

Original Research Paper

# *In-silico* Rational Protein Engineering and Design Approach to Improve Thermostability of a Haloalkane Dehalogenase Enzyme

<sup>1</sup>Raghunath Satpathy, <sup>2</sup>V. Badireenath Konkimalla and <sup>3</sup>Jagnyeswar Ratha

<sup>1,3</sup>School of Life Science, Sambalpur University, Jyoti vihar, Burla, Odisha, 768019, India

<sup>2</sup>School of Biological Sciences,

National Institute of Science Education and Research (NISER), Bhubaneswar, Odisha, 751005, India

## Article history

Received: 24-03-2015

Revised: 05-10-2015

Accepted: 05-10-2015

Corresponding Author:

Raghunath Satpathy

School of Life Science, Sambalpur

University, Jyoti vihar, Burla,

Odisha, 768019, India

Email: rnsatpathy@gmail.com

**Abstract:** Thermostability of enzymes is a major prerequisite for use in industrial enzymology. There are as such no simple general principles for achieving thermostability in case of enzymes as many factors are required to fulfil for different enzymes. The present study describes computational methods to design thermostable haloalkane dehalogenase enzyme using the crystal structure available Protein Data Bank (PDB ID: 1EDE). In *in silico* design strategy rule-based approaches such as disulfide bond geometry, new hydrophobic pocket design, new salt bridge construction and multiple mutations (combination of the above approaches) were introduced to the original enzyme. After each design strategy the functional effect was confirmed in terms of enzyme substrate binding by molecular docking using Autodock vina tool. Best design strategy was evaluated by comparative molecular dynamics simulation applying simulated annealing method at 8 ns using GROMACS tool. The surface hydrophobicity which is the key factor for thermostability in haloalkane dehalogenase was obtained from the simulation result. Upon optimizing the parameters, thermostability of mutant enzyme under consideration was also confirmed by the 5 ns molecular dynamics simulation at 400, 500 and 600 K.

**Keywords:** Thermostability, Haloalkane Dehalogenase, Hydrophobic Pocket Design, Docking, Surface Hydrophobicity, Molecular Dynamics Simulation

## Introduction

Haloalkane dehalogenase enzymes can be used as a potential catalyst for biodegradation of a number of halogenated compounds that are environmentally toxic expelled as industrial by-products (Fetzner and Lingens, 1994; Koudelakova *et al.*, 2013). Along with their potential applications in bioremediation, these enzymes have gained a limited interest in the industry for the production of optically pure alcohols. Therefore, the identification of novel dehalogenating enzymes is very much important in terms of their industrial utility (Bidmanova *et al.*, 2013; Hasan *et al.*, 2011). The catalytic reaction mechanism of the haloalkane dehalogenase enzymes (EC 3.8.1.5) involves the conversion of 1-haloalkane to a primary alcohol and halide ions following a hydrolysis reaction (Wackett,

1994; Prokop *et al.*, 2003). Structurally, these enzymes belong to alpha/beta hydrolase superfamily that specifically acts on halide bonds present in carbon-halide compounds. Their active site is found to be buried in a hydrophobic cavity at the interface formed by hydrolase core domain and another cap domain (Dijkstra, 1993; Oakley *et al.*, 2002) The general mechanism of the catalysis by haloalkane dehalogenase has been studied in a detailed manner where amino acids ASP-124 and HIS-289 align perfectly with the substrate and catalyze the hydrolysis reaction respectively (Marek *et al.*, 2000; Hesseler *et al.*, 2011). By understanding the importance of these enzymes, designing of a thermostable haloalkane dehalogenase would be very much important. The current *in silico* approaches in the protein engineering and design to enhance the thermostability in enzymes have been

employed as a potential method in many events as explained by many researchers (Lehmann and Wyss, 2001; Wijma and Janssen, 2013; Damborsky and Brezovsky, 2014; Basu and Sen, 2013). A typical sequence/structure based design to enhance thermostability assumes that the conserved amino acids observed in related sequences are singles which mainly contributes favoring stability. Thus, the engineering strategy in the former case consists of output from multiple alignments of similar sequences followed by implementing suitable design strategies in the target protein that resemble the consensus sequence. Otherwise from a structure-based method it can be summarized that a more rigid protein would be more stable at high temperatures (Watanabe *et al.*, 2006; Razvi and Scholtz, 2006). The different design strategies introduced to the protein structure are increasing the disulfide bonds, salt bridges and stabilizing the loop regions either by introducing residues such as proline or removing glycine. Similarly, another common strategy is finding the most flexible region of the target proteins from the experimental B-factors and then to focus mutagenesis on this aspect (Farias and Bonato, 2003; Kumar *et al.*, 2000; Wang and Lercher, 2010; Anbarasan *et al.*, 2010). One of the direct method in case of proteins of interest having enhanced thermal stability may be obtained either by protein engineering or by searching for homologs in thermophiles (Wijma *et al.*, 2013; Radestock and Gohlke, 2011). Protein engineering may involve directed evolution, design from first principles or some combination of these two strategies (Chakraborty *et al.*, 2013; Barrozo *et al.*, 2012; Lindberg *et al.*, 2013). Protein design strategy can be viewed just like an optimization problem: That considers certain criteria, by which an engineered protein will fold to the desired structure and that holds the novel function (Alvizo *et al.*, 2007). The most common design mechanism to achieve the thermo stability is engineering a desired protein by inserting disulfide bonds, substituting core amino acids, *de novo* design of the hydrophobic pocket, salt-bridge insertion etc. (Chirakkal *et al.*, 2001; Yokota *et al.*, 2006; Zhou *et al.*, 1993; Charbonneau and Beauregard, 2013). Following computational design it is necessary to evaluate the substrate binding affinity by the molecular docking method. For docking purpose Autodock Vina is one of the most preferred tools as described by many researchers world-wide (Azam and Abbasi, 2013; Chen and Ren, 2014; Frączek *et al.*, 2013). Furthermore the molecular dynamics simulation is used commonly to check the stability, energy and also to evaluate thermostability property of the enzyme at different temperatures (Paul *et al.*, 2014; Manjunath and Sekar, 2013; Noorbacha *et al.*, 2013).

The objective of the present work is to develop a computational protein design method using the crystal structure of haloalkane dehalogenase enzyme as a template to make it thermostable. Subsequently the design and validation of the designed protein was studied by various *in silico* approaches.

## Materials and Methods

### *System Architecture used in this Work*

The configuration of the PC used for performing computation studies were as follows; OS Windows XP, Dual Core, RAM 2 GB, 250 GB HD, 1.76 GZ and for docking and MD simulation HPC WIPRO Super Nova, 1 TF, Intel Xeon Quad Core 2.66 GHZ, with 16 Nodes were used.

### *Selection of Suitable Haloalkane Dehalogenase Enzyme Structure from PDB and In-Silico Protein Design*

Currently the PDB contains about 59 haloalkane dehalogenase 3D structures. The most suitable dehalogenase was searched from PDB server using keywords and Boolean criteria in the advanced search (<http://www.rcsb.org/pdb/search/advSearch.do>) such as “haloalkane dehalogenase”, chain length from 200-350 amino acid, single chain, CATH criteria having alpha-beta with no ligands and with a low crystallographic R value.

### *Disulfide Bond Design*

Disulfide bond design is an important biotechnological tool that has great interest in wide areas of research and the same has been applied extensively to improve protein stability, modify functional characteristics. DbD2 web server available at <http://cptweb.cpt.wayne.edu/DbD2/> is used for disulfide bond design. DbD2 is a web-based, platform-independent application that significantly extends functionality and visualization capacity. This feature facilitates identification of potential disulfides that are not only likely to form, but are also expected to provide improved thermal stability to the protein that are very reliable. Following the desired PDB file upload, suitable parameters were selected for our desired PDB (details of parameters discussed in the results and discussion section).

### *Hydrophobic Pocket Design*

Binding sites and active sites of proteins are most often associated with structural pockets and cavities. CastPserver (<http://stsfw.bioengr.uic.edu/castp/calculation.php>) was used for identification and measurements of surface

accessible pockets as well as interior inaccessible cavities, for proteins. This program analytically measures the area and volume of each pocket and cavity that can be visualized by the Jmol visualization tool. Castp predicted several pockets for the protein of interest, among them the suitable pocket was chosen based on criteria such as, area >100, cavity not containing any active site residues and has maximum number of hydrophilic residues. Then the selected hydrophilic residues were mutated to corresponding hydrophobic residues based on comparable molecular weight using PyMol tool.

### *Salt Bridge Mutant Construction*

The criteria for salt bridge formation is that at least one Asp or Glu side-chain carboxyl oxygen atom (containing atom type OD in Asp or OE in Glu) and one side-chain nitrogen atom of Arg, Lys or His (contain atom type NH in Arg, NZ in Lys or NE and ND in His) should be within a distance of 4.0 Angstroms. The salt bridge in the selected protein was analyzed by ESBRI web tool (<http://bioinformatica.isa.cnr.it/ESBRI/>). Other charged residue that does not take part in salt bridge formation in the core and surface were identified using the in-house built software. Other than the existing salt bridges, additional salt bridges were designed by suitable mutation at a specific range and the respective pairs were analyzed whether they are still within the range (about 4.0 angstrom) after the energy minimization.

### *Multiple Mutant Enzyme Design*

All the above mutation strategies mentioned in the earlier sections were introduced in combination to the protein structure template using PyMol tool.

### *Energy Minimization and Molecular Docking to Check the Substrate Affinity*

Following the design of the mutant structures it is necessary to analyze the different energy values obtained before and after minimization to gain an idea about their stability. For the various mutants, preliminary energy minimization was performed using SPD viewer tool followed by GROMACS software. Further molecular docking was performed on the designed protein, including the original (wild type). For docking purpose the chemical structures of substrates for haloalkane dehalogenase enzymes were obtained from the PubChem database (<http://pubchem.ncbi.nlm.nih.gov/>). Two substrate structures 1, 2, 3-trichloropropane and 1, 2-dichloroethane were retrieved based on the toxicity level that are also degraded by the haloalkane dehalogenase enzyme. These compounds were energy

minimized to respective 3D structure by PRODRG server (<http://davapc1.bioch.dundee.ac.uk/prodrg/>) followed by molecular docking using Autodock Vina tool.

### *Comparative Molecular Dynamics Simulation*

Molecular dynamics simulation was run to examine the stability of different designed proteins. This is one of the popular computational methods that calculates the time dependent behavior of protein and provides detailed information in terms of fluctuations and conformational changes. For MD simulation, GROMACS 4.5.0 (Groningen Machine for Advanced Chemical Simulations) software was used. It is a freely available and open source molecular dynamics simulation software mainly designed for simulations of biochemical molecules like proteins, lipids and nucleic acids. The force field used in this study was GROMOS96 53a6. The protocol for simulation includes placing the initial protein structure in a cubic box filled with water, followed by energy minimization of the system without restraint using the steepest descent method for maximum up to 300 steps. The molecular dynamics simulation was set at temperature 298 K, with decreasing harmonic position restraints on the protein atoms over a 50 pico second time interval, followed by 8 ns (up to 333 K) without restraints in the simulated annealing mode. The P-LINCS algorithm was used with a dielectric constant of 1 and a time step of 2 Femto second (fs) was used for the MD simulation purpose with pressure kept constant during the MD simulations. A simulated annealing protocol was performed by dynamically changing the reference temperature (from 298 to 333 K) in the thermostat algorithm. After selection of suitable mutants that achieved thermo stability, another round of molecular dynamics simulation was performed for 5 nano second at temperatures 400, 500 and 600 K respectively.

## **Results and Discussion**

Crystallographic structures of haloalkane dehalogenase from the source organism *Xanthobacter autotrophicus* GJ10 (PDB entry 1EDE) was used as a template for the initial coordinates for MD simulations (Verschueren *et al.*, 1993). This protein structure has single chain of 310 residues without any ligand whose structure was solved at a relatively low crystallographic R value (Angstrom) of 1.9. For all mutation purpose, the structural conformation of this protein was used as template.

### *Disulfide Design in the Structure by DbD Tool*

For calculating the possible amino acid pairs that can be mutated to form disulfide bonds by DbD server, intra

chain interaction was selected and  $\chi_3$  angle within range  $-87^\circ$  or  $+97^\circ \pm 10$  with  $C\alpha-C\beta-S\gamma$  angle range  $114.6^\circ \pm 8$  was considered. From the output a total of 12 pairs of predicted disulfide bond formers were obtained, out of which only 3 residue pairs (ASP 48 -ALA 75, VAL 69 -ALA 299 and TYR 38 -PRO 80) were chosen for mutation to form disulfide bonds due to its low energy (Kcal/Mol) value. Altogether the above mentioned residues were then mutated to cysteine using the Pymol mutagenesis module.

### Hydrophobic Pocket Design

From the output of the CastP server, only pocket ID #49 fulfils the criteria stated in materials and methods section and therefore selected. From the predicted pocket, residue ASP 9 and both PHE 12 and TYR 24 were not considered for mutagenesis as the former is present in an alpha helix and the later residue is hydrophobic in nature. All other residues of pocket #49 were converted into corresponding hydrophobic residues based on comparable molecular weight using PyMol mutagenesis module. The mutant protein in this type contains replacement of residue 6ARG/PHE, 7THR/VAL, 35 ARG/PHE, 87 LYS/ILE, 88 SER/VAL, 90 LYS/ILE as shown in Fig. 1.

### Result of Salt bridge Construction

The various charged amino acid residues within  $4\text{\AA}$  radius were selectively replaced by other appropriately charged residues (Table 1).

### Result of Multiple Mutant Construction

Multiple mutant protein contains all the changes that have been performed for disulfide bond design,

hydrophobic core design as well as the mutation that is done for salt bridge construction as explained above.

### Energy Minimization and Molecular Docking Study

The results indicate that all these mutants along with the original (wild type) protein possess a total energy that is negative, hence stable (Table 2). Followed by a docking protocol, the results show that the binding site of the ligand was found to be explicitly positioned near the crystallographic binding orientation within the protein cavity. From binding energy values in all cases it was concluded that in spite of mutation the mutant enzymes show almost equal-affinity to the substrates (Table 2).

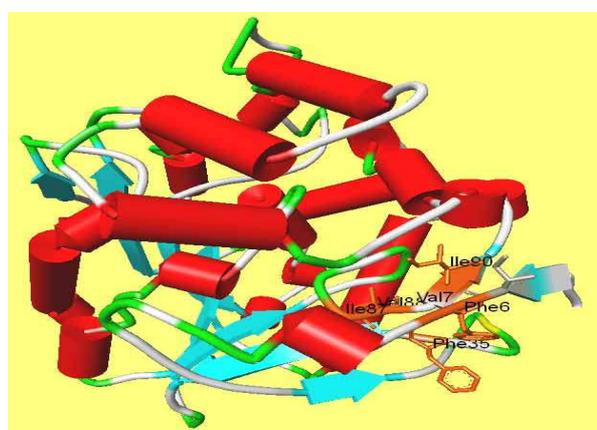


Fig. 1. Showing hydrophobic mutant of 1EDE contains specific mutant amino acids as described above

Table 1. Final selected residue for salt bridge construction followed by energy minimization

Protein region	Amino acid pair within $4\text{\AA}$ distance	Substituted amino acid	Distance ( $\text{\AA}$ )	
			Before minimization	After minimization
Surface	MET 258- GLU 283	LYS 258	2.4	3.3
Core	THR 132- ARG 103	ASP-132	4.9	4.0
Core	PHE 82- HIS102	ASP-82	4.0	4.0

Table 2. Energy value before and after the design and the binding energy with the substrates after docking

Protein Type	Energy before Minimization <sup>a</sup>	Energy after Minimization <sup>a</sup>	Binding energy for ligands (kcal/mol) <sup>b</sup>	
			1,2,3-trichloropropane	1,2-dichloroethane
Original	-18631.602	-19123.094	-3.0	-2.5
Disulfide mutant	-9211.067	-15828.719	-3.0	-2.8
Hydrophobic mutant	990231.750	-13642.758	-3.4	-3.0
Salt-bridge mutant	-14212.729	-17243.926	-3.2	-3.3
Multiple mutant	-13259.147	-16317.693	-3.2	-3.3

<sup>a</sup>Energy value (KJ/mol) was calculated by Spdviewer version 4.0.1 software by using GROMOS 96 force field. <sup>b</sup>Computed after docking of protein and ligands by Autodock Vina tool

### *Molecular Dynamics Simulation Results*

The molecular dynamics simulation was by placing the protein in a defined grid solvated in a water environment. The simulation was performed by the simulated annealing method for 8000 Pico second at different temperature levels ranging 298 to 333 K as discussed above. Various properties such as RMSD, Solvent-Accessible Surface Area (SASA) both hydrophobic and hydrophilic, Radius of Gyration (RG), Root Mean Square Fluctuation (RMSF), number of hydrogen bonds both in protein and protein-solvent (water) system were computed as a function of time to assess the stability of the proteins from the respective trajectory file of Gromacs output.

#### *Low Root Mean Square Deviation (RMSD) of Hydrophobic Mutant Enzyme*

RMSD is a parameter that shows how much a structure differs from its original protein structure and therefore it is considered as an important indicator for evaluating the thermostability of a protein in terms of structural conformation. In general, the smaller the RMSD value of a protein, smaller is its  $C\alpha$ -atomic displacement range and therefore more thermo stable is its conformation (Zhang *et al.*, 2014). The output was obtained as the  $C\alpha$ -atomic displacement parameter to access the stability. The RMSD during the simulations shows the average over the 8 ns simulations (Fig. 2). The overall C-alpha to C-alpha fluctuation obtained for the hydrophobic mutant is comparatively less. At the beginning of the trajectory ( $t = 0$ ), RMSD of all proteins displayed a value of 0.09 nm indicating further about the movements occurred during the thermalization and equilibration periods. Between 7000-8000 pico seconds (298-333 K), the deviation value of hydrophobic mutant decreases sharply than other mutant protein.

However, it is possible that at simulation times of 8ns, the hydrophobic mutant structure might be stable and reach at the folded state at a variable temperature. This is also due the substituted hydrophobic amino acids are responsible for enhancing the thermal stability of protein structure in this mutant of the haloalkane dehalogenase enzyme.

#### *Low Solvent Accessible Surface Area (SASA) of Hydrophobic Mutant*

Solvent Accessible Surface Area (SASA) is basically used to calculate the transfer of free energy that is required to move a protein from a polar solvent to a non-polar solvent and is also used to predict the secondary structure (Momen-Roknabadi *et al.*, 2008). Figure 3 and 4 show a significant decrease in total

solvent accessible surface value (for both hydrophobic and hydrophilic) in case of hydrophobic mutants until the end of simulation time. The solvent accessible surface areas of atoms are correlated with their hydrophobicity value and folding process in proteins.

A lower SASA value correlated proportionally to the stability in the protein structure has been explained earlier (Richmond, 1984). The replacement of hydrophilic residues with hydrophobic residues has increased the non-polar residues located inside the protein structure favouring protein packing and thus reducing the internal cavities within protein that providing overall protein thermostability (Salleh *et al.*, 2012).

#### *Distribution Salt Bridge in Mutant Enzyme Structure during MD Simulation*

Figure 5 indicates the trends of salt bridge formation and breakage in case of all considered mutants as well as the original protein. In general, surface salt bridges have been shown to contribute to protein stability at high temperatures due to decreased desolvation penalties of charged groups (Fei *et al.*, 2013).

A significant increase in the number of salt-bridges has been reported for most structures of thermostable enzymes, but this trend has not has been explained properly by site-directed mutagenesis experiments (Yip *et al.*, 1998; Trejo *et al.*, 2001). However in this case, at room temperature (0 ns/ 298 K) the number of salt bridges is more but as the temperature increases there is anomalous effect observed in all enzymes.

#### *Lower Root Mean Square Fluctuation (RMSF) Value of Hydrophobic Mutant*

Root Mean Square Fluctuation (RMSF) of each residue in the proteins was used to calculate the trajectory to determine the overall flexibility and behavior of the system at different temperatures (298-333 K). The conformational structure of haloalkane dehalogenase mutant structures continuously change during the MD simulations until at the end of the simulation (8 ns).

In this comparative study lower conformational changes were obtained for hydrophobic mutant (Fig. 6). However, due to lack of significant fluctuations (nm), the RMSF does not correlate much to the thermal stability for a specific mutant but indicate better compactness and high rigidity which are the major factors in protein thermal stability in all considered proteins (Kuzmanic and Zagovic, 2010).

#### *Lower Radius of Gyration of Hydrophobic Mutant*

The radius of gyration describes the general size of the protein and basically indicates the global changes that occur in the shape of the protein during MD simulation. The radii of gyration during the

simulations are shown in Fig. 3 and the average values over the 8 ns production simulations are shown in Fig. 7. It can be seen that most of the simulations give average radii of gyration of 1.78-1.86 nm. However among all selected proteins the hydrophobic mutant show slightly lesser radii of 1.79 to 1.82 nm between temperature range 298K/28°C and 333K/60°C. This indicates that, in general the mutations do not have any major impact on the general shape of the protein, except some rather small changes when the hydrophobic residues are inserted to the original

protein structure, which might have indirectly improved the hydrophobic interaction in the protein structure leading to the rigidity in the structure.

The radius of gyration as a feature for selecting thermo stability has been studied by many researchers (Irani *et al.*, 2013; Feng *et al.*, 2007; Liu *et al.*, 2008). It has been observed that the thermal stability of the proteins is inversely proportional to the radius of gyration. From our analysis hydrophobic mutant show comparatively lower value for radius of gyration indicating thermo stability.

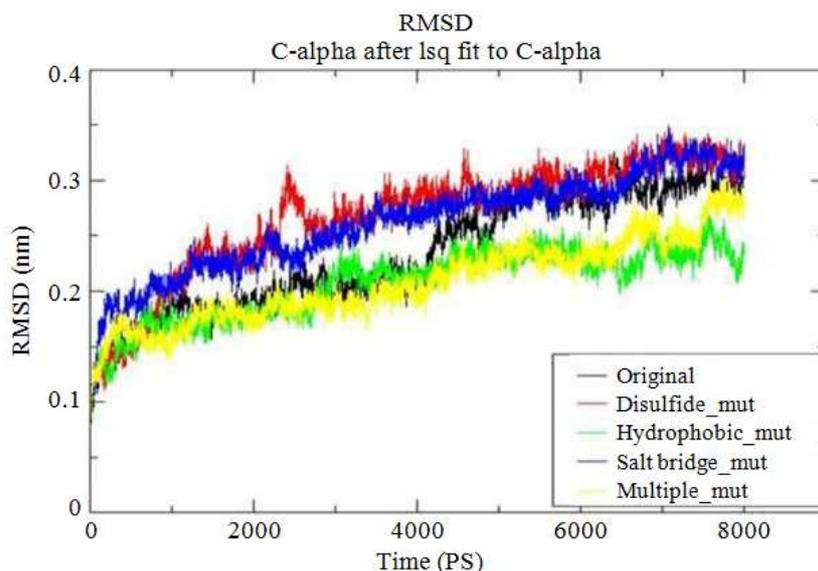


Fig. 2. Root Mean Square Deviation (RMSD) of the  $C\alpha$  position of mutant and original protein backbone in (nm) from 298K/25°C-333K/60°C as obtained from the period of 0-8 ns

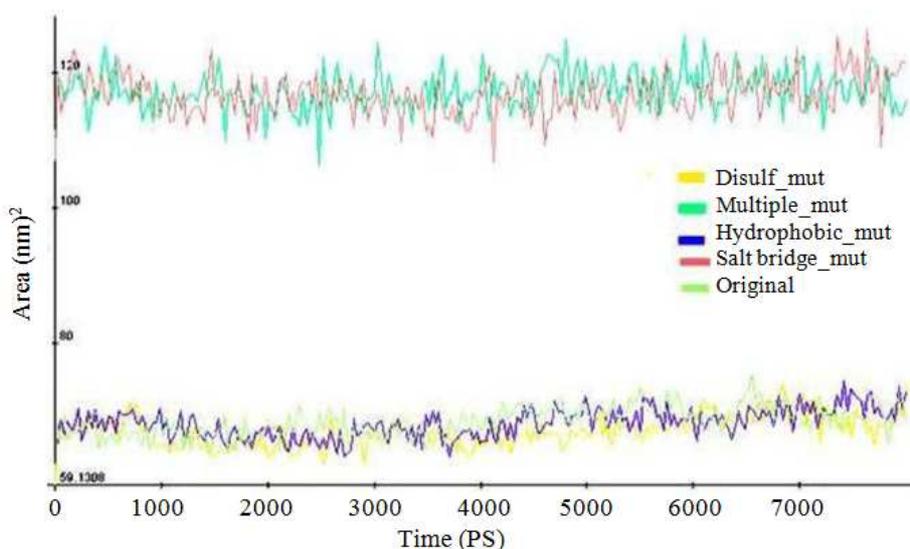


Fig. 3. Change in hydrophobic surface area during 8ns simulation

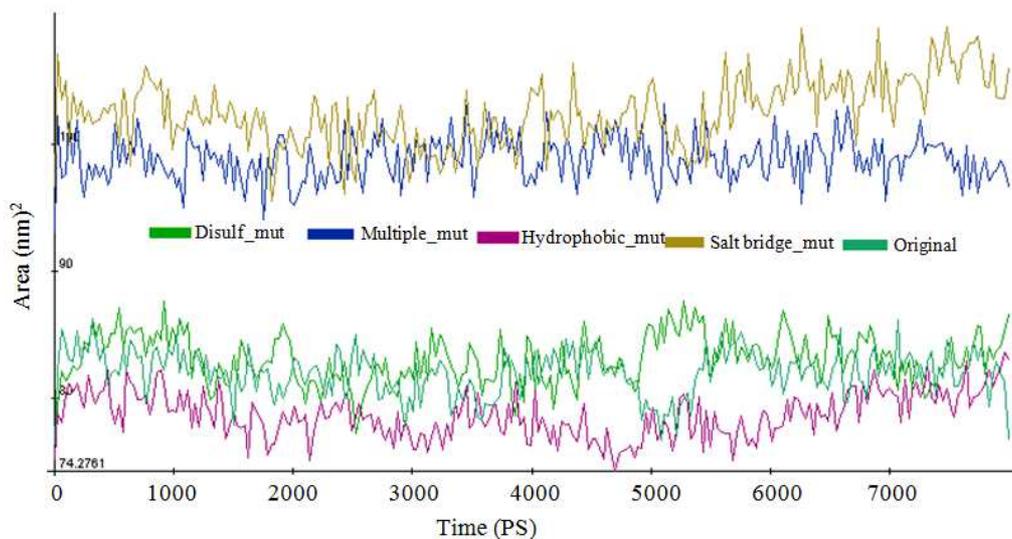


Fig. 4. Change in hydrophilic surface area during 8ns simulation

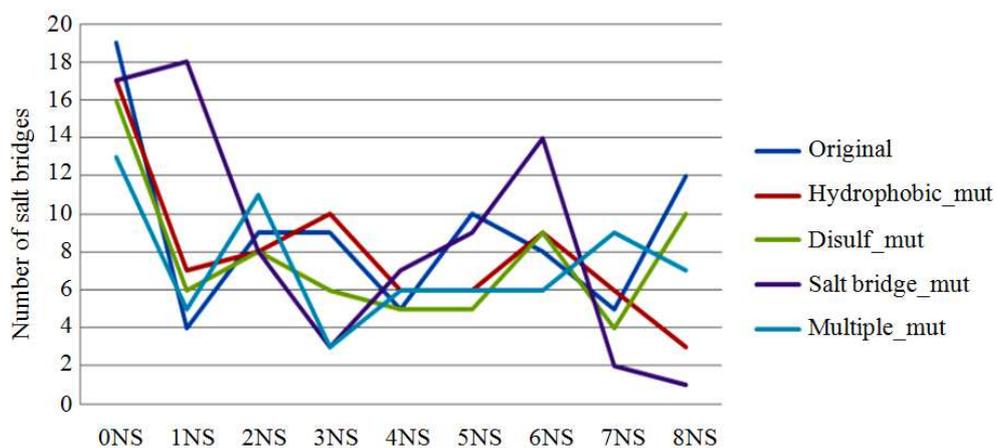


Fig. 5. Trends in the salt bridge construction and breakage during the simulation

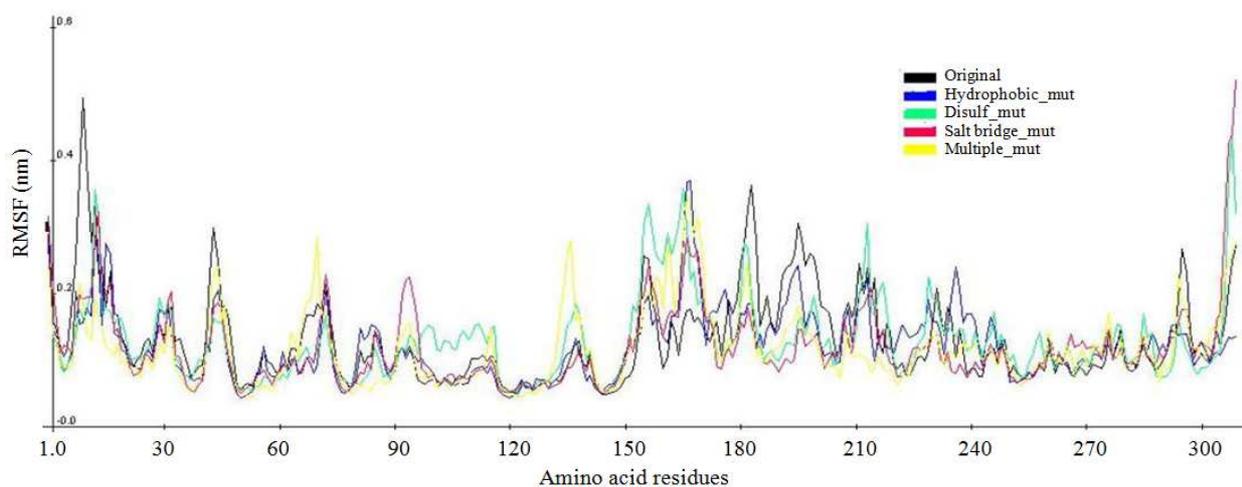


Fig. 6. Root Mean Square Fluctuations (RMSF) of the Ca position of lipase backbone versus residue in different temperature from 298/25°C and 333K/60°C during 8 ns simulations

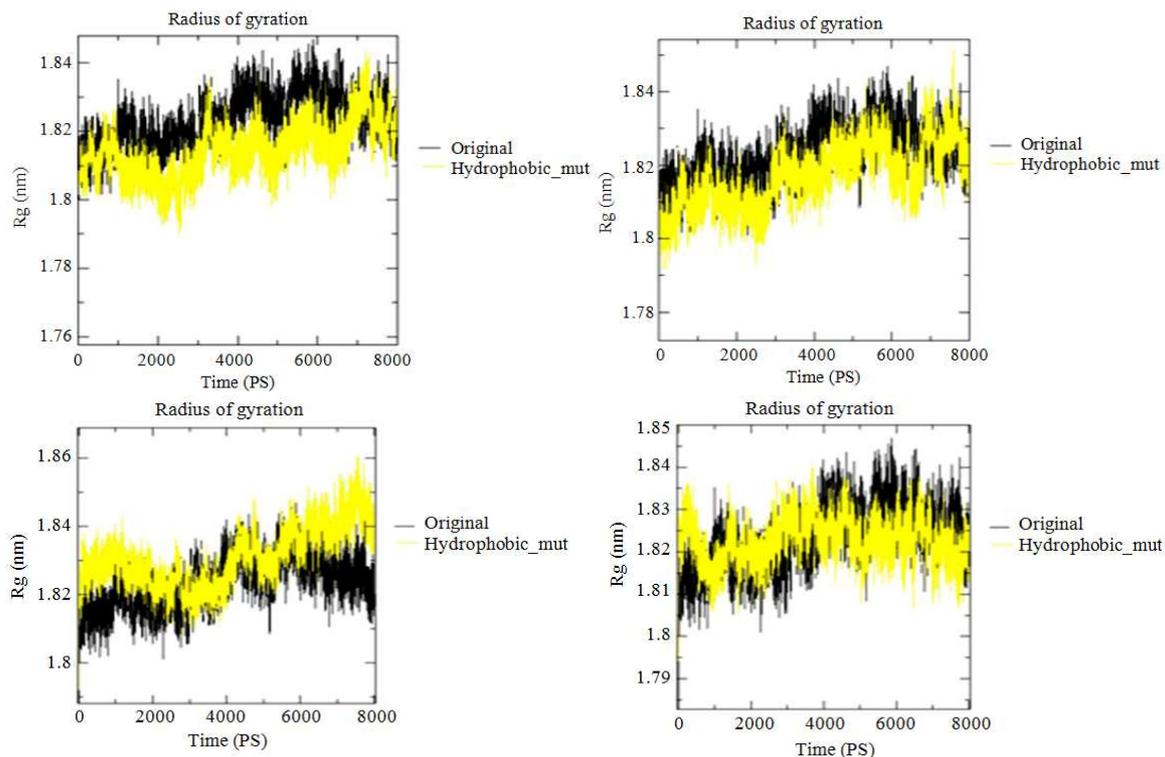


Fig. 7. Showing radius gyrations for original and other mutant protein

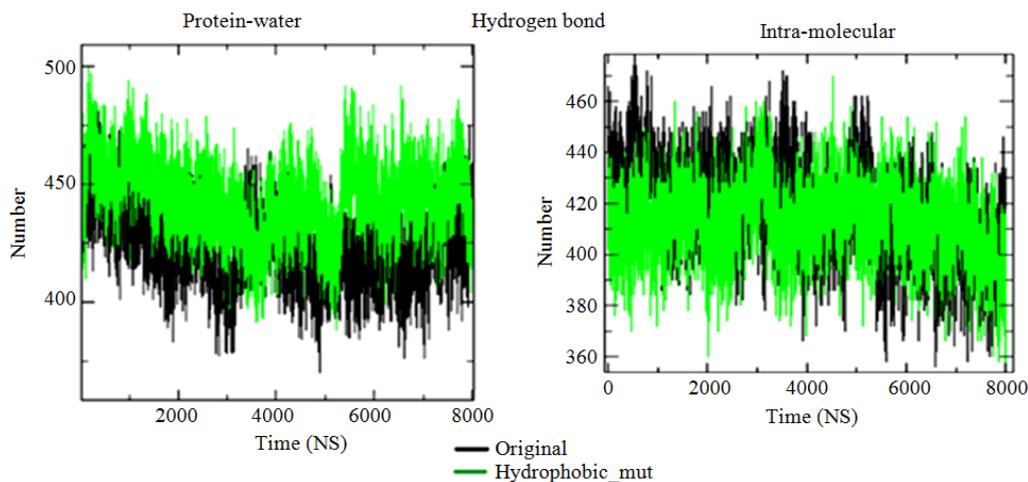


Fig. 8. Hydrogen bond formation and breakage during 8 ns simulation in protein-water and intra-molecular system

### Optimum Hydrogen bond Distribution in Hydrophobic Mutant

Hydrogen bond calculation from the trajectory file of Gromacs was computed at a distance cut-off of 3.5 Angstrom. As the temperature increases with the increase in time, the numbers of intra-molecular hydrogen bonds were observed to be slightly decreased in case of hydrophobic mutant in comparison to the un-mutated (Original) protein. This

is reasonable because the structure becomes more distorted as the simulation temperature is raised. It was also observed from the plot (Fig. 8) that, although the number of hydrogen bonds varies in different temperature still it is steadily maintained with the increase in simulation time and temperature (Choudhury *et al.*, 2010; Vogt *et al.*, 1997). Due to a large distortion of regular secondary structural elements and unpacking of the hydrophobic cores, some of water molecules are inserted in hydrophobic cores.

It can be predicted that, the enhanced intermolecular (water-protein) hydrogen bonding might compensate the breakage of intra-molecular Hydrogen bonding thereby providing the stability to the hydrophobic mutant.

### Dynamic Behavior of the Hydrophobic Amino Acids in Hydrophobic Mutant

In addition to all the features, the dynamic behavior of designed hydrophobic pockets was also analyzed and shown in Fig. 9. The amino acid residue PHE 6 with VAL 88, PHE 35 and ILE 90 was observed to interact periodically as the time and temperature increased. The hydrophobic interactions have more stabilizing effect in comparison to the salt bridges (Burg *et al.*, 1994; Karjiban *et al.*, 2009; Shirdel *et al.*, 2013).

For each nano second of simulation the non-polar SASA value was computed from Naccess program (<http://www.bioinf.manchester.ac.uk/naccess/>) to check the behaviour of the considered residues. Figure 10

indicates Val 88 and Ile 90 remaining almost buried in nature and Ile is little exposed (at 5 ns) during the simulation time with temperature increment. However in this condition, due to the exposure of the residues Phe 6 and Phe 35, it strongly promotes the flexibility, enhancing their interaction and burial of adjacent residues strongly inhibiting the flexibility of the central residues (interaction of Val 7 and Ile 87).

The present analysis has shown a new concept of enhancing haloalkane dehalogenase protein thermo stability by hydrophobic amino acid interaction at the surface. So along with the surface exposed hydrophilic residues and core hydrophobic residues, the surface hydrophobic residues in the loop region also play a significant role in attaining thermo stability (Zhang *et al.*, 2009; Moret and Zebende, 2007). This factor can be considered as a new protein design strategy for this particular enzyme to achieve thermo stability.

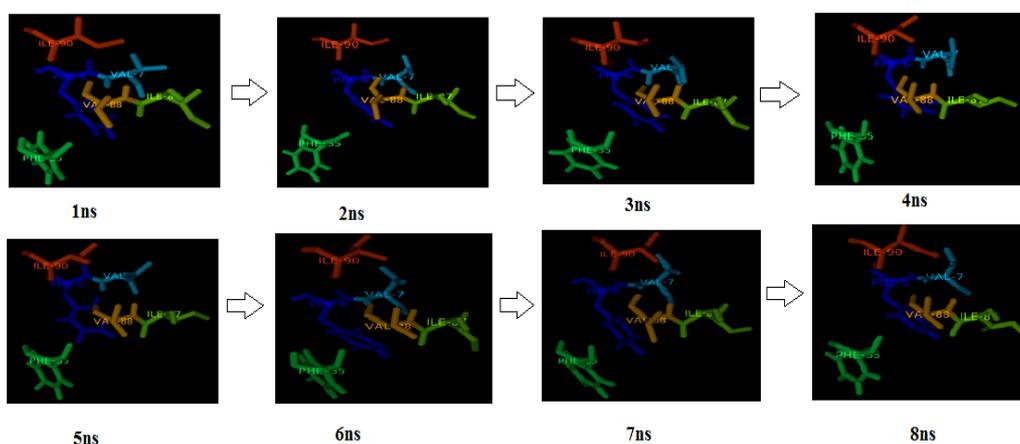


Fig. 9. Dynamic behavior of the designed hydrophobic pocket (interaction of Phe 6, Phe 35 and Ile 90 represented in blue, green and deep brown colour respectively)

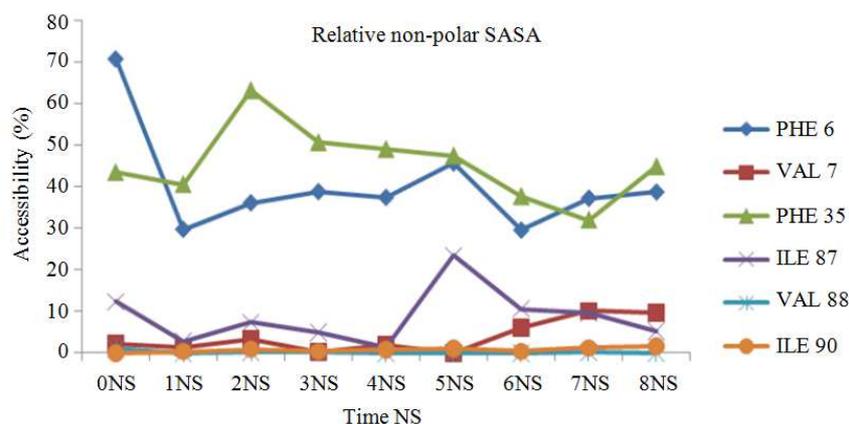


Fig. 10. Trends in relative non-polar SASA computed from NACCESS program for the mutated residues in the hydrophobic pocket

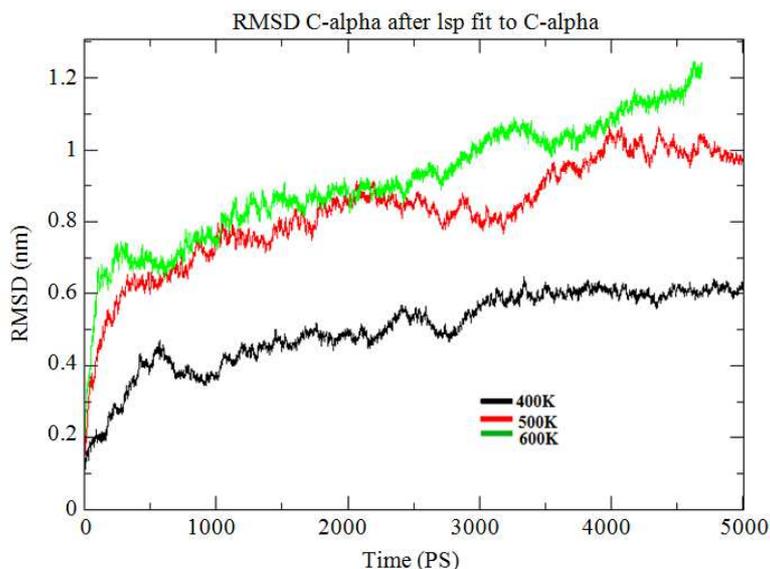


Fig. 11. RMSD calculation at higher temperature MD simulation of the protein that contains mutation in the hydrophobic pocket

### *MD Simulation of Hydrophobic Mutant at high Temperature Conditions*

In order to validate its thermal stability, further simulations at (constant) high temperature is necessary (Chang and Mahoney, 1995). The plots of RMSD of the protein versus time at different temperatures are shown in Fig. 11. The plot shows that at 400 and 500 K the enzyme was found to be stable throughout 5ns simulation time. In the trajectory run at 600 K, the starting conformation changed and the backbone RMSD slightly increased fluctuating between 0.2 and 1.5 nm.

The rise in RMSD indicates large changes in protein structure and some disruption in the tertiary structure of protein. Due to the thermostable property, the hydrophobic mutant can stay longer time in the folded state at higher temperature.

### **Conclusion**

In the current work a computational effort has been made to improve the thermostability of a haloalkane dehalogenase enzyme, thereby understanding the thermostability mechanism of the enzymes. After a rule-based mutant protein design, energy minimization and docking showed all the mutants to be stable not losing their substrate binding affinity with selected ligands. During MD simulation, various properties such as the average values of the Solvent-Accessible Surface Area (SASA) for both hydrophilic and hydrophobic, the Radius of Gyration (RG), the number of inter and intra-protein hydrogen bonds and the number of salt bridges were computed as a function of time. On the basis of the simulation result, it was observed that the hydrophobic pocket mutant enzyme that contains mutation of amino

acids 6ARG/PHE, 7THR/VAL, 35 ARG/PHE, 87 LYS/ILE, 88 SER/VAL, 90 LYS/ILE in the loop region show thermo stability. The mechanism of thermostability of hydrophobic mutant was studied that was mainly attributed due to the interaction of 6PHE with 35 PHE and 87 ILE in the surface. Furthermore, at high temperature simulation (at 400 and 500 K), the hydrophobic mutant shows low RMSD confirming the thermo stability nature. Surface hydrophobicity is linked to the thermo stability in the haloalkane dehalogenase enzyme has been established in this study. However this is a computational interpretation, therefore further standardized experimental methods as site directed mutagenesis and enzyme activity analysis is required to establish this concept. The whole study offers a general characterization of the computational approaches to design biotechnologically improved thermostable enzymes.

### **Acknowledgement**

We are thankful to chief executive of MITS group for providing MIRC laboratory for performing the Molecular dynamics simulation.

### **Funding Information**

No funding was available from any sources for performing the work

### **Author's Contributions**

**Satpathy Raghunath:** Designed the problem, performed Molecular dynamics simulation and docking experiments and writing the manuscript.

**V. Badireenath Konkimalla:** Developed an in house based software for SASA calculation of protein structure, Linking the concept of hydrophobicity to thermo-stability.

**Ratha Jagnyeswar:** Analysis of the result and overall hypothesis prediction, overall editing in the manuscript.

## Ethics

This manuscript is an original research article and contains unpublished material. The corresponding author declares that, the manuscript has been approved by all the three authors and there are no ethical issues involved.

## References

- Anbarasan, S., J. Jänis, M. Paloheimo, M. Laitaoja and M. Vuolanto *et al.*, 2010. Effect of glycosylation and additional domains on the thermostability of a family 10 xylanase produced by *Thermopolyspora flexuosa*. *Applied Environ. Microbiol.*, 76: 356-360. DOI: 10.1128/AEM.00357-09
- Azam, S.S. and S.W. Abbasi, 2013. Molecular docking studies for the identification of novel melatoninergic inhibitors for acetylserotonin-O-methyltransferase using different docking routines. *Theor. Biol. Med. Modell.*, 10: 63-63. DOI: 10.1186/1742-4682-10-63
- Alvizo, O., B.D. Allen and S.L. Mayo, 2007. Computational protein design promises to revolutionize protein engineering. *Biotechniques*, 42: 31-35. PMID: 17269482
- Bidmanova, S., J. Damborsky and Z. Prokop, 2013. Immobilization of Haloalkane dehalogenase LinB from *Sphingobium japonicum* UT26 for Biotechnological Applications. *Biocatal. Biotransform.* DOI: 10.4172/2324-9099.1000106
- Barrozo, A., R. Borstnar, G. Marloie and S.C.L. Kamerlin, 2012. Computational protein engineering: bridging the gap between rational design and laboratory evolution. *Int. J. Mol. Sci.*, 13: 12428-12460. DOI: 10.3390/ijms131012428
- Basu, S. and S. Sen, 2013. An *in silico* method for designing thermostable variant of a dimeric mesophilic protein based on its 3D structure. *J. Mol. Graphics Modell.*, 42: 92-103. DOI: 10.1016/j.jmgm.2013.02.011
- Burg, B., B.W. Dijkstra, G. Vriend, B. Vinne and G. Venema *et al.*, 1994. Protein stabilization by hydrophobic interactions at the surface. *Eur. J. Biochem.*, 220: 981-985. PMID: 8143751
- Chakraborty, S., R. Minda, L. Salaye, A.M. Dandekar and S.K. Bhattacharjee *et al.*, 2013. Promiscuity-based enzyme selection for rational directed evolution experiments. *Methods Mol. Biol.*, 978: 205-216. DOI: 10.1007/978-1-62703-293-3\_15
- Charbonneau, D.M. and M. Beauregard, 2013. Role of key salt bridges in thermostability of *G. thermodenitrificans* EstGtA2: Distinctive patterns within the new bacterial lipolytic enzyme family XV. *PLoS One*. DOI: 10.1371/journal.pone.0076675
- Choudhury, D., S. Biswas, S. Roy and J.K. Dattagupta, 2010. Improving thermostability of papain through structure-based protein engineering. *Protein Eng. Des. Sel.*, 23: 457-467. DOI: 10.1093/protein/gzq016
- Chirakkal, H., G.C. Ford and A. Moir, 2001. Analysis of a conserved hydrophobic pocket important for the thermostability of *Bacillus pumilus* Chloramphenicol Acetyltransferase (CAT-86). *Protein Eng.*, 14: 161-166. PMID: 113427120
- Chang, B.S. and R.R. Mahoney, 1995. Enzyme thermostabilization by bovine serum albumin and other proteins: Evidence for hydrophobic interactions. *Biotechnol. Applied Biochem.*, 22: 203-214. PMID: 7576258
- Chen, S.J. and J.L. Ren, 2014. Identification of a potential anticancer target of danshensu by inverse docking. *Asian Pac. J. Cancer Prev.*, 15: 111-116. PMID: 24528010
- Dijkstra, B.B.W., 1993. Crystallographic analysis of the catalytic mechanism of haloalkane dehalogenase. *Nature*, 363: 693-698. PMID: 8515812
- Damborsky, J. and J. Brezovsky, 2014. Computational tools for designing and engineering enzymes. *Curr. Opin. Chem. Biol.*, 19: 8-16. DOI: 10.1016/j.cbpa.2013.12.003
- Fetzner, S. and F. Lingens, 1994. Bacterial dehalogenases: Biochemistry, genetics and biotechnological applications. *Microbiol. Rev.*, 58: 641-685. PMID: 7854251
- Fei, B., H. Xu, F. Zhang, X. Li and S. Ma *et al.*, 2013. Relationship between *Escherichia coli* AppA phytase's thermostability and salt bridges. *J. Biosci. Bioeng.*, 115: 623-627. DOI: 10.1016/j.jbiosc.2012.12.010
- Feng, J.A., L.A. Tessler and G.R. Marshall, 2007. Chimeric protein engineering. *Int. J. Peptide Res. Therapeut.*, 13: 151-160. DOI: 10.1007/s10989-006-9058-8
- Farias, S.T. and M.C. Bonato, 2003. Preferred amino acids and thermostability. *Genet. Mol. Res.*, 2: 383-393. PMID: 15011142
- Frączek, T., A. Siwek and P. Paneth, 2013. Assessing molecular docking tools for relative biological activity prediction: A case study of triazole HIV-1 NNRTIs. *J. Chem. Inf. Model.*, 53: 3326-3342. DOI: 10.1021/ci400427a

- Hasan, K., A. Fortova, T. Koudelakova, R. Chaloupkova and M. Ishitsuka *et al.*, 2011. Biochemical characteristics of the novel haloalkane dehalogenase DatA, isolated from the plant pathogen *Agrobacterium tumefaciens* C58. *Applied Environ. Microbiol.*, 77: 1881-1884. DOI: 10.1128/AEM.02109-10
- Hesseler, M., X. Bogdanović, A. Hidalgo, J. Berenguer and G.J. Palm *et al.*, 2011. Cloning, functional expression, biochemical characterization and structural analysis of a haloalkane dehalogenase from *Plesiocystis pacifica* SIR-1. *Applied Microbiol. Biotechnol.*, 91: 1049-1060. DOI: 10.1007/s00253-011-3328-x
- Irani, M., U. Törnvall, S. Genheden, M.W. Larsen and R. Hatti-Kaul *et al.*, 2013. Amino acid oxidation of *Candida antarctica* lipase B studied by molecular dynamics simulations and site-directed mutagenesis. *Biochemistry*, 52: 1280-1289. DOI: 10.1021/bi301298m
- Koudelakova, T., S. Bidmanova, P. Dvorak, A. Pavelka and R. Chaloupkova *et al.*, 2013. Haloalkane dehalogenases: Biotechnological applications. *Biotechnol. J.*, 8: 32-45. DOI: 10.1002/biot.201100486
- Karjiban, R.A., M.B.A. Rahman, M. Basri, A.B. Salleh and D. Jacobs *et al.*, 2009. Molecular dynamics study of the structure, flexibility and dynamics of thermostable L1 lipase at high temperatures. *Protein J.*, 28: 14-23. DOI: 10.1007/s10930-008-9159-7
- Kuzmanic, A. and B. Zagrovic, 2010. Determination of ensemble-average pairwise root mean-square deviation from experimental B-factors. *Biophys. J.*, 98: 861-871. DOI: 10.1016/j.bpj.2009.11.011
- Kumar, S., C.J. Tsai and R. Nussinov, 2000. Factors enhancing protein thermostability. *Protein Eng.*, 13: 179-191. PMID: 10775659
- Liu, J., H. Yu and Z. Shen, 2008. Insights into thermal stability of thermophilic nitrile hydratases by molecular dynamics simulation. *J. Mol. Graph. Modell.*, 27: 529-535. DOI: 10.1016/j.jmgm.2008.09.004
- Lehmann, M. and M. Wyss, 2001. Engineering proteins for thermostability: The use of sequence alignments versus rational design and directed evolution. *Curr. Opin. Biotechnol.*, 12: 371-375. PMID: 11551465
- Lindberg, H., A. Johansson, T. Härd, S. Ståhl and J. Löfblom, 2013. Staphylococcal display for combinatorial protein engineering of a head-to-tail affibody dimer binding the Alzheimer amyloid- $\beta$  peptide. *Biotechnol. J.*, 8: 139-145. DOI: 10.1002/biot.201200228
- Momen-Roknabadi, A., M. Sadeghi, H. Pezeshk and S.A. Marashi, 2008. Impact of residue accessible surface area on the prediction of protein secondary structures. *BMC Bioinform.*, 9: 357-357. DOI: 10.1186/1471-2105-9-357
- Marek, J., J. Vévodová, I.K. Smatanová, Y. Nagata and L.A. Svensson *et al.*, 2000. Crystal structure of the haloalkane dehalogenase from *Sphingomonas paucimobilis* UT26. *Biochemistry*, 39: 14082-14086. PMID: 11087355
- Manjunath, K. and K. Sekar, 2013. Molecular dynamics perspective on the protein thermal stability: A case study using SAICAR synthetase. *J. Chem. Inf. Model.*, 53: 2448-2461. DOI: 10.1021/ci400306m
- Moret, M.A. and G.F. Zebende, 2007. Amino acid hydrophobicity and accessible surface area. *Phys. Rev. E*, 75: 011920-011920. DOI: 10.1103/PhysRevE.75.011920
- Noorbach, I.A., A.M. Sultan, H.M. Salleh and A. Amid, 2013. Understanding thermostability factors of *Aspergillus niger* PhyA phytase: A molecular dynamics study. *Protein J.*, 32: 309-316. DOI: 10.1007/s10930-013-9489-y
- Oakley, A.J., Z. Prokop, M. Bohác, J. Kmuniček and T. Jedlicka *et al.*, 2002. Exploring the structure and activity of haloalkane dehalogenase from *Sphingomonas paucimobilis* UT26: Evidence for product-and water-mediated inhibition. *Biochemistry*, 41: 4847-4855. PMID: 11939779
- Prokop, Z., M. Monincová, R. Chaloupková, M. Klvaňa and Y. Nagata *et al.*, 2003. Catalytic mechanism of the haloalkane dehalogenase LinB from *Sphingomonas paucimobilis* UT26. *J. Biol. Chem.*, 278: 45094-45100. PMID: 12952988
- Paul, M., M. Hazra, A. Barman and S. Hazra, 2014. Comparative molecular dynamics simulation studies for determining factors contributing to the thermostability of chemotaxis protein "CheY". *J. Biomol. Struct. Dyn.*, 32: 928-49. DOI: 10.1080/07391102.2013.799438
- Razvi, A. and J.M. Scholtz, 2006. Lessons in stability from thermophilic proteins. *Protein Sci.*, 15: 1569-1578. PMID: 16815912
- Radestock, S. and H. Gohlke, 2011. Protein rigidity and thermophilic adaptation. *Proteins Struct. Funct. Bioinform.*, 79: 1089-1108. DOI: 10.1002/prot.22946
- Richmond, T.J., 1984. Solvent accessible surface area and excluded volume in proteins: Analytical equations for overlapping spheres and implications for the hydrophobic effect. *J. Mol. Biol.*, 178: 63-89. PMID: 6548264
- Shirdel, S.A., K. Khajeh, S.M. Asghari and H.R. Karbalaeei-Heidari, 2013. Enhancement of thermostability and resistance against autolysis in a zinc metalloprotease. *Eng. Life Sci.* DOI: 10.1002/elsc.201200218
- Salleh, A.B., A.S.M.A. Rahim, L.T. Rahman Rnzra and M. Basri, 2012. The role of Arg157Ser in improving the compactness and stability of ARM lipase. *J. Comput. Sci. Syst. Biol.*, 5: 039-046. DOI: 10.4172/jcsb.1000088

- Trejo, F., J.L. Gelpi, A. Ferrer, A. Boronat, M. Busquets and A. Cortés, 2001. Contribution of engineered electrostatic interactions to the stability of cytosolic malate dehydrogenase. *Protein Eng.*, 14: 911-917. PMID: 11742111
- Verschuere, K.H., S.M. Franken, H.J. Rozeboom, K.H. Kalk and B.W. Dijkstra, 1993. Refined X-ray structures of haloalkane dehalogenase at pH 6.2 and pH 8.2 and implications for the reaction mechanism. *J. Mol. Biol.*, 232: 856-872. PMID: 8355275
- Vogt, G., S. Woell and P. Argos, 1997. Protein thermal stability, hydrogen bonds and ion pairs. *J. Mol. Biol.*, 269: 631-643. PMID: 9217266
- Wackett, L.P., 1994. Dehalogenation in environmental biotechnology. *Curr. Opin. Biotechnol.*, 5: 260-265. PMID: 7765008
- Wijma, H.J. and D.B. Janssen, 2013. Computational design gains momentum in enzyme catalysis engineering. *FEBS J.*, 280: 2948-2960. DOI: 10.1111/febs.12324
- Wijma, H.J., R.J. Floor and D.B. Janssen, 2013. Structure- and sequence-analysis inspired engineering of proteins for enhanced thermostability. *Curr. Opin. Struct. Biol.*, 23: 588-594. DOI: 10.1016/j.sbi.2013.04.008
- Watanabe, K., T. Ohkuri, S.I. Yokobori and A. Yamagishi, 2006. Designing thermostable proteins: ancestral mutants of 3-isopropylmalate dehydrogenase designed by using a phylogenetic tree. *J. Mol. Biol.*, 355: 664-674. PMID: 16309701
- Wang, G.Z. and M.J. Lercher, 2010. Amino acid composition in endothermic vertebrates is biased in the same direction as in thermophilic prokaryotes. *BMC Evol. Biol.*, 10: 263-263. DOI: 10.1186/1471-2148-10-263
- Yokota, K., K. Satou and S.Y. Ohki, 2006. Comparative analysis of protein thermostability: Differences in amino acid content and substitution at the surfaces and in the core regions of thermophilic and mesophilic proteins. *Sci. Technol. Adv. Mater.*, 7: 255-262. DOI: 10.1016/j.stam.2006.03.003
- Yip, K.S., K.L. Britton, T.J. Stillman, J. Lebbink and W.M. de Vos *et al.*, 1998. Insights into the molecular basis of thermal stability from the analysis of ion-pair networks in the Glutamate Dehydrogenase family. *Eur. J. Biochem.*, 255: 336-346. PMID: 9716374
- Zhou, N.E., C.M. Kay and R.S. Hodges, 1993. Disulfide bond contribution to protein stability: Positional effects of substitution in the hydrophobic core of the two-stranded alpha-helical coiled-coil. *Biochemistry*, 32: 3178-3187. DOI: 10.1021/bi00063a033
- Zhang, H., J. Li, J. Wang, Y. Yang and M. Wu, 2014. Determinants for the improved thermostability of a mesophilic family 11 xylanase predicted by computational methods. *Biotechnol. Biofuels*, 7: 3-3. DOI: 10.1186/1754-6834-7-3
- Zhang, H., T. Zhang, K. Chen, S. Shen and J. Ruan *et al.*, 2009. On the relation between residue flexibility and local solvent accessibility in proteins. *Proteins Struct. Funct. Bioinform.*, 76: 617-636. DOI: 10.1002/prot.22375