Indian Journal of Biochemistry & Biophysics
Vol. 58, June 2021, pp. 284-291

Mutational analysis of resveratrol-cleaving dioxygenase towards enhancement of vanillin synthesis

Pinki Dey & Pravin Kumar*

School of Computational and Integrative Sciences, Jawaharlal Nehru University, New Delhi-110 067, Delhi, India

Received 09 September 2019; *revised 13 January 2020*

To create better enzyme for enhanced yield of industrial products is a challenging task in protein engineering. It requires lots of resources to manipulate or mutate enzymes experimentally. Computational biology helps in studying the mutational effects on enzymes' structure and stability. The enzyme selected for our study is a Resveratrol-cleaving Dioxygenase that convert resveratrol to vanillin. Using comprehensive computational investigation, we studied the catalytic activity of Resveratrol-cleaving Dioxygenase at the molecular level. We introduced site directed mutations to create mutants of the Resveratrol-cleaving Dioxygenase and studied its interactions with resveratrol by Molecular docking and Molecular dynamics simulations. Here, by investigating the interactions between the mutant Resveratrol-cleaving Dioxygenase and the substrate allow us to highlight the improved performance of mutants over the wild type Resveratrol-cleaving Dioxygenase. It was observed that mutant1 is as stable as wild type of enzyme and has a better affinity toward substrate as the formation of hydrogen bond has increased as compared to wild type. Hence we propose that mutant1 if created in lab could provide better yield.

Keywords: Protein engineering, Resveratrol-cleaving dioxygenase, Vanillin

Vanillin is a crystal powder with vanilla like odour. It is sweet in taste and is one of the most used flavouring agents in the food industry[1]. The uses of vanillin are widespread ranging from cosmetics to pharmaceutical industries. Moreover, vanillin is also reported to have anti-cancer properties[2,3] making it a potential antitumor drug candidate.

Typically vanillin is extracted from the seed pods of *Vanilla planifolia*, generally grown in the tropical regions[4]. However, the tradition process of vanillin extraction is very time consuming and could not meet the demand of end consumers[5]. Therefore, new methods such as chemical synthesis and biotransformation etc. are devised to enhance vanillin production. As the product of chemical synthesis is not considered as natural and leads to environmental pollution, the alternative approach of biotransformation is being preferred for vanillin production[6].

Currently, various enzymes are employed for the biotransformation of natural compounds, like lignin, phenolic aldehydes etc. into vanillin[7–9]. Resveratrol-cleaving Dioxygenase converts a wide range of stilbenes compounds into vanillin. Resveratrol is derived from the stilbenes which are abundantly found in grapes and peanuts. Moreover, consumption of resveratrol is associated with many health benefits as it lowers the rate of cardiovascular disease and cancer[10]. It was first initially found in *Sphingomonas paucimobilis*[11]. Subsequently other Resveratrol-cleaving Dioxygenase enzymes also known as NOV1 was extracted from *Novosphingobium aromaticivorans*[12] that cleaves stilbenes into vanillin in presence of oxygen. The schematic chemical reaction is shown in (Fig. 1).

However, the process of biotransformation is expensive with low yield. Hence, computational approaches are used to analyse the binding region of Resveratrol-cleaving Dioxygenase and design mutants that could increase the yield of vanillin. In our study, the motivation is to engineer an enzyme with better substrate affinity to facilitate vanillin release. Here, we employ a computational approach to enhance the enzyme activity before it can be modified experimentally. This in turn will save a lot of time and resources and provide a molecular view of the substrate binding by the protein.

In this paper, we introduce mutations in the catalytic site of Resveratrol-cleaving Dioxygenase and investigated the structural changes in the enzyme on binding with the ligand *i.e.*, resveratrol. Using

Correspondence:
E-mail: pinki16_sit@jnu.ac.in

molecular dynamics simulations, we hypothesized that change in the active site of the enzyme that promotes better interactions with substrates and facilitates vanillin production[13]. In our study, we selected two mutants and investigated their structural changes upon ligand binding. We selected the mutants in such a way that they promote better hydrogen bond formation between the enzyme and the ligands. In addition to the changes in the stability to the mutated structures, we also investigated the variations in the active site flexibility of mutated protein using principal component analysis and dynamic cross correlation maps.

Materials and Methods

Preparation of protein and ligand

The X-ray crystal structure of Resveratrol-cleaving Dioxygenase was reported (PDB ID: 5J54.pdb) by McAndrew*et al.* in 2016[14]. This structure was used to create two mutants by mutating the binding site amino acids. The mutants are grouped as mutant1 and mutant2. The list of mutated amino acids is presented in (Table 1).

The mutation studies were performed based on the prescription followed by McAndrew *et al.*[15]. The mutants were prepared by PyMol's mutagenesis wizard. The Phenylalanine and Histidine residues were replaced with Glutamic acid in mutant1 and Serine in mutant2, respectively, to promote the hydrogen bond formation with the ligand.

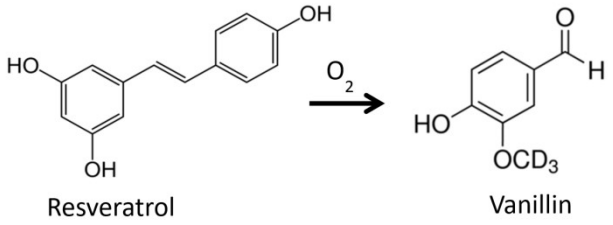

Fig. 1 — Reaction pathway of Resveratrol-cleaving Dioxygenase where resveratrol is converted into vanillin in the presence of oxygen

Table 1 — Table describing the positions of the active site amino that are mutated in Resveratrol-cleaving Dioxygenase

Mutants	Amino acid in wild type enzyme	Amino acids in Mutated enzymes
Mutant1	PHE39	GLU59
	HIS218	GLU218
	HIS284	GLU284
	PHE307	GLU307
Mutant2	PHE39	SER59
	HIS218	SER218
	HIS284	SER284
	PHE307	SER307

The three-dimensional structure of the wild type Resveratrol-cleaving Dioxygenase and the mutants are presented in (Fig. 2). The binding site amino acid residues are highlighted and labelled for all three structures (wild type and the mutants).

We extracted the structure of the ligand from Zinc Database[16] and MGL Tools converted the Mol2 format of the ligand to PDBQT structure format.

Molecular docking studies

We performed the docking study by Autodockvina v1.2[17] due to its accuracy and high speed[18–20]. The docking grid box of 25Å × 25Å × 25 Å was selected such that it covers the entire protein's active site and Broyden-fletcher-Gold-farb-shanno algorithm is used for local optimization of the docked structures. The values were set at default and the torsional angles were allowed to rotate during the docking process. From the AutoDock analysis, we see that the binding of the wild type protein and its mutants to the ligand resveratrol are favourable as they display a negative binding energy (kcal/mol) and can be used for the further analysis (see Table 2).

Molecular dynamic simulation of Resveratrol-cleaving Dioxygenase and its mutants and Mutant-ligand complex

We performed the molecular dynamic simulations using GROMACS 5.5 package andAmber99sb-ildn

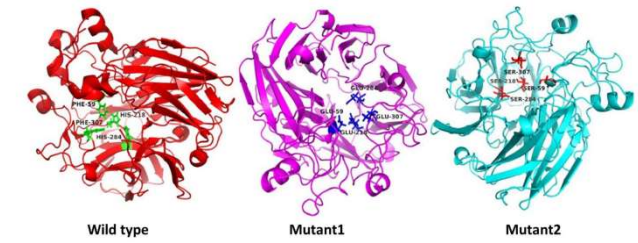

Fig. 2 — Schematic representation of the tertiary structure of Resveratrol-cleaving Dioxygenase (shown in red), mutant1 (shown in violet) and mutant 2 (shown in cyan). The binding site amino acid residues are highlighted and labelled in all the three structures

Table 2 — The binding energy of the wild type protein and its variants as obtained from AutoDock vina

Protein Type	Binding Energy (kcal/mol)
Wild type	–7.7
Mutant1	–7.9
Mutant2	–8.0

force field[20]. The structures were solvated with TIP3P[21] water model. The WT protein was solvated in a cubic box while a dodecahedron solvation box was used for the mutants. We kept a minimum distance of 0.9 nm between the protein surface and solvation box. Steepest descent algorithm was used for energy minimization.

The protein and ligand topology files were created using Dundee Prodrg 2.5[22]. The protein-ligand was then immersed in a hexagonal box and solvated using SPC/E water model. The complex was then neutralized using sodium or chlorine atoms. After solvation, the conflicting contacts were relaxed by energy minimization using steepest descent algorithm. The energy minimized structures were equilibrated followed by subsequent production phase. The wild type, mutants, counter ion and ligand was position-restrained during the equilibration. We performed 50ns MD simulations at 300 K temperature and 1 bar pressure using periodic boundary condition.

We used the GROMACS utilities for analysis of the simulation trajectories and data visualization and representation was performed using Pymol, VMD[23], Xmgrace and MATLAB.

Dynamic Cross Correlation map

Dynamic cross-correlation maps (DCCM)[24] give the time correlated motions of all the C-alpha atom pairs in the protein molecule. The following equation was used to calculate the dynamic cross-correlation matrix:

$$C_{ij} = \frac{\langle \Delta r_i \Delta r_j \rangle}{(\langle \Delta {r_i}^2 \rangle \langle \Delta {r_j}^2 \rangle)^{1/2}} \quad \ldots (1)$$

where i and j denotes the i-th and j-th atom, respectively, and r_i andr_j are the displacement vector. <.....> indicates the ensemble average. Positive values of C_{ij} indicate correlated atomic motions in the same direction whereas negative value of C_{ij} denotes anti-correlated motion of the atoms. The DCCM matrix was plotted using gnuplot[25].

Principal component analysis

We performed Principal Component analysis (PCA) analysis following the prescription of Amadei *et al*.[25] in order to investigate the motion of C-alpha atoms throughout simulation. A covariance matrix of the fluctuations of C-alpha atoms was constructed and the eigenvectors and eigenvalues calculated. The eigenvector gives the direction of motion whereas, eigen values indicates the mean square fluctuation of the atoms. We plotted the PCA using the following equation:

$$C_{ij} = \langle (r_i(t) - \langle r_i \rangle_t)(r_j(t) - \langle r_j \rangle_t) \rangle_t \quad \ldots (2)$$

where <> denotes the average fluctuation over time and r_i, $r_j$denotes the cartesian coordinates of the i-th and j-th atom respectively.

Results

The active site's detailed analysis of the Resveratrol-cleaving Dioxygenase and its binding with resveratrol is demonstrated using extensive molecular dynamics simulations. Here, we introduced mutations at four active site residues of Resveratrol-cleaving Dioxygenaseenzyme. After creating the mutants, the conformational stability of wildtype protein and its mutants and their effective binding to the ligand were studied using molecular dynamic simulation.

Structural analysis of Resveratrol-cleaving Dioxygenaseand its mutants

The stability of the structures was monitored during simulation by calculating the root mean square deviation (RMSD) from the starting structures. The RMSD plot is reported in (Fig. 3). The initial increase

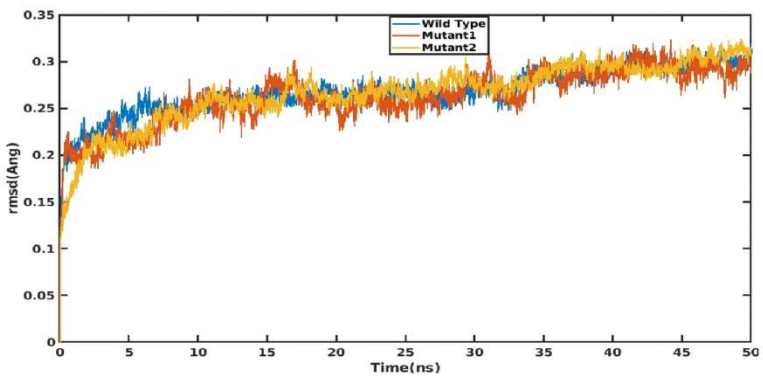

Fig. 3 — Variations in the RMSD of wild type (blue), mutant1 (red) and mutant2 (yellow) as a function of simulation time. All the three structures are stable during the 50 ns simulation

in the value of RMSD of all the models (both wild type and mutants) is due to relaxation of the structures. The RMSD values reached a plateau after 10 ns after which all models were stable during the 50 ns simulation.

This indicates that the introduction of mutations in the WT protein doesn't impact the overall conformational stability of the protein molecule. In order to identify the higher flexibility region, we calculate the per residue root mean square fluctuation (RMSF) as a function of simulation time. From Figure 4, we found a similarity in the per residue fluctuation of the mutant1 and the wild type protein whereas the fluctuation is higher in the case of mutant2 as compared to that of the wild type protein. It has been reported the active site residues of Resveratrol-cleaving Dioxygenase that are involved in cleaving resveratrol into vanillin and are F59, Y101, K134, H218, H284, S283, F307, E353. Out of these amino acids E353, S283, Y101 and K134 form hydrogen bond with resveratrol[14]. The position of the active site residues that are mutated are shown by the arrows which indicates a higher fluctuation for the mutant 2 active site residues as compared to mutant1.

Effect of mutations on the motion of the protein

We calculated the Dynamic cross correlation Map (DCCM) to investigate the change in correlated motion of the wild type and the mutated proteins. The DCCM analysis provided a complete picture of the correlated motion of atoms as presented in (Fig. 5). The strongly correlated region is represented in blue while the region coloured in red is strongly anti-correlated.

In the wild type (Fig. 5A), it was observed that there is a strong correlation between residue 1 to 210. Anti-correlation can be observed between residues 210-350. The position of the active site groove is such that it lies between residues 1-210 and residues 210-479. The position of the active amino acids are 59, 101, 121, 134, 135, 216, 218 283, 284, 307, 353 and 354. In mutant1 (Fig. 5B), we observe an anti-correlated motion between residue 1 to 210, same for the residue 210 to 479. In mutant2 (Fig. 5C), the correlation between residue 1 to 370 has increased while the anti-correlation has increased between residue 350-479.

The overall movement of the protein was identified using Principal component analysis(PCA). The PCA of the C_α carbon atom of the wild type protein and its mutants are shown in (Fig. 6). The porcupine plot of the wild type of enzyme is presented in (Fig. 6A). PCA identifies the pattern of the movement of the C_α carbon atoms in the protein. It helps in identifying the total coordinated motion of the C_α atoms in the protein expressed by the eigenvectors of the covariance matrix

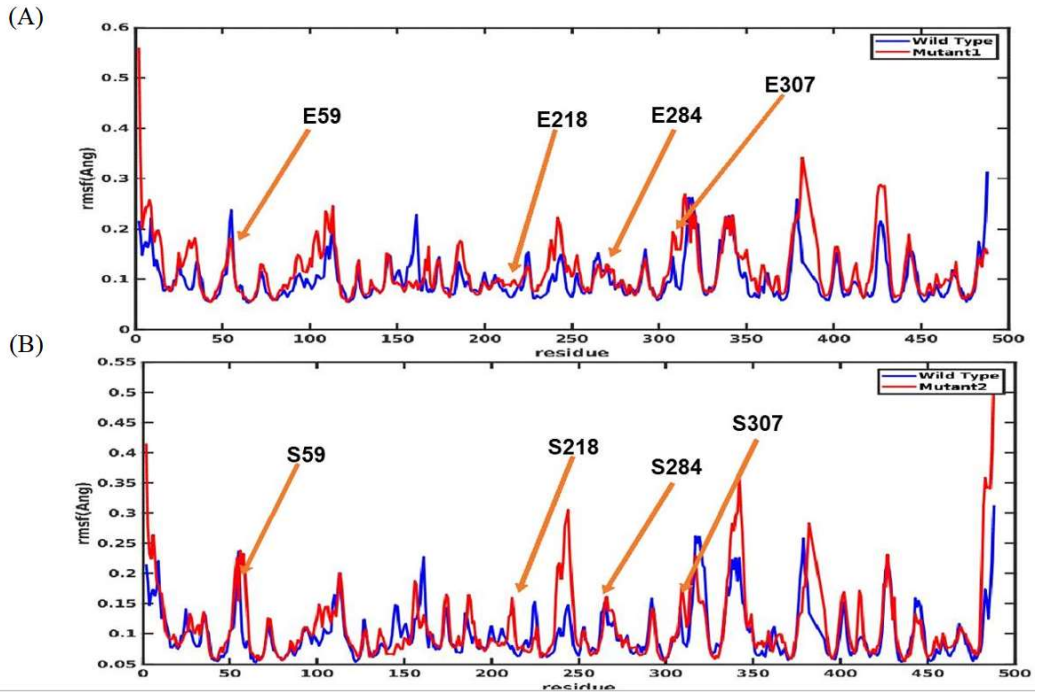

Fig. 4 — Variation in the RMSF of the residues for (A) wild type (blue) and mutant1 (red) (B) wild type (blue) and mutant2 (red) as a function of the protein residue number

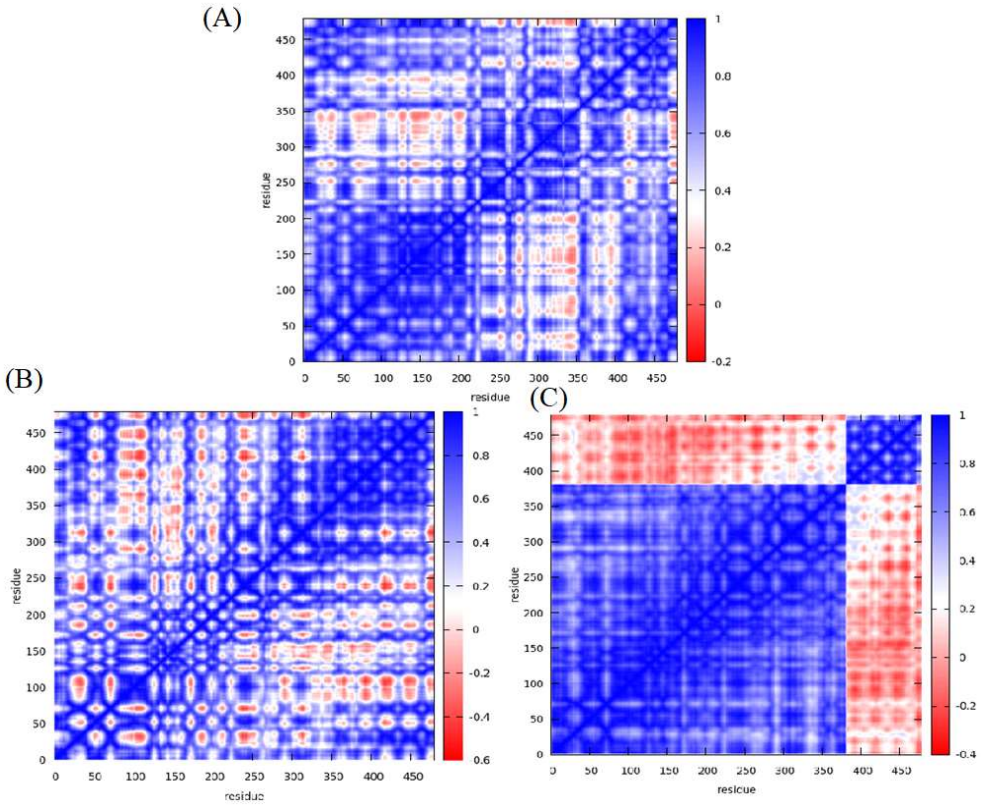

Fig. 5 — Residue based correlated motion for the C_α carbon atoms in the wild and mutated models during the 50 ns of time average. The blue coloured region blue represents the strong correlated motion while the red represents anti-correlated region of (A) wild type; (B) mutant1; and (C) mutant2

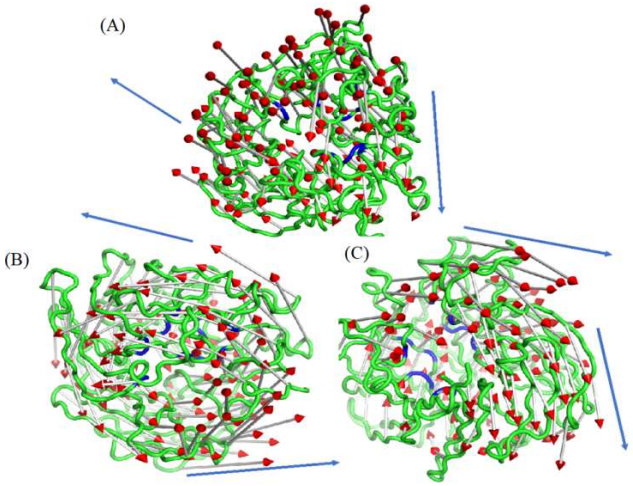

Fig. 6 — Porcupine plot of (A) wild Type (B) Mutant1 (C) Mutant2 as generated from the first eigenvector of Resveratrol-cleaving Dioxygenase and its mutants. The arrow indicates the direction of motion. The active site residues are highlighted in blue

which is implied by the corresponding eigen values. Figure6 depicts the motion of the protein along its first eigenvector. In the wild type protein as shown by the blue arrow the two parts of the protein are moving opposite to each and hence facilitates the opening of the active site so that the ligand molecule can bind to it. In mutant1 same type of movement can be seen. However, in mutant2 movement along the first eigenvector might not help in the opening of the enzyme active site for substrate binding.

Molecular dynamic simulation of Resveratrol-cleaving Dioxygenase and its variants in complex with resveratrol

The MD simulation of wild type-ligand complex and the mutant-ligand complexes were performed to find out their corresponding stability. The trajectories were analysed by the RMSD plot and time evolution of hydrogen bond formations at the active site with ligand during 50 ns simulation. The RMSD plot is presented in (Fig. 7). It can be observed from the (Fig. 7) that all the structures have plateaued after 7 ns. Mutant2 has more fluctuation than wild type and mutant1. But overall, all the structures are stable during 50 ns simulation.

The hydrogen bond formation between the wild type and mutants' active site and ligand during 50 ns simulation has been presented in (Fig. 8).

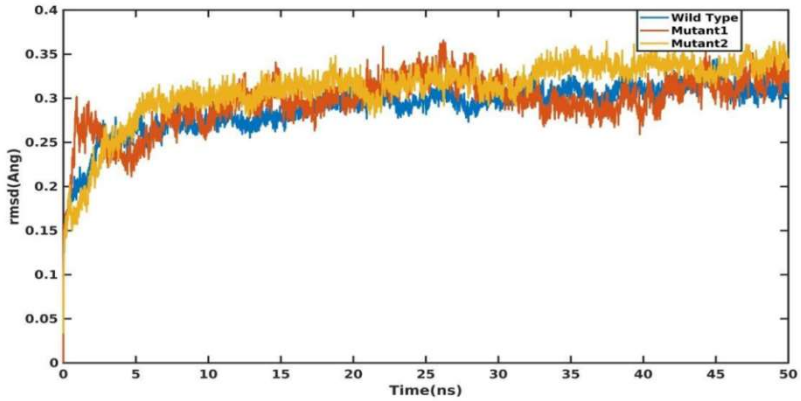

Fig. 7 — RMSD *vs* time plot of phenolic stilbenes and variants after docking. The wild type Resveratrol-cleaving Dioxygenase is shown in blue, mutant1 is shown in red and mutant2 is shown in yellow

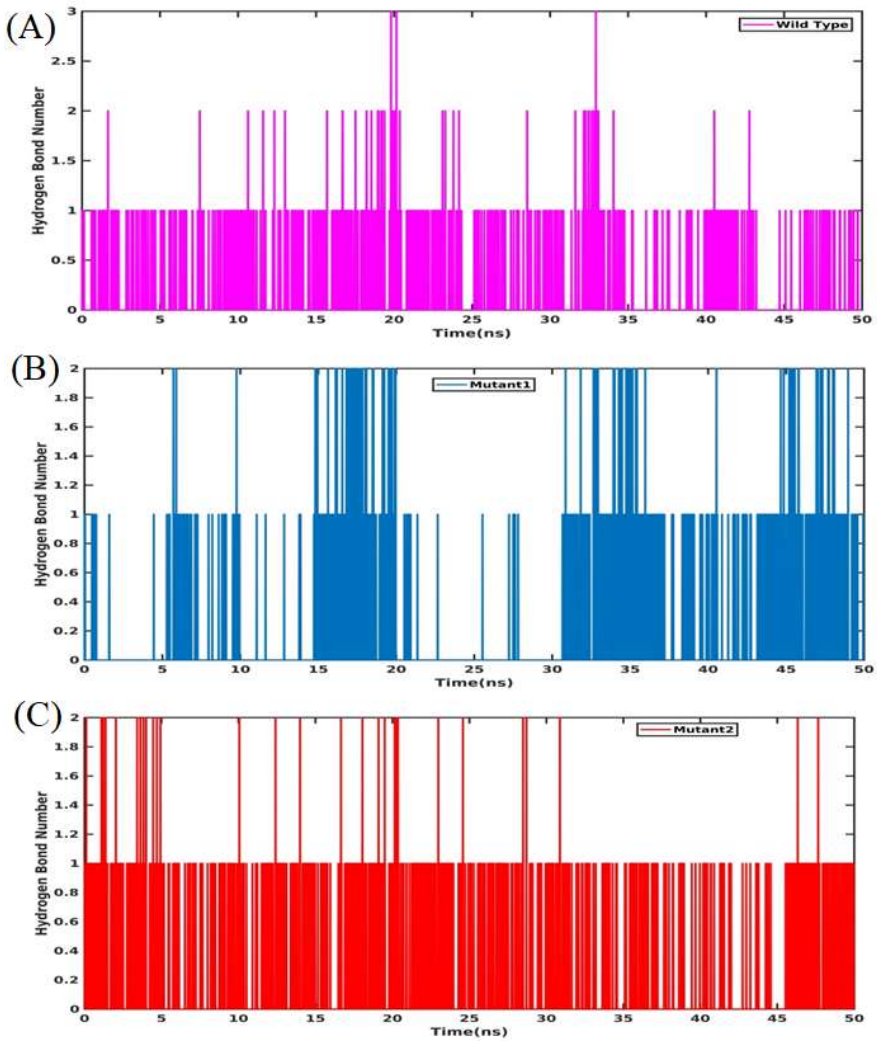

Fig. 8 — Variation in the hydrogen bond formation for (A) wild type, (B) mutant1, (C) mutant2 as a function of simulation time

The average number of hydrogen bond formed during 50 ns is tabled in (Table 3). We see that the average number of a hydrogen bond is highest in the mutant1 as presented in (Table 3) and can be observed from (Fig. 8). Mutant2 forms less number of hydrogen bond than wild enzyme and mutant1.

Table 3 — The average number of hydrogen bond formation between the protein and ligand

Protein Type	Average number of hydrogen bond formation during simulation
Wild type	0.168
Mutant1	0.210
Mutant2	0.147

Discussion

In order to increase the efficiency of vanillin production by Resveratrol-cleaving Dioxygenase enzyme by promoting better interactions with substrates, two mutants were prepared and named as mutant1 and mutant2. To check the stability of these mutants compared to wild type Resveratrol-cleaving Dioxygenase enzyme, RMSD and RMSF plot was plotted. A lower RMSD indicates the stability of protein[26,27] and it was found that the average RMSD of the wild type and variants is lower than 0.3 Å which indicates the stability of the mutant structures[28]. Also, RMSF plot indicates high fluctuation in the mutant2 whereas the mutant1's fluctuation is similar to that of the wild type enzyme. Moreover, the DCCM studies indicate that for mutant2, the overall correlation has increased which might affect the opening of the active site but in mutant1 DCCM plot has not deviated too much from the wild type of enzyme. Further, to check the overall movement of mutants, principal component analysis was performed. It was found that mutant1 protein's movement is similar to the wild type while the movement of mutant2 is such that it doesn't help in the opening of the active site.

To further elucidate the stability of mutants after docking with resveratrol, RMSD was plotted and it was found all the structure are stable as average RMSD value of all structure is below 0.3 Å. But, hydrogen bonding analysis between resveratrol and mutants has shown that the hydrogen bond formation has increased for mutant1 and resveratrol complex as compared to wild type and mutant2. In conclusion it can be said that mutant1 which has glutamine at 59 (original Phe), 218 (original His), 284 (original His) and 307th (original Phe) position is able to improve the original wild type of enzyme as the stability of mutant1 has not been affected and it is able to form more hydrogen bonds with resveratrol thereby increasing its efficiency to cleave the ligand, resveratrol.

Conclusion

In our study, we present an in depth computational analysis of the structural basis of Resveratrol-cleaving Dioxygenase substrate binding and product release. By generating two different mutants of the wild typeResveratrol-cleaving Dioxygenase, we investigate the requirements of the wild type enzyme in the catalytic site for better interaction with substrate resveratrol. The RMSD and RMSF studies of wild type enzyme and mutants showed that all the structures are stable. To check the movement of mutants, DCCM and PCA analysis were performed. The DCCM and PCA analysis of mutant1 indicate that its binding region's movement is similar to that of the wild type enzyme. But, considerable changes are observed in the movement of the binding region of mutant2 suggesting that it might affect resveratrol binding. Stability of the enzymes after docking was checked by plotting RMSD value and all enzymes were found to be stable during the simulation. However, mutant1 shows a maximum number of a hydrogen bond formation with resveratrol. From these results it can be concluded that mutant1 has a better affinity toward substrate *i.e*. resveratrol.

Acknowledgement

We gratefully acknowledge the financial support from DST India (DST PURSE, DST/INSPIRE/04/2013/000100, DST-SERB ECR/2016/000188) and JNU (UPoE) research grant.

Conflict of interest

All authors declare no conflict of interest.

References

1 Priefert H, Rabenhorst J & Steinbüchel A, Biotechnological Production of Vanillin. *Appl. Microbiol Biotechnol*, 56 (2001) 296.

2 Lirdprapamongkol K, Kramb JP, Suthiphongchai T, Surarit R, Srisomsap C, Dannhardt G & Svasti J, Vanillin suppresses metastatic potential of human cancer cells through pi3k inhibition and decreases angiogenesis *in vivo*. *J Agric Food Chem*, 57 (2009) 3055.

3 Bezerra DP, Soares AKN & de Sousa D, Overview of the Role of vanillin on redox status and cancer development. *Oxid Med Cell Longev*, (2016) Article ID 9734816.

4 Yang H, Barros-Rios J, Kourteva G, Rao X, Chen F, Shen H, Liu C, Podstolski A, Belanger F, Havkin-Frenkel D & Dixon RA, A re-evaluation of the final step of vanillin biosynthesis in the orchid *vanilla planifolia*. *Phytochemistry*, 139 (2017) 33.

5 Krings U & Berger RG, Biotechnological Production of Flavours and Fragrances. *Appl Microbiol Biotechnol*, 49 (1998) 1.

6 Baqueiro-Peña I & Guerrero-Beltrán JÁ, Vanilla (*Vanilla planifolia* Andr.), Its residues and other industrial by-products for recovering high value flavor molecules: A Review. *J Appl Res Med Aromat Plants*, 6 (2017) 1.

7 Zamzuri NA & Abd-Aziz S, Biovanillin from agro wastes as an alternative food flavour. *J Sci Food Agric*, 93 (2013) 429.

8 Vaithanomsat P & Apiwatanapiwat W, Feasibility study on vanillin production from *Jatropha Curcas* stem using steam explosion as a pretreatment. *Inter J Chem Biolo Engr*, 2 (2009) 211.
9 Fraaije M W& Laane C, Enzymatic synthesis of vanillin. *J Agric Food Chem*, 49 (2001) 2954.
10 Yu W, Fu YC & Wang W, Cellular and molecular effects of resveratrol in health and disease. *J Cell Biochem*, 113 (2012) 752.
11 Kamoda S & Saburi Y, Cloning, Expression, and Sequence Analysis of a Lignostilbene-α, β-Dioxygenase gene from *Pseudomonas Paucimobilis* TMY1009. *Biosci Biotechnol Biochem*, 57 (1993) 926.
12 Marasco EK & Schmidt-Dannert C, Identification of bacterial carotenoid cleavage dioxygenase homologues that cleave the interphenyl α, β double bond of stilbene derivatives via a monooxygenase reaction. *ChembioChem*, 9 (2008) 1450.
13 Trylska J, Grochowski P & McCammon JA, The role of hydrogen bonding in the enzymatic reaction catalyzed by HIV-1 protease. *Protein Sci*, 13 (2004) 513.
14 McAndrew RP, Sathitsuksanoh N, Mbughuni MM, Heins RA, Pereira JH, George A, Sale KL, Fox BG, Simmons BA & Adams PD, Structure and mechanism of nov1, a resveratrol-cleaving dioxygenase. *Proc Natl Acad Sci*, 113 (2016) 14324.
15 McAndrew RP, Sathitsuksanoh N, Mbughuni MM, Heins RA, Pereira JH, George A, Sale KL, Fox BG, Simmons BA & Adams P, Structure and mechanism of NOV1, A resveratrol-cleaving dioxygenase. *Proc Natl Acad Sci*, 113 (2016) 14324.
16 Irwin JJ & Shoichet BK, ZINC− A free database of commercially available compounds for virtual screening. *J Chem Inf Model*, 45 (2005) 177.
17 Trott O & Olson AJ, AutoDock vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*, 31 (2010) 455.
18 Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS & Olson AJ, AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem*, 30 (2009) 2785.
19 Abraham MJ, van der Spoel D, Lindahl E &Hess B, The GROMACS development team, GROMACS User Manual Version 5.1. 2. 2016.
20 Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO & Shaw D, Improved side-chain torsion potentials for the amber ff99sb 139protein force field. proteins: Structure, Function, and bioinformatics. *Proteins*, 78 (2010) 1950.
21 Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW & Klein ML, Comparison of simple potential functions for simulating liquid water. *J Chem Phys*, 79 (1983) 926.
22 SchuÈttelkopf AW & Van Aalten DM, PRODRG: A Tool for high-throughput crystallography of protein–ligand complexes. *Acta Crystallogr Sect D Biol Crystallogr*, 60 (2004) 1355.
23 Humphrey W, Dalke A & Schulten K, VMD: Visual molecular dynamics. *J Mol Graph*, 14 (1996) 33.
24 Ichiye T & Karplus M, Collective motions in proteins: A covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations. *Proteins Struct Funct Bioinforma*, 11 (1991) 205.
25 Amadei A, Linssen A & Berendsen HJ, Essential dynamics of proteins. *Proteins Struct Funct Bioinforma*, 17 (1993) 412.
26 Hemachandran H, Anantharaman A, Mohan S, Mohan G, Kumar DT, Dey D, Kumar D, Dey P, Choudhury A, Doss CGP & Ramamoorthy S, Unraveling the inhibition mechanism of cyanidin-3-sophoroside on polyphenol oxidase and its effect on enzymatic browning of apples. *Food Chem*, 227 (2017) 102.
27 Priya R, Sneha P, Rivera Madrid R, Doss C, Singh P & Siva R, Molecular modeling and dynamic simulation of *Arabidopsis Thaliana* carotenoid cleavage dioxygenase gene: A comparison with *Bixa Orellana* and *Crocus Sativus*. *J Cell Biochem*, 118 (2017) 2712.
28 Kumar P, Sachan SG & Poddar R, Mutational analysis of microbial hydroxycinnamoyl-coa hydratase-lyase (hchl) towards enhancement of binding affinity: A computational approach. *J Mol Graph Model*, 77 (2017) 94.