically, these proteins have been described as specialized oxygen carrier molecules. Recent *in vitro* studies have detected phenoloxidase activity associated with hemocyanin suggesting its possible role in innate defence mechanisms involving the production of microbicidal pigment, melanin [4]. This activation of enzymatic properties is mediated *in vitro* by various agents that include proteolytic cleavage by trypsin and chymotrypsin and structural disruption by denaturants like sodium dodecyl sulphate (SDS) [4-7].

Structurally, hemocyanins are very closely related to phenoloxidases i.e. tyrosinase (Figure 1(A)) and catecholoxidase (Figures 1(B)) [8-10]. These proteins show similar active site geometry with binuclear copper centers (Figure 1). Despite this structural similarity, the phenoloxidase activity in native hemocyanin is reported to be significantly less compared to tyrosinase and catecholoxidase. The reduced enzyme activity in hemocyanin is associated with the presence of additional domains (Figures 1(C) & 1(D)) that sterically hinder the access of bulky phenolic substrates to the protein's active site [8-10]. These additional domains include a C-terminal  $\beta$ -rich domain seen in molluscans (domain-III in Figures 1(C)) and an N-terminal  $\alpha$ -rich domain seen in arthropods (domain-I in Figure 1(D)) [8, 10]. Together, these domains also provide the frame work for positioning highly conserved 'placeholder' amino acids, L2830 in 1JS8 in molluscan and F49 in 1OXY in arthropodan hemocyanins (Figures 1(C) & 1(D)) which further limit active site access to the substrates [8, 10]. It is suggested that the proteolytic removal of these domains along with their placeholder amino acids result in an increased substrate access to the protein's active site with a concomitant increase in its phenoloxidase activity [8, 10]. In *Octopus* hemocyanin, it is shown that the proteolytic cleavage can occur in

the loop region (Figure 1(C)) connecting  $\alpha$  and  $\beta$  domains (II & III in Figure 1(C)). This results in the removal of 20 kDa  $\beta$ -domain along with the ‘placeholder’ amino acid increasing phenoloxidase activity of the protein for L-DOPA, a monophenolic substrate [11]. Yet another mechanism described for SDS-mediated activation of scorpion hemocyanin using cryo electron microscopy involves the twisting of domain-I from II and III [12].

In spite of all these studies, it is still not clear whether the change in the structure of the protein caused by either proteolytic cleavage or SDS-treatment, is sufficient enough to enable bulkier substrates to access the active site of hemocyanin. Moreover, studies describing the mechanism of SDS-mediated activation of phenoloxidase activity in molluscan hemocyanin are lacking.

To address these lacunae, an *in silico* molecular docking approach is used on molluscan hemocyanins to characterize the active site binding properties of various phenolic substrates ( $n=12$ ) that differ in their functional groups and bulkiness (Figure 2). Further, using *in silico* molecular dynamics simulations the mechanism of SDS-mediated activation of phenoloxidase activity in molluscan hemocyanin via perturbations in its  $\alpha/\beta$  domains is also examined.

## 2. Methodology:

### 2.1. Molecular Docking Studies:

*In silico* docking studies using AutoDock Vina [13] were carried out to understand the interaction between molluscan hemocyanins and various phenolic substrates, and the LigPlot+ [14] was used to identify the various active site residues involved in the interaction. The docking studies were performed using three-dimensional structures of two model molluscan hemocyanins from *Enteroctopus dofleini* and *Megathura crenulata* to substantiate the role of  $\beta$ -rich sandwich domain in regulating the active site access for phenolic substrates. Towards this, the binding energies of various substrates were calculated for molluscan hemocyanins both in the presence and absence of  $\beta$ -rich sandwich domain (Figure – 3(A) & 3(B)) and compared with those calculated for tyrosinase and catecholoxidase.

#### 2.1.1 Preparation of receptors:

The crystal structures of hemocyanins from *Enteroctopus dofleini* (PDB code: 1JS8) and *Megathura crenulata* (KLH, PDB code: 3QJO), as well as tyrosinase from *Agaricus bisporus* (PDB code: 2Y9W) and catecholoxidase from *Ipomoea batatas* (PDB code: 1BT1), were used as receptors. For the docking studies only the chain-A (Figures – 1 (A) to 1(D)) and the active site hetero-atoms containing Cu and peroxide were retained after manually deleting the coordinates for B/C chains. The water molecules and other bound ligands were also removed from the data. The PDB files containing the modified structures were then imported into MGL tools 1.5.4., where the hydrogen atoms as well as the Kollman and Gasteiger charges were added. The resulting PDB files were used in subsequent calculations.

### **2.1.2 Preparation of ligands:**

The structure of each substrate was drawn in two-dimensions and converted into three-dimensional structures using Hyperchem-7 [15, 16]. The geometry of these structures was optimized using Polak–Ribiere (conjugate gradient) algorithm, in Hyperchem-7, following the force field molecular mechanics (MM+) method [17]. The resulting structures were then used as the starting point for energy re-minimization using Polak–Ribiere optimization technique following AM1 semi-empirical quantum mechanical method [18]. The energy minimizations were performed until the absolute value of its largest partial derivative with respect to the coordinates was below  $0.04 \text{ kJ mol}^{-1} \text{ Å}^{-1}$ . Following all of the above analysis, the resulting structure files were saved in a format (.mol2) compatible with Autodock Vina [13].

### **2.1.3 Docking:**

All the docking studies were carried out on PyMol molecular visualisation platform using AutoDock/Vina PyMol plugin [19]. The files containing both the ligand and the protein structures were imported into PyMol and converted to .PDBQT files using the plugin. The grid search space was then set to a size of 15 Å in x, y and z directions with reference to the central dicopper active site. The grid box was centered with the x,y,z co-ordinates for each receptor: 3QJO (-25.65, 48.30, 44.98); 1JS8 (56.65,

43.50, 68.69); 2Y9W (5.38, 28.50, 93.20); 1BT1 (19.00, 99.50, 7.50). The docking was performed for each ligand using default parameters in AutoDock Vina [13]. For each docked ligand, binding energy values were obtained at ten different binding confirmations. The binding energy corresponding to the most optimal confirmation was determined by the software and used for downstream analysis. The receptor-ligand complexes were analysed using LigPlot+ software [14] to delineate the various interactions between the receptor and the ligand. The interaction diagrams displaying hydrogen bonding and hydrophobic interactions were generated and the amino acid residues taking part in the interaction were identified. The binding energies of various substrates bound to hemocyanins, both in the presence and absence of the  $\beta$ -domain, were computed and compared with those of catecholoxidase and tyrosinase.

## **2.2. Solvent accessibility calculations:**

Once the amino acid residues in hemocyanins that interact with the substrates were identified, their solvent accessibility was calculated and compared for structures both in the presence and absence of the  $\beta$ -domain. The relative solvent accessibility (RSA), a measure of substrate accessibility, was expressed as percentage of the accessible surface observed in ALA-X-ALA tripeptide, where X represents any given amino acid [20]. Naccess version 2.1.1, a stand-alone FORTRAN based program was used for the calculation of relative solvent accessibilities. It calculates the accessible surface area when a probe of a given radius is rolled around the van der Waals surface of the protein molecule. The path traced out by the centre of a given probe with a van der Waals radius similar to that of water (1.4 Angstroms) is labelled as the solvent accessible surface [20].

## **2.3. MD simulations:**

MD simulations were carried out with the structure of *Octopus* hemocyanin (PDB id: 1JS8) using GROMACS 4.6.3 run on a Linux cluster. This structure was selected for its better resolution compared to the other structures of molluscan hemocyanins available from PDB. As the objective was to study the effect of SDS on overall protein structure, Cu<sub>2</sub>O<sub>2</sub> group in the active site was removed from 1JS8. G43a1

GROMOS molecular mechanics force field was chosen and a cubic box of dimensions 10 nm x 10 nm x 10 nm was defined around the protein for all simulations. The simulations were carried out in three different systems where the protein is surrounded by : **a.** 41424 water molecules and 13 Na<sup>+</sup> ions, **b.** 40876 water molecules, 32 dodecyl sulphate ions (~ 0.075 %) and 45 Na<sup>+</sup> ions and **c.** 40353 water molecules, 64 dodecyl sulphate ions (~ 0.15%) and 77 Na<sup>+</sup> ions. In all cases Na<sup>+</sup> ions were added to neutralize the charge in the systems. Further, the systems were subjected to energy minimization for 500 steps by using steepest descent algorithm and the equilibration was conducted for the systems in two phases – i. under an NVT (constant Number of particles, Volume and Temperature) and ii. an NPT (constant Number of particles, Pressure and Temperature) ensemble. The equilibration simulation was run under position restraint condition wherein the atoms of the protein were held fixed and those of the molecules surrounding the protein were allowed to move freely such that they reach equilibrium states. For systems ‘**a**’ and ‘**b**’ both NVT and NPT equilibration runs were carried out for 500 pico seconds (ps) each whereas for system ‘**c**’, it was for 1000 ps with a time step of 1 femto second (fs). The periodic boundary conditions set to x,y,z coordinates were used to keep all the atoms within the simulation box. A constant temperature of 300 K was maintained during the simulations with the help of modified Berendsen temperature coupling algorithm [21]. Leap-frog integrator [22] was used to integrate Newton’s equations in MD simulations. LINCS algorithm [23] was used to fix the covalent bonds between atoms of protein molecule while SETTLE algorithm [24] was used to fix the same in water molecules. Finally MD simulations were carried out for 7500 ps with each system without any restraint on the movement of the atoms.

#### **2.4. Trajectory analysis:**

GROMACS [25, 26] was used to analyse the trajectory files obtained after MD simulations. To start with the RMSD of C<sup>α</sup> chain of the protein was calculated at a time interval of 1 ps. The overlay of the RMSD plots comparing the α-carbon (C<sup>α</sup>) chain of the protein in systems **a**, **b** and **c** were generated. MD trajectory was analyzed using UCSF Chimera [27] and PyMol [28].

## 2.5. Enzyme kinetics experiments:

Experiments to calculate enzyme kinetic parameters were carried out using a model molluscan hemocyanin i.e. the keyhole limpet hemocyanin (KLH) [Sigma chemicals] to show the effect of SDS on the binding properties of phenolic substrates to the protein.  $K_m$ , the Michealis Menton constant that reflects the enzyme-substrate affinity, was calculated for two model substrates, catechol and dopamine, under two different conditions, one in the presence and the other in the absence of SDS. A microplate based protocol using an assay volume of 250  $\mu$ l was designed based on the method reported earlier for tyrosinase [29]. The kinetic readouts were obtained using a Spectramax M-5 multimode microplate reader (Molecular Devices, USA). A typical reaction mixture consisted of components at the given final concentrations: Enzyme (KLH) – 0.6 mg/ml; DMF – 4%; MBTH – 1mM; and substrates (ranging from 1.25 mM to 15 mM) in 50 mM sodium phosphate buffer, pH 6.8. The reaction was started by adding the required volume of enzyme stock solution. The protein was incubated in 0.1% SDS for 5 minutes prior to its addition for studying the kinetics under SDS-activated condition. The kinetic parameters –  $K_m$  and  $V_{max}$  were calculated through non-linear fitting of the kinetic data using GraphPad Prism [30].

## 3. Results and discussion:

### 3.1. Molecular Docking to understand active site binding properties of substrates

*In silico* molecular docking studies were carried out using 12 phenolic substrates (Figure 2). Both monophenols and diphenols were chosen as substrate ligands to study their interaction with the molluscan hemocyanins. Figure 3(A) shows the structure of hemocyanin with its active site capped by a lid-like  $\beta$ -rich domain where the  $\beta$ -rich sandwich domain, is held in place by the placeholder amino acid residue – Leucine (L2830). The removal of the  $\beta$ -domain allowed docking of phenolic substrates like dopamine (Figure 3(B)). Next the affinity of the various ligands to the active sites of hemocyanins (with and without the  $\beta$ -domain) from two different molluscans *E. dofleini* (1JS8) and *M. crenulata* (3QJO) as well as tyrosinase from *A. bisporus* (2Y9W) and catecholoxidase from *I. batatus* (1BT1) was examined.

This was done by comparing their relative binding energies. The binding energy values become highly negative for 1JS8\* and 3QJO\* (receptors lacking β-domain) compared to native 1JS8 and 3QJO respectively (Figure 4). This indicates the enhancement of ligand binding affinity to the active site of hemocyanins after the removal of the C-terminal β-domain and confirms the steric hindrance to phenolic substrates associated with the presence of β-domain in the protein. The removal of the β-domain allows the binding affinity of phenolic substrates to hemocyanin to be comparable with tyrosinase and catecholoxidase (Figure 4). Following *in silico* molecular docking, the amino acid residues which are involved in receptor-ligand interaction were identified using LigPlot+ software. Table 1 lists the active site amino acids involved in ligand-receptor interaction, their binding energies for various ligands and the number of potential hydrophobic/hydrogen bonds that constitute the interaction. Four (two Hs, N & S) out of five active site residues were found to be conserved between 1JS8 and 3QJO, while Threonine and Leucine constituted the fifth amino acid in *E. dofleini* and *M. crenulata* hemocyanin respectively.

### **3.2. Solvent accessibility:**

The solvent accessibility was examined for the five conserved active site amino acid residues as a surrogate to define the ease with which the ligands can access the protein's active site. As shown in Figures 5(A) & 5(B), the removal of the C-terminal β-domain in molluscan hemocyanins resulted in an increased solvent and ligand access to the five key amino acid residues in the protein's active site, in turn, substantiating the enhancement in the binding affinity values observed in the docking studies.

### **3.3. Molecular dynamics simulation studies reveal SDS-induced structural changes in hemocyanin:**

Having identified an improvement in ligand binding properties upon the removal of the C-terminal β-rich sandwich domain of molluscan hemocyanin, the next step was to find out if this could explain SDS-induced activation of phenoloxidase activity in the protein.

To address this, MD simulations lasting for 7500 ps were carried out with *Octopus* hemocyanin using either water alone or a solution of SDS in water and the change in

the structure of the protein during the course of simulation was monitored. Figure – 6 shows the alteration in the root mean square deviation (RMSD) values for the C<sup>α</sup> chain of the protein in relation to its original starting structure. It is evident from the plot that the protein structure simulated in the environment with water containing 0.15% SDS ('c' in Figure - 6) shows a prominent variation in the RMSD value compared to the ones simulated in environments with water alone ('a' in Figure - 6) and water with 0.075 % SDS ('b' in Figure - 6). It is evident that the RMSD value increases in case of 'c' starting from the time point of ~1000 picoseconds until ~3500 picoseconds and stabilizes thereafter for a period of 4000 ps. Further, the MD study was intended to observe mean local fluctuations in the native structure of the protein under the effect of SDS rather than major unfolding of the protein that may result at higher time scales. Based on this, the simulations were terminated at 7500 ps. The trajectory file for 'c' was then used to extract structural snapshots and gain insights into the alterations in the protein caused by SDS during the course of the MD simulation. For this, the structural snapshot of the protein obtained at the beginning of the simulation (0 ps) was compared with the one resulted at the end (7500 ps) by carrying out structural alignment using PyMol. Specifically the α-rich core domain (A2503 to Y2769) was aligned in order to highlight the displacement of the β-rich sandwich domain which plays important role in regulation of the active site access. It is clearly evident that the β-domain in the structure corresponding to the time point - 7500 ps has got displaced from its original position observed in the structure at the time point - 0 ps (Figure – 7(A)). It is further highlighted in the inset of figure – 7(B) that the side chain of the placeholder amino acid residue – L2830 is displaced by a distance of 5.4 Å from its native location. As a consequence, it probably breaks the distance dependent van der Waals interaction, proposed between the 'placeholder' and the active site histidine (His2562) that may be crucial for holding the β-rich sandwich domain in place [8, 31].

Further, to investigate the effect of the displacement of the β-domain on the interaction of the substrates with the active site, the molecular docking studies using various phenolic substrates were carried out with the structures extracted at different time points of the MD simulation i.e., 0, 1000, 3000, 5000 and 7500 ps. It was found that the binding affinities of the substrates progressively improved from 0 to 7500 ps (Figure – 8). This shows that the active site of the protein becomes more accessible

to the substrates with the displacement of the  $\beta$ -domain under the influence of SDS. Further analysis using VEGA [32] suggest that high positive binding energies exhibited by ligands 3, 9 and 12 could be explained by their large molecular volumes (refer Supplementary table S1). Our observations from the MD simulations with the structure of *Octopus* hemocyanin in water/SDS environment support the earlier hypothesis [8] relating to the tilting of  $\beta$ -rich sandwich domain through possible disruption of inter-domain interactions in molluscan hemocyanins by detergents to open an access route to the active site.

### **3.4. Experimental results on o-diphenoloxidase like activity of KLH:**

The use of SDS to enhance the phenoloxidase activity in hemocyanins has been reported in numerous studies [4]. In general, this process involves pre-incubation of the protein with 0.1% SDS for 5 minutes prior to assaying for enzyme activity. Earlier reports based on the analysis of crystal structures of molluscan hemocyanins have described the access of ligands to the protein's active site as a key determinant for its enzyme activity [8-10]. The MD studies described here confirm that SDS-induced enhancement of phenoloxidase activity in molluscan hemocyanin is accomplished by the displacement of the protein's  $\beta$ -domain allowing better ligand access and concomitantly higher binding affinity. To validate this, the phenoloxidase activity of KLH was measured using normal Michealis-Menten kinetics both in the presence and absence of SDS (Figures – 9(A) & 9(B)). The  $K_m$  values calculated using a non-linear fitting of the kinetic data show an enhancement in the activity of the protein upon the addition of SDS (Figures – 9(A) & 9(B), & Table 2) corroborating our *in silico* predictions. Notably, upon addition of SDS, the  $K_m$  value of hemocyanin for catechol and dopamine has improved by 1.5 and 1.6 folds respectively.  $K_{cat}$ , that defines the enzyme turnover, did not change for catechol while it increased by 2.5 fold for dopamine (Table 2). Overall, our data show that the catalytic efficiency of the enzyme, described by  $K_{cat}/K_m$  values, shows 1.4 fold improvement (from 63.8 to 88.6) for catechol and a 4.0 fold increase (from 7.9 to 32.0) for dopamine, upon addition of SDS. In other words, the addition of SDS has minimal to no effect on binding properties of catechol while it prominently enhances the binding characteristics of dopamine. This is consistent with the less complex structure of catechol, allowing it to access the active site probably even in presence of the  $\beta$ -

domain while the more complex structure of dopamine makes it sterically unable to access the catalytic site. Although the data obtained from enzyme kinetic studies are consistent with the results seen with MD, the caveat remains that the former was carried out using native multimeric protein while the latter was done using structural information derived from a single functional unit of the protein.

#### 4. Conclusions:

Using *in silico* molecular docking the C-terminal  $\beta$ -domain in the functional unit of molluscan hemocyanins was identified to cause steric hindrance to bulky phenolic substrates accessing the protein's active site. MD simulations coupled to *in vitro* kinetic assays reveal that SDS enhances phenoloxidase activity in molluscan hemocyanins by sideways displacement of the  $\beta$ -domain thus releasing the steric hindrance and allowing easy access for phenolic substrates to the protein active site. Though the displacement of the domain along with the placeholder amino acid (Leu 2830) *per se* is consistent with the hypothetical mechanism of activation of molluscan hemocyanin, our observations show that the domain tilts sideways rather than upwards as proposed by Decker and Tuczak [8]. A cryo electron microscopic study has described a similar mechanism involving displacement of the N-terminal  $\alpha$ -rich domain, for SDS-induced activation of hemocyanin in *Pandinus imperator* (scorpion, arthropod) [12].

Catechol and other phenolic derivatives that are toxic and hazardous are common components of industrial effluents. Prior studies have reported different versions of tyrosinase-based biosensors for detection of phenols in environmental samples [33-36]. Two recent reports, one from our group [7] and the other by Raynova et al [5] described hemocyanin-based optical biosensors built by immobilizing molluscan hemocyanin on a thin film of chitosan coated on a glass slide. This first-generation sensor was able to detect catechol at concentrations of 5 ppm to 90 ppm. Importantly, unlike tyrosinase, hemocyanin can be purified with greater ease, obtained in larger quantities and is more thermostable. However, it suffers from lower affinity for bulky phenolic substrates. Currently, optimal phenoloxidase activity in hemocyanins is obtained by either carrying out limited proteolysis [5] or pre-treatment with SDS [7], prior to its immobilization on sensors. Thus, it is important to understand the structural changes caused by the action of proteolytic agents or SDS

resulting in an increase in the affinity of the protein for bulky phenolic substrates. Here, using an *in silico* MD guided approach coupled to molecular docking studies, we were able to determine the changes in the structure of hemocyanin caused by these agents that were associated with concomitant increase in its substrate affinity. Importantly, the *in silico* method described here is universal and can be used to determine the effect of various chemical agents on the binding properties of any enzyme-substrate pair. Using this information, it is possible to inform the design of biosensors with enhanced substrate sensitivity.

## **5. Acknowledgements:**

The authors dedicate this paper to their Beloved Founder Chancellor Bhagawan Sri Sathya Sai Baba who offers them free and value-based education at the University. Financial support from University Grants Commission, Department of Science & Technology and Department of Biotechnology, Govt. of India, New Delhi, India under SAP-DRS-Level-II, FIST-Level-I and BIF-Level-I schemes respectively are gratefully acknowledged. The authors also thank Prof. V. Chandrashekaran, Head, Department of Mathematics and Computer Science, SSSIHL, Prasanthi Nilayam for extending the computational facility to carry out MD simulations.

## **6. References:**

- [1] van Holde, K.E., Miller, K.I. Haemocyanins. *Advances in Protein Chemistry*. 1995, 47, 1-81.
- [2] van Holde, K.E., Miller, K.I., Decker, H. Hemocyanins and Invertebrate Evolution. *Journal of Biological Chemistry*. 2001, 276, 15563-6.
- [3] Markl, J. Evolution of molluscan hemocyanin structures. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*. 2013.
- [4] Coates, C.J., Nairn, J. Diverse immune functions of hemocyanins. *Developmental & Comparative Immunology*. 2014, 45, 43-55.
- [5] Raynova, Y., Doumanova, L., Idakieva, K.N. Phenoloxidase activity of *Helix aspersa maxima* (garden snail, gastropod) hemocyanin. *The protein journal*. 2013, 32, 609-18.
- [6] Dolashki, A., Voelter, W., Dolashka, P. Phenoloxidase activity of intact and chemically modified functional unit RvH1: a from molluscan *Rapana venosa* hemocyanin. *Comparative biochemistry and physiology. Part B, Biochemistry & molecular biology*. 2011, 160, 1-7.
- [7] Naresh, K.N., Krupanidhi, S., Rajan, S.S. Purification, spectroscopic characterization and o-diphenoloxidase activity of hemocyanin from a freshwater gastropod: *Pila globosa*. *The protein journal*. 2013, 32, 327-36.
- [8] Decker, H., Tuczek, F. Tyrosinase/catecholoxidase activity of hemocyanins: structural basis and molecular mechanism. *Trends in Biochemical Sciences*. 2000, 25, 392-7.

- [9] Decker, H., Jaenicke, E. Recent findings on phenoloxidase activity and antimicrobial activity of hemocyanins. *Developmental & Comparative Immunology*. 2004, 28, 673-87.
- [10] Decker, H., Schweikardt, T., Nillius, D., Salzbrunn, U., Jaenicke, E., Tuczek, F. Similar enzyme activation and catalysis in hemocyanins and tyrosinases. *Gene*. 2007, 398, 183-91.
- [11] Campello, S., Beltramini, M., Giordano, G., Di Muro, P., Marino, S.M., Bubacco, L. Role of the tertiary structure in the diphenol oxidase activity of *Octopus vulgaris* hemocyanin. *Archives of biochemistry and biophysics*. 2008, 471, 159-67.
- [12] Cong, Y., Zhang, Q., Woolford, D., Schweikardt, T., Khant, H., Dougherty, M., et al. Structural mechanism of SDS-induced enzyme activity of scorpion hemocyanin revealed by electron cryomicroscopy. *Structure (London, England : 1993)*. 2009, 17, 749-58.
- [13] Trott, O., Olson, A.J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*. 2010, 31, 455-61.
- [14] Laskowski, R.A., Swindells, M.B. LigPlot+: Multiple Ligand–Protein Interaction Diagrams for Drug Discovery. *Journal of Chemical Information and Modeling*. 2011, 51, 2778-86.
- [15] Froimowitz, M. HyperChem: a software package for computational chemistry and molecular modeling. *BioTechniques*. 1993, 14, 1010-3.
- [16] HyperChem7. Hypercube, Inc., 1115 NW 4th Street, Gainesville, Florida 32601, USA.
- [17] Khoda, K.M., Liu, Y., Storey, C. Generalized Polak-Ribière algorithm. *J Optim Theory Appl*. 1992, 75, 345-54.
- [18] Dewar, M.J.S., Zoebisch, E.G., Healy, E.F., Stewart, J.J.P. Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. *Journal of the American Chemical Society*. 1985, 107, 3902-9.
- [19] Seeliger, D., de Groot, B.L. Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *Journal of computer-aided molecular design*. 2010, 24, 417-22.
- [20] Hubbard, S.J., Thornton, J.M. 'NACCESS', Computer Program, Department of Biochemistry and Molecular Biology, University College London. 1993.
- [21] Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A., Haak, J.R. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*. 1984, 81, 3684-90.
- [22] Hockney, R.W., Goel, S.P., Eastwood, J.W. Quiet high-resolution computer models of a plasma. *Journal of Computational Physics*. 1974, 14, 148-58.
- [23] Hess, B., Bekker, H., Berendsen, H.J.C., Fraaije, J.G.E.M. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*. 1997, 18, 1463-72.
- [24] Miyamoto, S., Kollman, P.A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of Computational Chemistry*. 1992, 13, 952-62.
- [25] Berendsen, H.J.C., van der Spoel, D., van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications*. 1995, 91, 43-56.
- [26] Lindahl, E., Hess, B., van der Spoel, D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J Mol Model*. 2001, 7, 306-17.
- [27] Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*. 2004, 25, 1605-12.
- [28] De Lano, W.L. The PyMOL molecular graphics system. 2002.
- [29] Espín, J.C., Varón, R., Fenoll, L.G., Gilabert, M.A., García-Ruiz, P.A., Tudela, J., et al. Kinetic characterization of the substrate specificity and mechanism of mushroom tyrosinase. *European Journal of Biochemistry*. 2000, 267, 1270-9.
- [30] Non-linear fitting of enzyme kinetics data using Michealis Menten equation was performed using GraphPad Prism version 6.00 (30 days trial) for Windows, G.S., La Jolla California USA, [www.graphpad.com](http://www.graphpad.com).
- [31] Cuff, M.E., Miller, K.I., van Holde, K.E., Hendrickson, W.A. Crystal structure of a functional unit from *Octopus* hemocyanin. *Journal of molecular biology*. 1998, 278, 855-70.

- [32] Pedretti, A., Villa, L., Vistoli, G. VEGA--an open platform to develop chemo-bio-informatics applications, using plug-in architecture and script programming. *Journal of computer-aided molecular design.* 2004, 18, 167-73.
- [33] Wang, G., Xu, J.-J., Ye, L.-H., Zhu, J.-J., Chen, H.-Y. Highly sensitive sensors based on the immobilization of tyrosinase in chitosan. *Bioelectrochemistry.* 2002, 57, 33-8.
- [34] Abdullah, J., Ahmad, M., Karuppiah, N., Heng, L.Y., Sidek, H. Immobilization of tyrosinase in chitosan film for an optical detection of phenol. *Sensors and Actuators B: Chemical.* 2006, 114, 604-9.
- [35] Wang, S., Tan, Y., Zhao, D., Liu, G. Amperometric tyrosinase biosensor based on Fe<sub>3</sub>O<sub>4</sub> nanoparticles–chitosan nanocomposite. *Biosensors and Bioelectronics.* 2008, 23, 1781-7.
- [36] Yang, L., Xiong, H., Zhang, X., Wang, S. A novel tyrosinase biosensor based on chitosan-carbon-coated nickel nanocomposite film. *Bioelectrochemistry.* 2012, 84, 44-8.
- [37] Ismaya, W.T., Rozeboom, H.J., Weijn, A., Mes, J.J., Fusetti, F., Wicher, H.J., et al. Crystal structure of *Agaricus bisporus* mushroom tyrosinase: identity of the tetramer subunits and interaction with tropolone. *Biochemistry.* 2011, 50, 5477-86.
- [38] Klabunde, T., Eicken, C., Sacchettini, J.C., Krebs, B. Crystal structure of a plant catechol oxidase containing a dicopper center. *Nature structural biology.* 1998, 5, 1084-90.
- [39] Magnus, K.A., Hazes, B., Ton-That, H., Bonaventura, C., Bonaventura, J., Hol, W.G. Crystallographic analysis of oxygenated and deoxygenated states of arthropod hemocyanin shows unusual differences. *Proteins.* 1994, 19, 302-9.

Legends for figures:

### **Fig. 1.**

- a.** Crystal structure of the active form of tyrosinase from *Agaricus bisporus* (PDB code: 2Y9W) [37] and **b.** catecholoxidase from *Ipomoea batatas* (PDB code: 1BT1) [38].
- c.** Crystal structure of hemocyanin from a mollusc – *Octopus dofleini* (PDB code: 1JS8) [31]. Cup-like α-rich core domain (blue) with the active site (two copper atoms with two oxygen atoms represented as spheres, co-ordinated by six histidine residues represented in yellow) is covered by lid-like C-terminal β-rich sandwich domain (red). Flexible loop that covalently connects the two domains is indicated in magenta. The placeholder amino acid residue L2830 is indicated in green. The sandwich domain along with the placeholder residue is considered to play a critical role in regulating the active site access for bulky phenolic substrates.

**d.** Crystal structure of hemocyanin from an arthropod - *Limulus polyphemus* (PDB code: 1OXY) [39]. Central core domain (Domain – II, blue colored) with the dicopper active site (two copper atoms with two oxygen atoms represented as spheres, coordinated by six histidine residues represented in yellow) is flanked by an N-terminal α-rich domain (Domain-I, cyan colored) and C-terminal β-rich domain (Domain-III, red colored). Placeholder amino acid residue – F49 – is indicated in green. All images were generated using PyMol molecular graphics system [28].

**Fig. 2.** Structures of the phenolic substrates used for docking with hemocyanins: (1 to 5 - Monophenols and 6 to 12 - Diphenols) **1.** p-Cresol **2.** Tyramine **3.** L-Tyrosine **4.** p-hydroxy phenyl acetic acid **5.** p-hydroxy phenyl propionic acid **6.** Catechol **7.** 4-Methylcatechol **8.** Dopamine **9.** L-3,4-dihydroxyphenylalanine **10.** 3,4-Dihydroxyphenylacetic acid **11.** 3-(3,4-Dihydroxyphenyl)propionic acid **12.** 4-ter-butyl catechol.

**Fig. 3. a.** Structure of *Octopus dofleini* hemocyanin (1JS8): the N-terminal α-rich core domain with the active site is shown as surface representation and the C-terminal β-rich sandwich domain as cartoon representation in red color. The sandwich domain held in place by close interaction of ‘placeholder’ amino acid residue – Leu2830 (green) with the active site cup is highlighted in the zoomed cross-eyed stereo image.

**b.** *Octopus* hemocyanin devoid of the β-rich sandwich domain (1JS8\*) docked with the substrate dopamine. Dopamine docked well into the active site cup in the absence of the ‘placeholder’ amino acid residue is highlighted in the zoomed cross-eyed stereo image.

**c & d** LigPlot diagrams of the docked complex of 1JS8\* with the substrate catechol & dopamine (in yellow) respectively showing the amino acid residues interacting with the substrates through hydrogen bonding (---) and hydrophobic interactions (\*\*\*\*).

**Fig. 4.** Comparison of the binding energies of the phenolic substrates (1 to 12) obtained after docking with the structure of *Octopus* hemocyanin with β-rich domain (1JS8), keyhole limpet hemocyanin with β-rich domain (3QJO), 1JS8 and 3QJO

without  $\beta$ -rich domain (1JS8\* and 3QJO\* ) and phenoloxidases i.e. mushroom tyrosinase (2Y9W) and plant catecholoxidase (1BT1).

**Fig. 5. a & b.** Comparison of the RSA values of the substrate interacting active site amino acid residues in 1JS8 & 1JS8\* and 3QJO & 3QJO\* respectively.

**Fig. 6.** Comparison of RMSD plots of C<sup>a</sup> of the structure of *Octopus* hemocyanin (1JS8) obtained after subjecting it to MD simulations in three different systems a, b and c. System ‘a’ –hemocyanin in water alone; system ‘b’ –hemocyanin in water with 0.075% SDS and system ‘c’ – hemocyanin in water with 0.15% SDS. The MD simulation was terminated at 7500 ps as the RMSD remains stable after 3500 ps for a time period of 3000 to 4000 ps.

### **Fig. 7.**

**a.** Cross-eyed Stereo image of the alignment of the structural snapshot of *Octopus* hemocyanin obtained at the end of MD simulation (7500 ps) with the structure at the beginning (0 ps). It can be noted that the  $\beta$ -domain in structure at 7500 ps (yellow) has slid out from its original location in structure at 0 ps (red).

**b.** Side view of the alignment showing the displacement of the  $\beta$ -domain. Inset: zoomed image showing the displacement of the ‘placeholder’ amino acid residue Lue2830 from its native location at 0 ps (red) by a distance of 5.4 angstroms in the structure at 7500 ps (yellow).

**Fig. 8.** Comparison of the binding energies of the phenolic substrates (1 to 12) obtained after docking with the structure snapshots of *Octopus* hemocyanin extracted at time points 0, 1000, 3000, 5000 and 7500 ps of the MD simulation.

**Fig. 9. a & b.** A Comparison of Michealis-Menten kinetics plots for conversion of catechol & dopamine to their respective quinones by keyhole limpet hemocyanin in assay conditions described in section 2.5 both in the presence and absence of SDS.

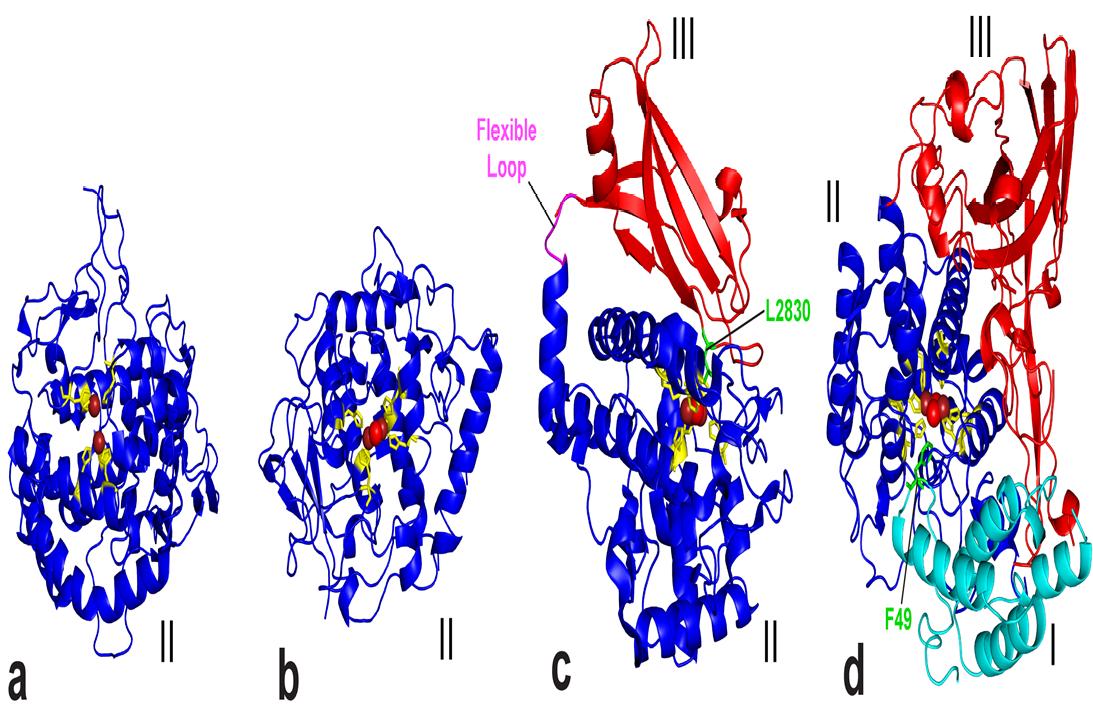
id	Binding Energy	No. of Hydrophobic Interactions	Amino acid residues involved in Hydrophobic interactions	No. of H-Bonding interactions	Residues involved
<b>Table - 1:</b> Tabulation of the binding energy values for the phenolic substrates along which form hydrophobic interactions and hydrogen bonding with the substrates hemocyanin with the β-domain removed; 3QJO* - <i>Megathura crenulata</i> hemocyanin 2Y9W – Tyrosinase from <i>Agaricus bisporus</i> ; 1BT1 – Catecholoxidase from <i>Ipomoea</i>					

ACCEPTED MANUSCRIPT					
E	-5.0	8	N2672, H2675, S2687, T2688, T2692, H2543, M2686, L2689	2	H2562, CuO888
MN	-5.3	7	H2562, N2672, S2687, T2688, CuO888, M2686, L2689	3	H2675, T2692, E2668
R	-6.0	6	N2672, S2687, T2688, T2692, H2543, L2689,	3	H2562, H2675, CuO888
A	-5.8	6	H2562, S2687, T2688, M2686, L2689, CuO888	3	N2672, H2675, T2692
A	-5.2	6	H2975, N3086, H3089, S3101, L3103, T3102	2	E3082, S3106
	-5.1	9	H2562, N2672, H2675, S2687, T2688, H2543, H2671, M2686, L2689	2	CuO888, T2692
	-5.5	5	H2562, N2672, S2687, T2688, H2543,	4	H2675, T2692, M2686
N	-5.6	10	H2562, N2672, H2675, T2688, S2676, H2543, H2671, M2686, S2687	2	CuO888, T2692,
PA	-6.1	7	H2562, N2672, S2687, T2688, H2543, E2668, L2689	4	H2675, M2686, T2692
PAC	-6.1	5	H2562, S2687, T2688, H2543, L2689	5	N2672, H2675, M2686
PPA	-6.2	5	H2562, S2687, T2688, H2543, L2689	5	N2672, H2675, M2686
	-6.0	6	H2562, N2672, S2687, H2543, M2686, L2689	4	H2675, T2688, T2692
n ing cid s			<b>H2562, N2672, H2675, S2687, T2688</b>		
E	-4.5	5	H2975, N3086, S3101, L3103, S3106	1	H3089
MN	-4.5	4	H2975, N3086, S3101, L3103	2	H3089, S3106
R	-5.5	5	H2975, N3086, H3089, L3103, T3102	3	S3101, E3082, S3106
A	-5.4	3	H2975, S3101, L3103	4	N3086, H3089, E3082
A	-5.2	6	H2975, N3086, H3089, S3101, L3103, T3102	2	E3082, S3106
	-4.4	4	H2975, N3086, H3089, T3102	2	S3101, L3103
	-4.5	4	S3101, L3103, H3089, H3085	2	H2975, N3086
N	-4.6	4	H2975, S3101, L3103, S3106	3	N3086, H3089, E3082
PA	-5.4	5	H2975, H3089, L3103, M3100, T3102	3	N3086, S3101, S3106
PAC	-5.6	5	H2975, H3089, S3101, L3103, T3102,	3	N3086, E3082, S3106
PPA	-5.8	6	H2975, H3089, S3101, L3103, T3102, S3106	3	N3086, E3082, M3100
	-5.1	6	H2975, N3086, S3101, L3103, T3102, Cu3408	3	H3089, M3100, S3106
n ing cid s			<b>H2975, N3086, H3089, S3101, L3103</b>		
E		9	N260, H263, F264, H259, M280, G281, V283, A286, Cu401	2	H61, H85
MN		11	N260, H263, F264, H259, M280, G281, S282, V283, A286, Cu400, Cu401	2	H61, H85
R		7	N260, H263, F264, H259, V283, Cu400, Cu401	2	H61, H85
A		10	N260, H263, F264, H259, M280, G281, S282, A286, Cu400, Cu401	2	H61, H85
A		9	N260, H263, F264, H85, M280, G281, S282, V283, A286	1	H244
		8	N260, H263, F264, H61, G281, S282, V283, A286	1	M280
			N260, H263, F264, H61, H259, M280, G281, S282, V283, A286,		
N		9	N260, H263, F264, H61, H259, G281, S282, V283, A286	1	M280
PA		10	H263, F264, H61, H244, H259, M280, G281, S282, V283, A286	1	N260
PAC		9	N260, H263, F264, H61, H259, G281, S282, V283, A283	1	M280
PPA		6	N260, F264, G281, S282, V283, A286	3	H244, H263, M280
		10	N260, H263, F264, H259, G281, V283, A286, F292, Cu400, Cu401	2	H61, H85
n ing cid s			<b>N260, H263, F264</b>		
E		3	N260, F261, I241,	1	G259
MN		6	N260, F261, I241, H244, Arg245, G259		
R		3	N260, F261, Q107	4	Cys92, N110, S229, Q107
A		2	N260, F261	3	Q107, N110, G259
A		2	N260, F261	3	Q107, N110, G259
		3	N260, I241, G259	1	F261
		5	N260, I241, H244, Arg245, G259,	1	F261
N		3	N260, I241, G259,	2	F261, N110
PA		4	N260, F261, H109, I241	3	Q107, N110, G259
PAC		1	N260,	4	F261, Cys92, Q107, N110
PPA		2	N260, I241	3	F261, N110, G259
		3	N260, F261, N110	2	Q107, S229
n ing cid s			<b>N260, F261</b>		

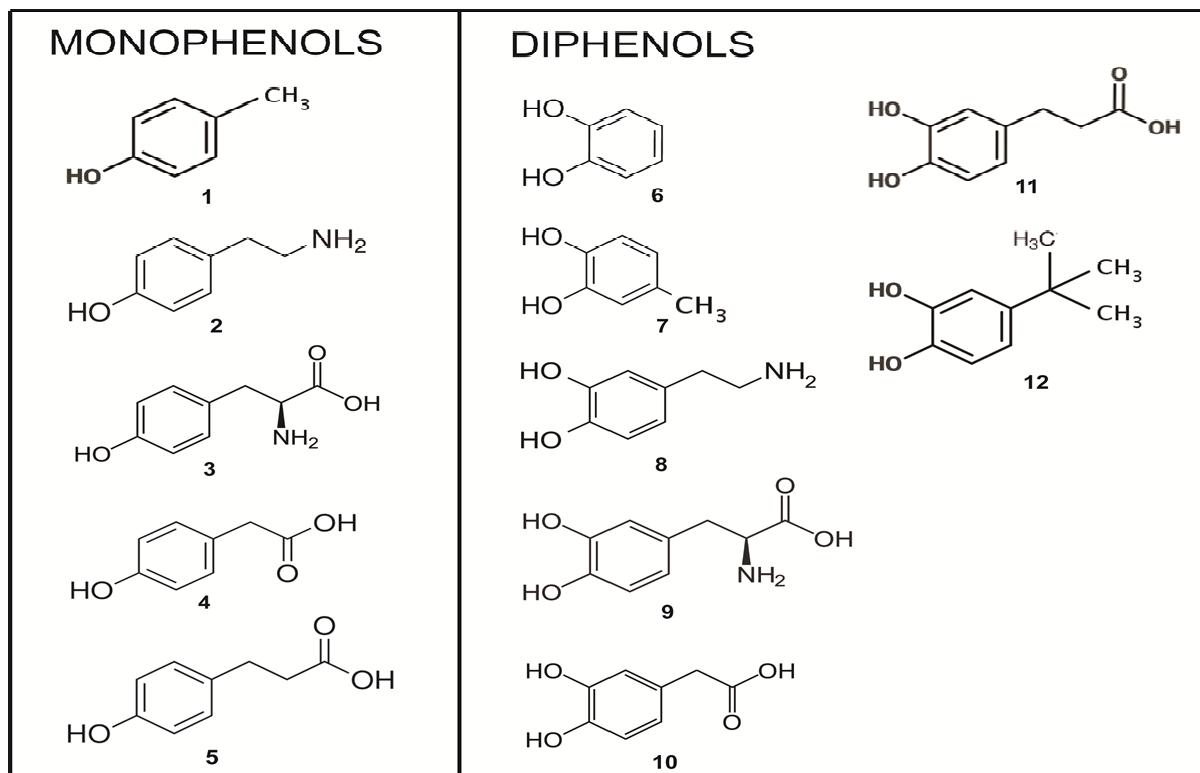
\*The active site residues that commonly interact with all the substrates are highlighted in red.

**Table – 2:** Kinetics parameters determined for native KLH using enzyme assays both in the absence and presence of SDSs.

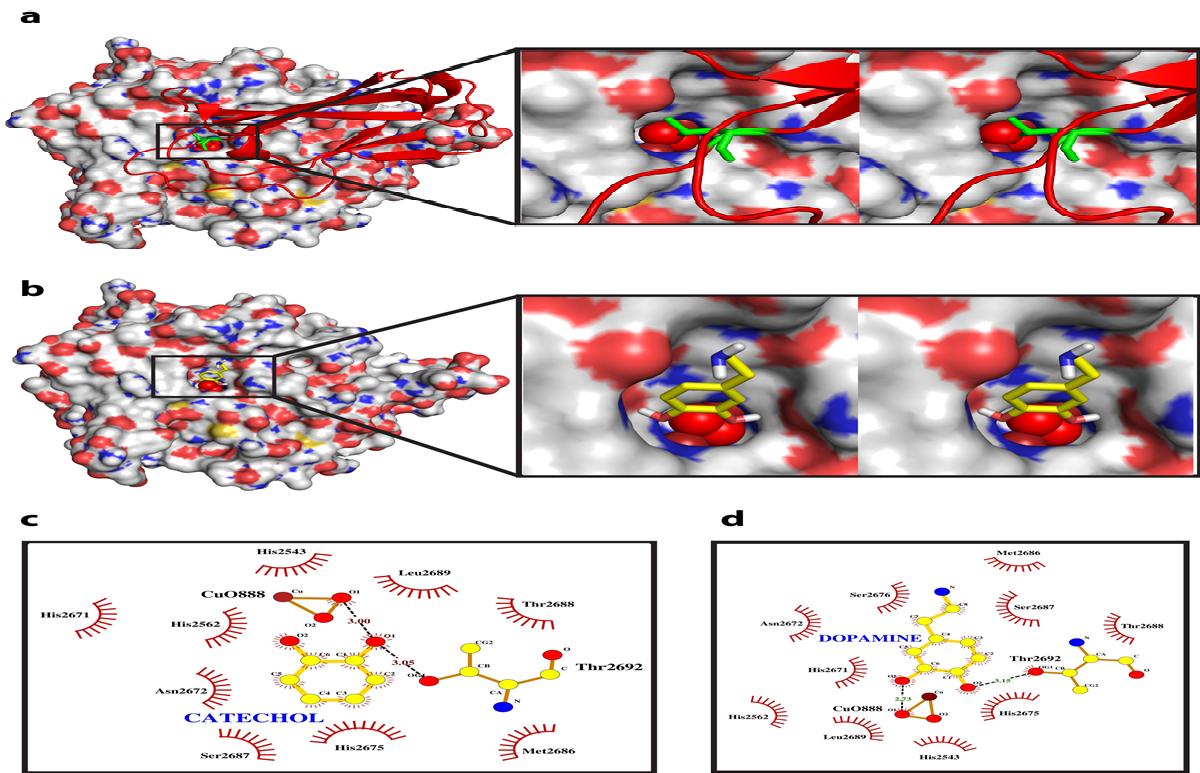
Kinetics Parameters	Catechol		Dopamine	
	Without SDS	With SDS	Without SDS	With SDS
$K_m$ (mM)	23.5	15.8	25.2	15.6
$V_{max}$ (mM.min <sup>-1</sup> .mg <sup>-1</sup> )	0.09	0.084	0.012	0.03
$K_{cat}$ (per second)	1500	1400	200	500
$K_{cat}/K_m$	63.8	88.6	7.9	32.0



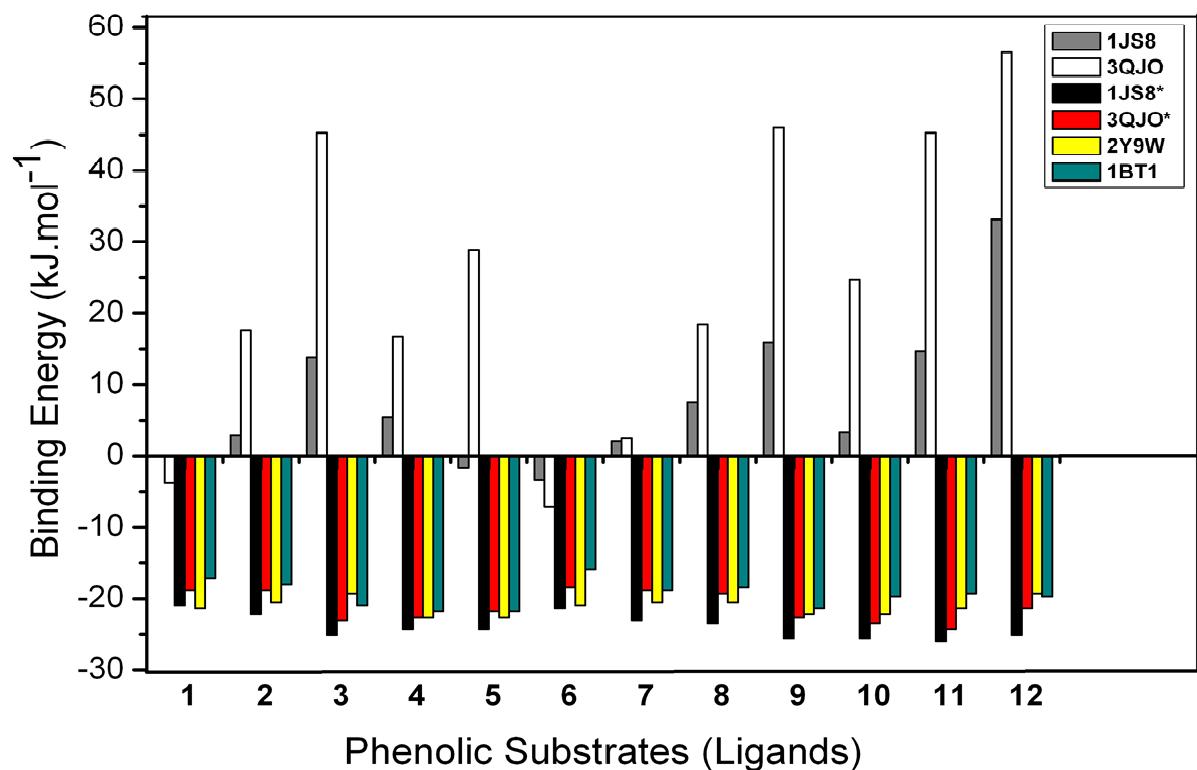
Figure\_1\_Naresh\_ArunSreekumar\_Rajan .



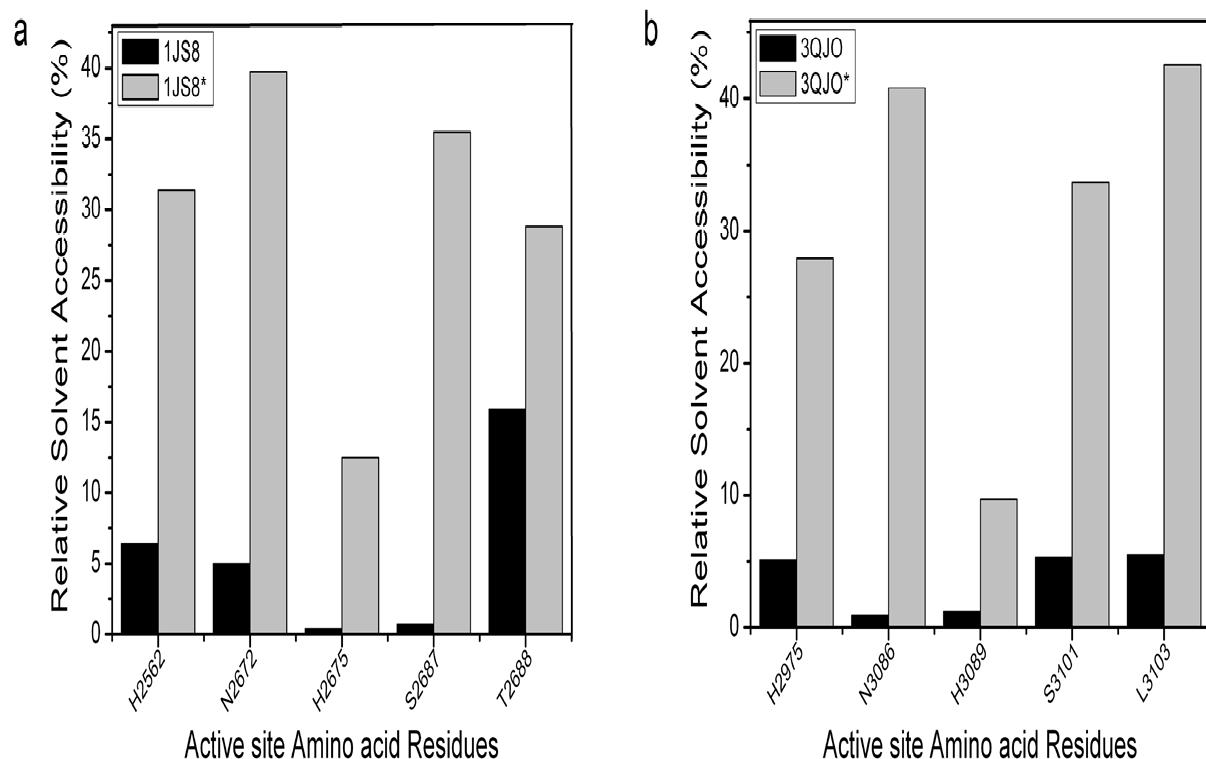
Figure\_2\_Naresh\_ArunSreekumar\_Rajan .



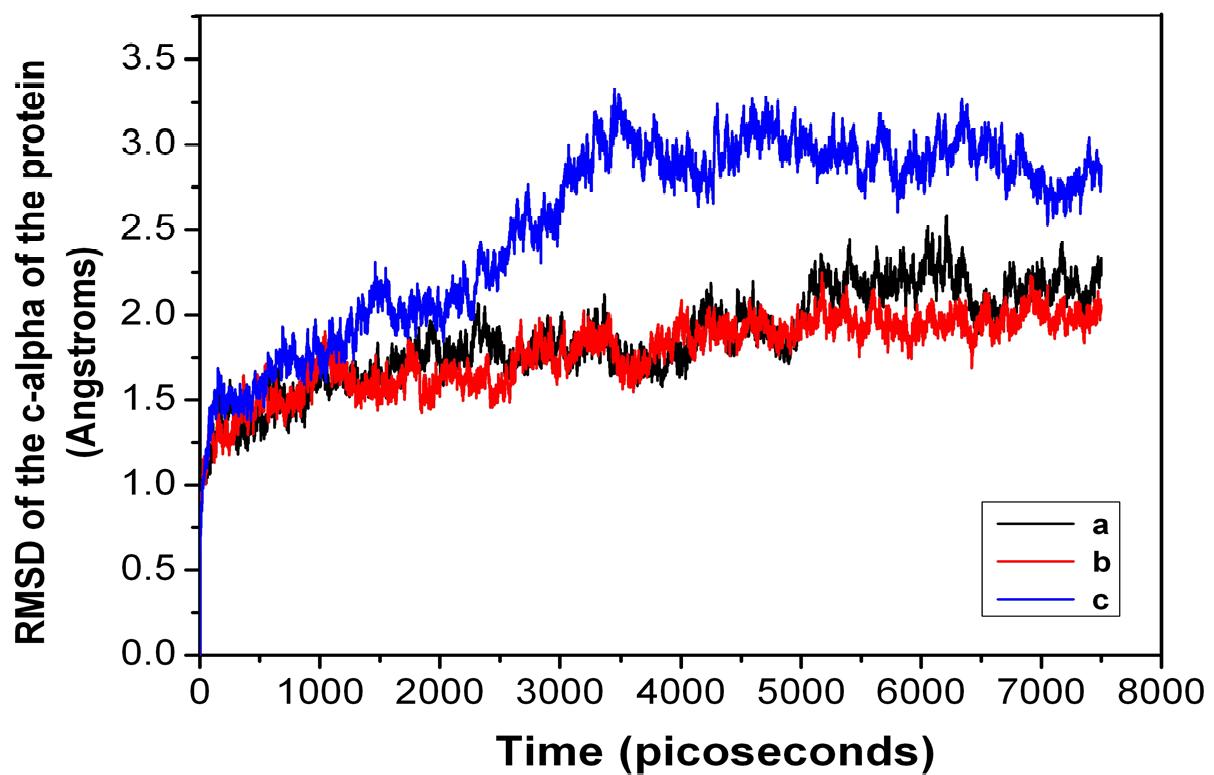
Figure\_3\_Naresh\_ArunSreekumar\_Rajan .



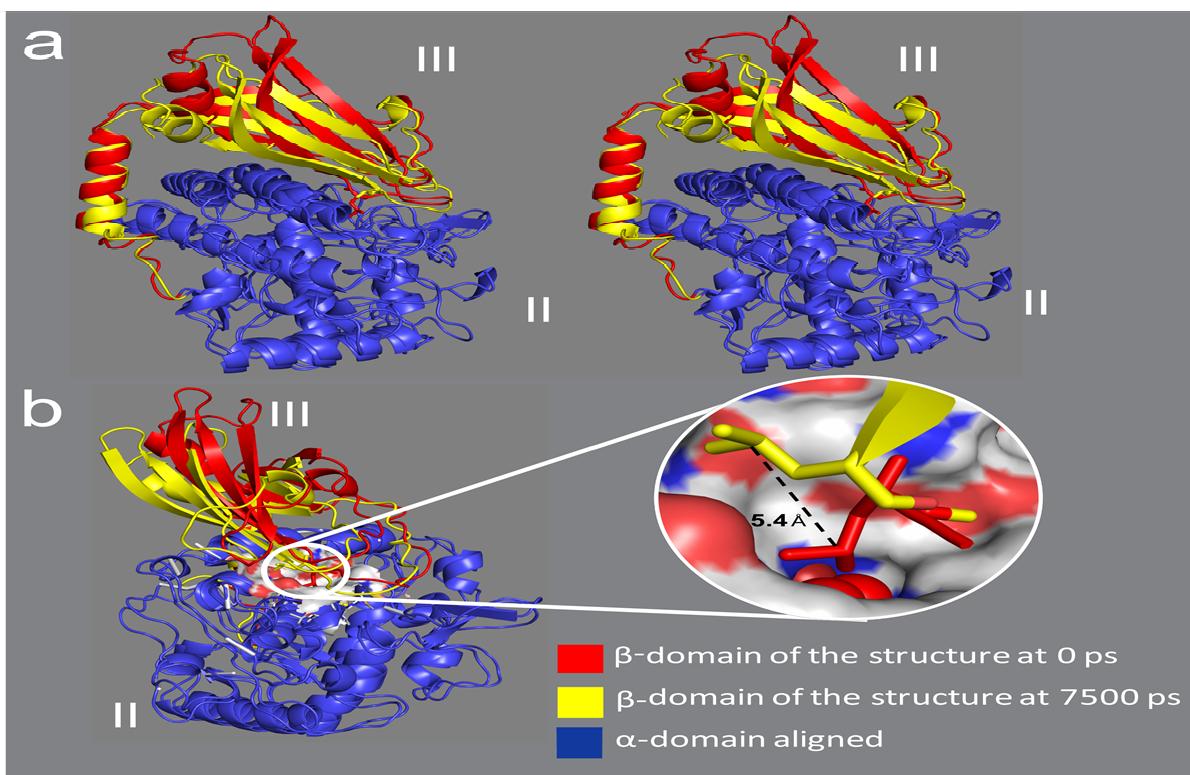
Figure\_4\_Revised .



Figure\_5\_Naresh\_ArunSreekumar\_Rajan .

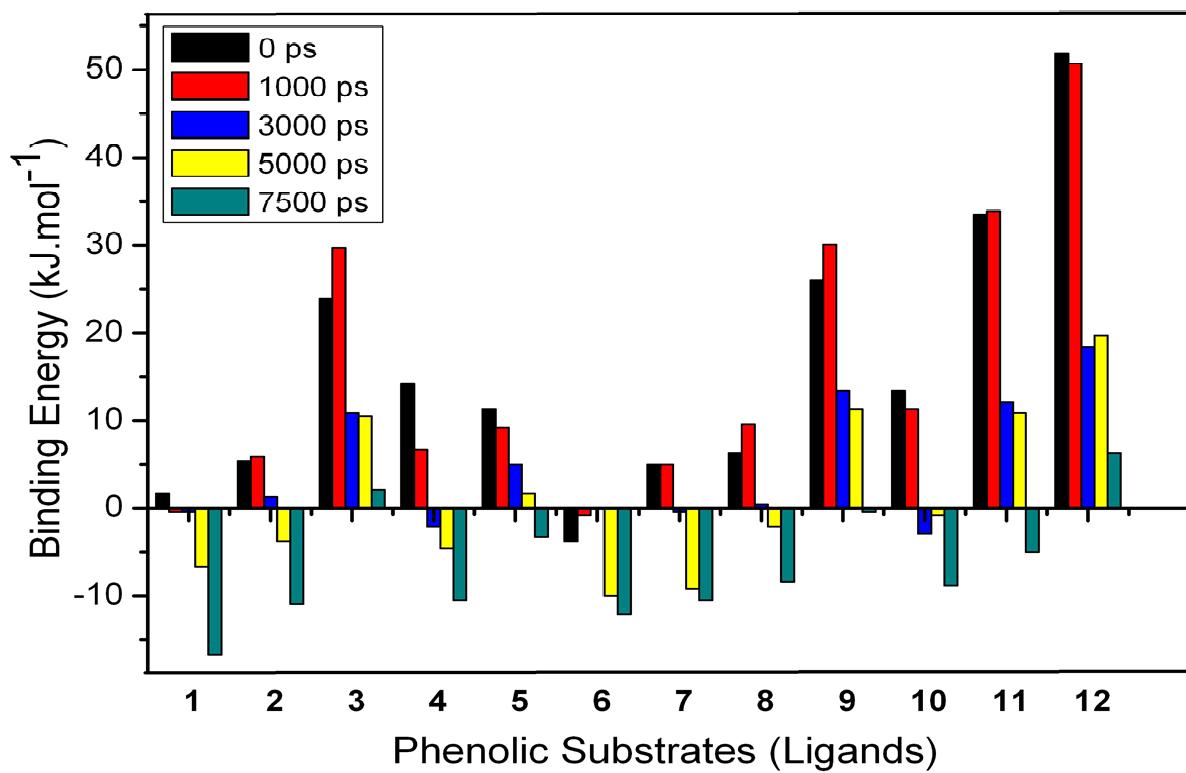


Figure\_6\_Naresh\_ArunSreekumar\_Rajan .

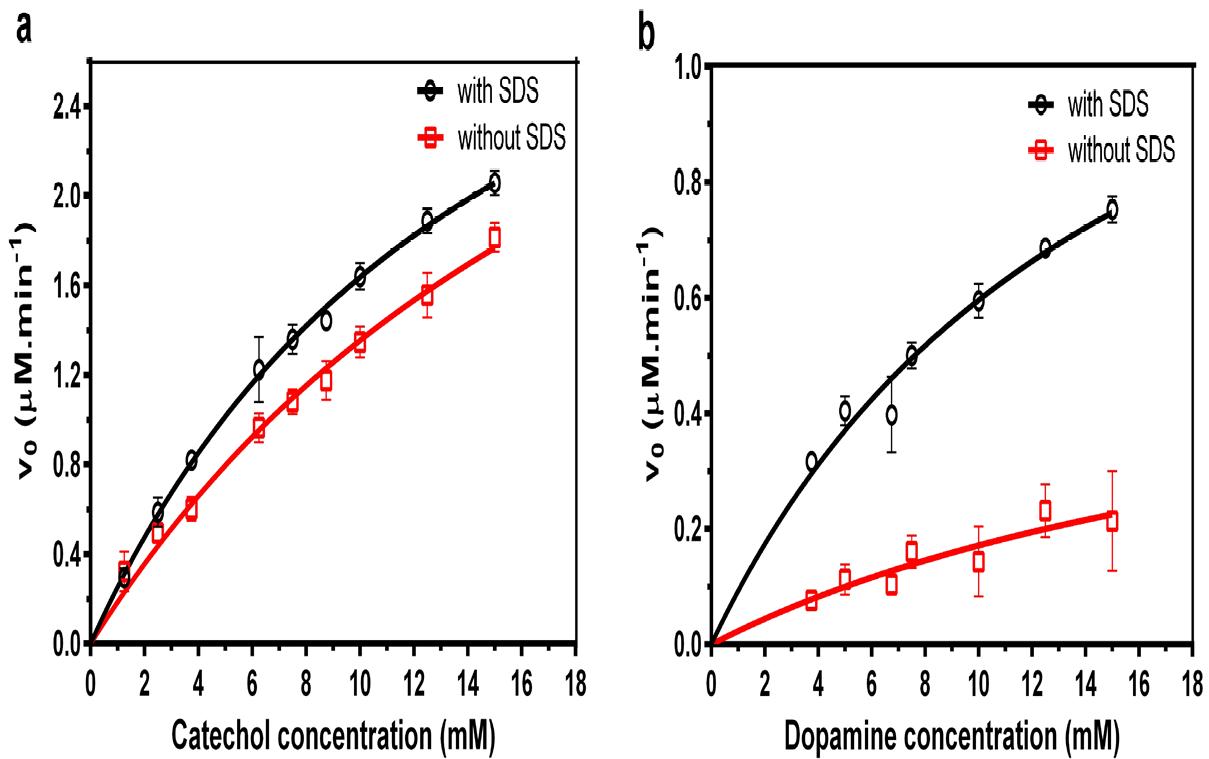


Figure\_7\_Naresh\_ArunSreekumar\_Rajan .

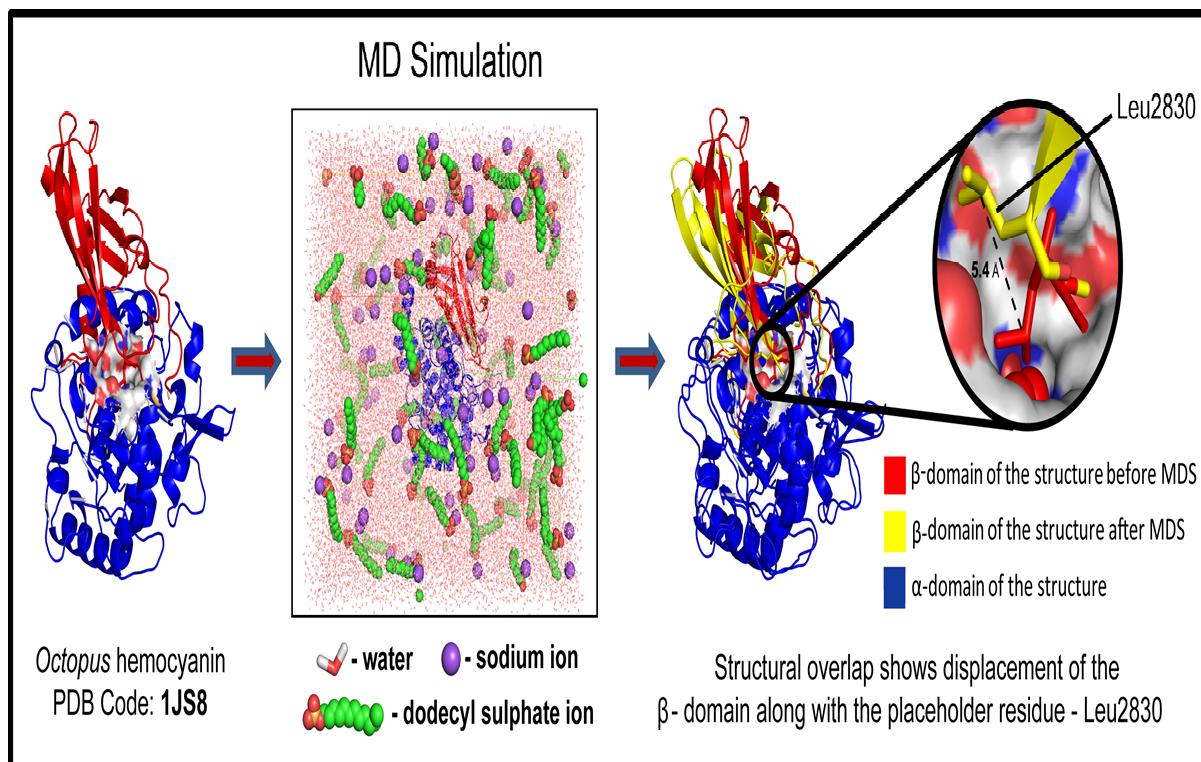
Figure\_7\_Naresh\_ArunSreekumar\_Rajan .



Figure\_8\_Revised .



Figure\_9\_Naresh\_ArunSreekumar\_Rajan .



Graphical\_Abstract\_Naresh\_ArunSreekumar\_Rajan .

Graphical\_Abstract\_Naresh\_ArunSreekumar\_Rajan .