Front. Biol. 2015, 10(6): 508–519
DOI 10.1007/s11515-015-1379-6

RESEARCH ARTICLE

Selection of effective and highly thermostable *Bacillus subtilis* lipase A template as an industrial biocatalyst–A modern computational approach

B. Senthilkumar, D. Meshachpaul, Rao Sethumadhavan, R. Rajasekaran (✉)

School of Biosciences and Technology, Bioinformatics Division, VIT University, Vellore 632014, Tamil Nadu, India

Abstract **Biocatalysts are intrinsically reactive and hence their operational stability is of vital significance for any bioprocess. The setback in biocatalyst stability has been tackled from diverse prospects. Inherently, stable biocatalysts are markedly realized and a regular attempt is being made to seek out new organisms that harbor them. Here, we analyzed the industrial biocatalyst lipase A (Native) of Bacillus subtilis and its six thermostable mutants (2M, 3M, 4M, 6M, 9M and 12M) computationally using conformational sampling technique. Consequently, the various structural events deciphering thermostability like root mean square deviation, root mean square fluctuation, radius of gyration and polar surface area showed mutant 12M to be highly stable with statistical validation. Besides, static model analysis involving intra-molecular interactions, secondary structure, solvent accessibility, hydrogen bond pattern, simulated thermal denaturation and desolvation energy also supported 12M comparatively. Of note, the presence of high secondary structural rigidity and hydrogen bonds increased thermostability and functionality of 12M, thus selecting it as a best template for designing thermostable lipases in future. Also, this study has a significant implication toward a better understanding of conformational sampling in enzyme catalysis and enzyme engineering.**

Keywords thermophilic, *Bacillus subtilis*, lipase A, conformational analysis, docking

Introduction

Enzyme production in large scale industries has been in use for several decades. Several basic techniques have been utilized widely for the screening of microbes and mutating it to obtain high and pure yield of the desired enzyme. Thermophilic enzymes have been a widely studied subject in past few decades. They serve as a remarkable tool in industrial biotechnology processes replacing the mesophilic enzymes (Zeikus et al., 1998). They also help in the better understanding of protein stability thus enlightening ways to stabilize other enzymes. Various microorganisms are used to produce these unique biocatalysts for retaining their activity at extreme conditions that prevail in industrial processes (Bruins et al., 2001). *Bacillus subtilis*, a ubiquitous, aerobic, spore forming, gram positive bacterium is largely used in the production of lipase, an industrial enzyme. It occurs naturally as mesophilic enzymes. Mutating the bacterium, thermostable enzymes with increased catalytic activity are produced, thereby finding its position in the field of biotechnology, (Tian et al., 2014). Few lipolytic enzymes secreted by *Bacillus subtilis* are lipase LipA (EC 3.1.1.3), esterase LipB (EC 3.1.1.1) and phospholipase. The atomic mass for *Bacillus subtilis* lipase A (LipA) is 19.3 kDa with a compact α/β hydrolase fold. It is composed of six β strands arranged analogous to each other with five α helices, two on one side and three on the other side of sheets (Eggert et al., 2001; van Pouderoyen et al., 2001).

Lipases (EC 3.1.1.3, triacylglycerol acylhydrolases) serve as a major class of hydrolase in the field of biotechnology. It aids in the hydrolysis of triacylglycerols releasing fatty acids and glycerol and it acts in reverse at oil-water interface. In addition, they not only possess the capability to hydrolyze various lipids, but also synthesize other esters or transesterification. Thus, they can be utilized as both hydrolases and

Received October 26, 2015; accepted December 4, 2015
Correspondence: R. Rajasekaran
E-mail: rrajasekaran@vit.ac.in

synthetases in various industrial areas. These distinguishing features of lipases are remarkably utilized in a wide array of industrial applications such as detergent industry (Ito et al., 1998; Sharma et al., 2001), pulp and paper industry (Jaeger and Reetz, 1998), resolution of racemic acids and alcohols (Klibanov, 1990; Klibanov, 1997), production of biodegradable polymers (Linko et al., 1998), diagnostic tool (Munoz and Katerndahl, 2000), cosmetics (Larios et al., 2004), medical applications (Annenkov et al., 2004), biosensors (Kynclova et al., 1995), food industry (Houde et al., 2004), tea processing (Hasan et al., 2006), degreasing of leather, oil biodegradation (Houde et al., 2004) and production of biodiesel (Noureddini et al., 2005). LipA includes a conserved region with Asp, Ser, and His constituting the catalytic triad (Gupta et al., 2004; Ma et al., 2006).

Industrial enzymes must possess both functional activity and structural integrity as they involve high temperature processes. The preservation of unique chemical and spatial structure of a polypeptide at higher temperature conditions is referred to as protein thermostability. The importance of protein thermostability, both experimentally and computationally has been studied for ages. Various factors are found to be involved in the stability of proteins, among them a few are cited consistently; they are increased hydrogen bonds, secondary structure stabilization, presence of disulphide bonds, number of proline residues, increase in polar surface area, short and stabilized loops, increase in buried surface area and surrounding hydrophobicity. The prologue mutations can affect the whole protein structure, hence mutational study of whole protein serves to be an eminent issue (Rajasekaran et al., 2011; Rajasekaran and Chen, 2012; Srivastava and Sinha, 2014).

The predominant factors for the stabilization of a protein are hydrophobic effect and hydrogen bonding (Pace et al., 1996). There are not much of computational studies precisely to support the parameters responsible for thermostablity of LipA. Day by day, the industrial demand for enzymes generated by microorganisms is growing swiftly (Jaeger and Reetz, 1998; Gupta et al., 2004). This study focuses on various parametrical key factors concerned with thermostablity and functionality of LipA in a computational aspect. In this study, native and six thermostable mutants of *Bacillus subtilis* LipA protein was selected. Previously thermostability of LipA was studied through Network and Dynamics approach (Srivastava and Sinha, 2014; Singh et al., 2015), thus explaining the presence of network residues and free energy landscape at different thermal conditions. The present study analyzes various stability enhancing parameters, specifically hydrogen bonds and its denaturation property to decipher the thermostability and functionality of LipA using conformational sampling technique. Additionally, the results will aid in identifying best LipA template computationally, which can serve as a starting point for experimental biologists to design *in vivo* thermostable lipases in future.

Materials and methods

Optimized structural data set

Structural coordinates of LipA native (1I6W) and six thermostable mutants (1T4M, 1T2N, 3D2A, 3D2B, 3D2C and 3QMM) were retrieved from PDB database (Berman et al., 2000). It is noted that, all six thermostable mutants attained their mutation by directed evolution. Moreover, each mutant preserved the earlier mutations with an increase in thermostability along with their mutations (Srivastava and Sinha, 2014). The PDB structure was optimized and structurally refined by energy minimization using Gromacs force field embedded in NOMAD-Ref program (Suhre and Sanejouand, 2004). The thermostability and functionality of LipA was analyzed using these optimized structures (Fig. 1).

Conformational sampling of LipA

NMSim, a multi-scale modeling program for protein conformational transitions, work on the basis of geometric simulation, involving three steps; (I) Coarse-grain of protein using FIRST software, (II) Normal-mode analysis of the rigid clusters and (III) Geometric simulations (NMSim). A stereochemical ensemble of protein structures, their corresponding conformational trajectory and data plots were speculated via structural analysis (Krüger et al., 2012). NMSim program used for conformational sampling, is said to be an efficient alternate to molecular dynamics simulations computationally (Ahmed et al., 2011).The Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF) along with the corresponding trajectory file of the given protein was obtained. The trajectory file was further evaluated for Radius of Gyration (Rg) and Polar Surface Area (PSA) using Vega ZZ program (Pedretti et al., 2004). The variations in values pronounced structural deviations seen in the proteins.

Additionally CABS-Flex, a coarse-grained protein modeling program resolves the structural flexibility of proteins via simulation (Jamroz et al., 2013). The given PDB file undergoes simulation yielding a dynamics trajectory, which in turn was subjected to structural clustering and flexibility analysis. These ensemble models were reconstructed based on atomic resolution and optimization (BBQ and ModRefiner tools) yielding 12 clusters by K-means clustering method. The program yields 12 individual PDB structures from the clusters representing highest iteratively possible stable structure. Subsequently, these structures were subjected to Bayesian Analysis Conformation Hunt (BACH) to discriminate the best model by BACHSCORE. BACH uses knowledge-based potential to compare the given protein structure with reference structure of the same, obtained from different databases on the basis of global distant test and atom energy score. Accordingly, the program displayed models and their

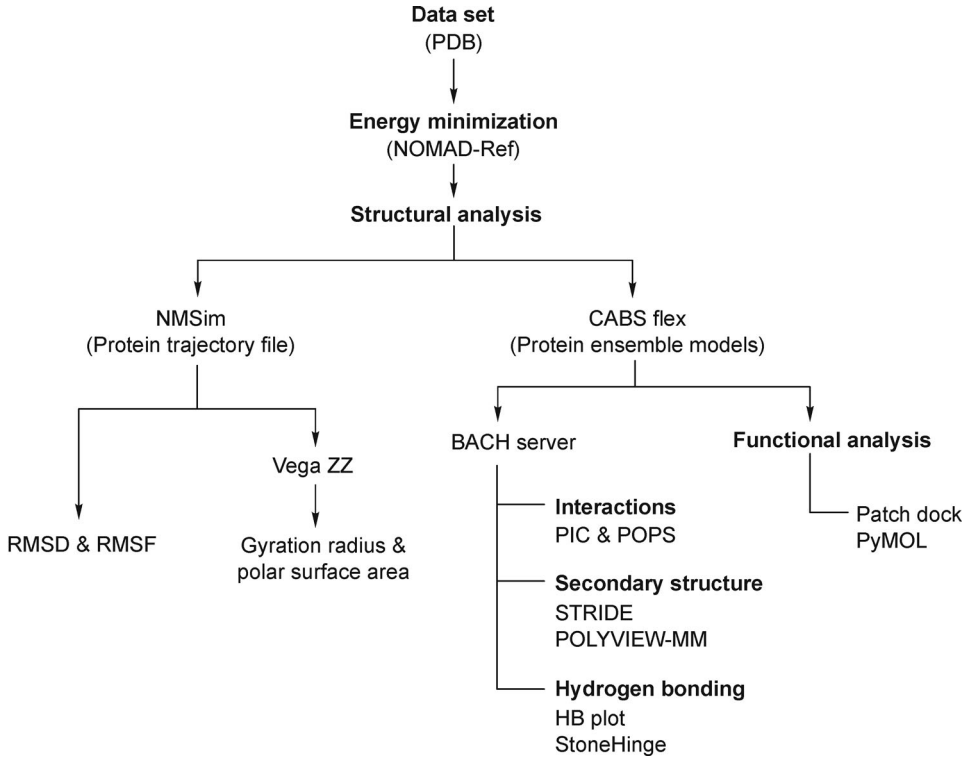

Figure 1 Flowchart illustration of the methodology and various bioinformatics tools used in this study.

corresponding energy scores as output. Thus, the conformer with least energy is predicted to be the best conformation (Cossio et al., 2012), which was utilized for single model analysis.

Structural interactions and properties

The various intra-molecular interactions of protein structure LipA was calculated using Protein Interaction Calculator (PIC) (Tina et al., 2007). The interactions used empirical and semi-empirical set of rules for calculations. The output displays various types of interactions and their corresponding distance between the residues (donor and acceptor) of the given protein. The structures of LipA was further computed with Parameter OPtimsed Surfaces (POPS) to calculate Solvent Accessible Surface Area (SASA), with the algorithm presented by Cavallo (Cavallo et al., 2003). Consequently, the program displays the hydrophobicity, hydrophilicity and their total values as an output. STRIDE (STRuctural IDEntification) program was used for the identification of protein secondary structures of given protein, based on φ and ψ angles of individual residues (Heinig and Frishman, 2004). The secondary structure distribution and their relative solvent accessibility was viewed two dimensionally using POLYVIEW-MM, a visualization tool (Porollo and Meller, 2010).

Hydrogen bonding and simulation of thermal denaturation

HB plot (Bikadi et al., 2007) was used to visualize the secondary structure of proteins with their corresponding stabilizing hydrogen bonds and amino acid residues. It computed the type of hydrogen bonds formed by the chain interactions as main, inter-chain, bifurcated and multiple hydrogen bonds. The results were displayed as Main chain-Main chain (M-M), Main chain-Side chain (M-S), Side chain-Main chain (S-M) and Side chain-Side chain (S-S) interactions based on the position of donor and acceptor molecules respectively. Besides, it categorized the range of hydrogen bonding into 3 classes; Short/Class I (distance between donor and acceptor is below 2.5Å), Intermediate/Class II (distance between 2.5Å and 3.2Å) and Long/Class III (distance above 3.2Å). Consequently, the hydrogen bonds involved in secondary and tertiary structural stabilization of LipA was studied.

Moreover, MSU ProFlex (formerly called FIRST) embedded in Stonehinge a computational tool designed for identification of rigid and flexible regions in the given proteins was used in this study. StoneHingeP defines the energy of protein decomposition, breaking it into two rigid areas held together by a flexible hinge using ProFlex. StoneHingeD defines the hinges with fixed residues, using DomDecomp analysis. Both these analyses aid in identification of flexible and rigid regions of the protein, along with their percent rigidity and residues contributing rigidity (Keating et al., 2009).

Functional analysis by docking

The role of receptor-ligand interaction is a well known

concept in biologic community, which is attained through docking analysis. A geometry-based molecular docking algorithm proposed by Schneidman-Duhovny et al.(2005) was used for docking our enzyme LipA with the substrate triglyceride. Triglyceride is the best substrate known for LipA both *in vitro* and *in vivo*. Hence, SMILES string was obtained for the molecule Triglyceride from PubChem (http://PubChem.ncbi.nl.nih.gov/) and 3D structure of the substrate was generated using CORINA (Gasteiger et al., 1990). The objective of docking is to identify the transformation of one molecule at which the other molecule can fit exactly without any steric changes (Chou et al., 2010; Lou et al., 2014). This not only depends on the geometric form of the molecules but also on its complex stability. The Patchdock program receives both receptor and ligand molecule as PDB files. The algorithmic calculations were based on three stages as molecular shape representation, surface patches matching and ranked scoring. The final outputs of files were ranked based on shape complementarity, interface shape and size of molecules. The output of program displayed geometric score, desolvation energy, area size and docked PDB file. This enabled to read the best template among mutants with high enzymatic activity (Duhovny et al., 2002). The docked molecules were visualized using PyMOL (The PyMOL Molecular Graphics System, Version 1.7.4 Schrödinger, LLC).

Statistical analysis

The non-parametric value obtained via Vega ZZ was analyzed using Kruskal–Wallis test for statistical significance. Kruskal–Wallis test (Kruskal, 1952) is an ordinary one-way ANOVA test. It was done using MS Excel without assuming Guassian distribution. The non parametric mean value of native and mutant lipases was compared simultaneously with each other. The significance of given parameter was detected using their probability value (*P*-value), where a *P*-value less than 0.05 is considered being statistically significant.

Results and discussion

Lipase A is a globular protein widely used in industrial biotechnology firm. The seven high resolution ($R \leqslant 2.5$Å) structures of native LipA and its thermostable mutants (van Pouderoyen et al., 2001; Acharya et al., 2004; Ahmad et al., 2008; Kamal et al., 2011) were collected from PDB for analysis (Table 1). In this study, native and six thermo stable mutants of LipA were studied on the basis of (I) compactness, (II) solvent accessibility and secondary structure distribution, (III) hydrogen bond classification and dilution and (IV) binding affinity.

Degree of variation and compactness of LipA

The optimized structure of LipA was initiated for conformation sampling using NMSim. The ensemble of LipA conformation obtained through NMSim was analyzed further. First, structural parameter RMSD and RMSF was analyzed. RMSD which shows deviation in backbone structure of the given protein model and RMSF that shows relative propensity of protein residues to deviate from an average dynamics structure was observed. Both play a vital role in defining thermostability of proteins (Unsworth et al., 2007). The calculated mean RMSD values of mutants, viz., 2M, 3M, 4M, 6M, 9M and 12M were 1.107Å, 1.171Å, 1.117Å, 1.197Å, 1.344Å and 1.147Å respectively, failed to show any marginal changes with that of native (1.205Å) ($p < 0.0001$) comparatively (Fig. 2A). The RMSD data that lacked significance showed similarity to results obtained earlier using classical molecular dynamics (Srivastava and Sinha, 2014). Subsequently, RMSF mean values of the mutants, viz., 2M, 3M, 4M, 6M, 9M and 12M were 0.580Å, 0.497Å, 0.489Å, 0.508Å, 0.598Å and 0.469Å respectively, which also failed to represent a marked rigidity in residues of mutants compared to that of native (0.500Å). Among six mutants, 6M alone showed a significant increase in rigidity, thus showing similarity to earlier classical molecular dynamics results (Kamal et al., 2011; Srivastava and Sinha, 2014). Overall, both RMSD and RMSF values failed to show structural deviations among the LipA native and six mutants (Fig. 2B), but their results increased the reliability of conformational sampling compared to classical molecular dynamics in protein dynamics. Thereby, other structural parameters that aid in stability of LipA was further analyzed.

Radius of Gyration (R_g), a major parameter for protein

Table 1 *Bacillus subtilis* LipA structures along with their computed intra-molecular interactions (IMM), hydrophobic and hydrophilic values of SASA

SI No.	PDB ID	Mutant name	Number of mutations	Total No. of IMM	Hydrophobic (Å)	Hydrophilic (Å)
1	1I6W	Native	Native	152	5782.87	4558.73
2	1T4M	2M	2	154	5428.72	4131.86
3	1T2N	3M	3	150	5651.67	4194.25
4	3D2A	4M	4	141	5913.04	4213.21
5	3D2B	6M	6	154	5528.73	4359.09
6	3D2C	9M	9	152	5683.36	4199.47
7	3QMM	12M	12	160	5497.66	4101.97

SI: serial; PDB: Protein Data Bank.

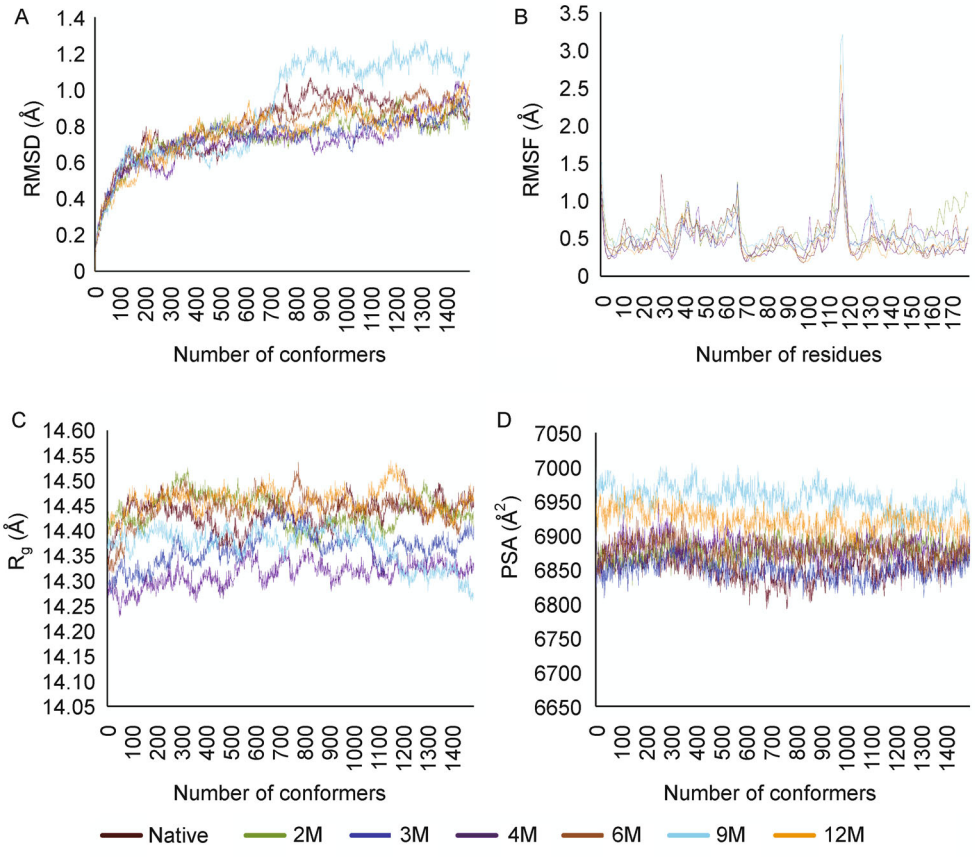

Figure 2 Geometrical observables of native Lip A and its mutants (A) Root Mean Square Deviation (RMSD), (B) Root Mean Square Fluctuation (RMSF), (C) Radius of Gyration (R_g) and (D) Polar surface area (PSA)

stability calculates geometrical size, molecular spatial packing and compactness of the given protein (Lobanov et al., 2008). A reduced R_g value corresponds to increase in compactness of corresponding protein. The analyzed mean values ($p < 0.0001$) for the following mutants, viz., 2M, 3M, 4M, 6M, 9M and 12M was 14.44Å, 14.37Å, 14.31Å, 14.45Å, 14.37Å and 14.46Å respectively, compared to that of native (14.43Å). The LipA mutants showed similar R_g values compared to native, except few mutants that showed an increase in compactness (Fig. 2C) comparatively. The compactness of protein also depends on the presence of polar and non-polar residues. PSA is defined as the sum of surfaces containing polar atoms (usually oxygens, nitrogens and attached hydrogens) in the given molecule. The apolar atoms C and H bonded to C; polar atoms O, S, N, P and H not bonded to C was calculated by their presence on atom surfaces (Pace et al., 2011). The computed PSA mean data showed an increase in mutant values ranging between 6876.163Å^2 and 6954.319Å^2 than that of native (6851.38Å^2) using Vega ZZ with statistical significance ($p < 0.0001$), which is graphically represented in Fig. 2D (Gaillard et al., 1994). The PSA mean value for the following mutants, viz., 2M, 3M, 4M, 6M, 9M and 12M was 6876.16Å^2, 6850.00Å^2, 6881.90Å^2, 6879.60Å^2, 6954.32Å^2 and 6922.14Å^2 respectively. Accordingly, the PSA values of the LipA mutant were found to be higher than the native. As per previous studies, the PSA is said to increase along with the hydrogen bonding density to water, thus constituting increased thermostability (Vogt and Argos, 1997). Hence, an analysis determining the presence and function of hydrogen bonds among thermostable mutants will enlighten the knowledge of LipA thermostability.

The values obtained via structural parameters, RMSF and RMSD of LipA ensembles enabled to monitor the reliability of conformational sampling to that of classical molecular dynamics. The Rg and PSA values showed an increased stability in LipA mutants compared to that of native, whereas to determine the highly stable mutant among for future development of thermostable lipases, a static model analysis of LipA is eminent. A static model analysis will provide needed structural and functional stability details of Lip A.

Intramolecular characteristics of LipA

CABS-Flex program was used to obtain a set of protein models reflecting the most dominant structural fluctuations in near-native ensemble. The best represented model of CABS-Flex program obtained through BACH score was utilized as

an input for static model analyses of LipA. First, the intramolecular interaction exhibited by structural base of LipA was calculated using PIC. The interactions involving hydrogen bonds were exempted as it's been studied in-depth later. Consequently, the number of interactions exhibited by native was 152 and that of six mutants, viz., 2M, 3M, 4M, 6M, 9M and 12M was 154, 150, 141, 154, 152 and 160 respectively as shown in Table 1. The marked differences among mutants reflected structural and functional change of LipA, consequently on the rigidity of protein (Mahalingam et al., 2014a, 2014b). The compactness and rigidity of a protein can also be calculated based on its solvent accessible surface area (Srivastava and Sinha, 2014). The ratio of buried non polar surface area to total non polar surface area of a protein is identified as hydrophobicity. Hydrophobic forces play a vital role in regard to thermostability and molecular folding of a protein (Goodenough and Jenkins, 1991; Pace et al., 2011). According to the SASA data given in Table 1, the native (55.9%) showed a decrease in hydrophobicity whereas mutants showed an increased variation of 56.8% (2M), 57.4% (3M), 58.4% (4M), 57.5% (9M) and 57.3% (12M), except the mutant 6M (55.9%), whose value was similar to that of native. Among the six thermostable mutants, four mutants, viz., 3M, 4M, 9M and 12M exhibited an increased hydrophobicity.

The distribution of relative solvent accessibility was depicted in Fig. 3. The black shades represented buried residues and white shades the exposed residues. While, the varying shades of black was characterized based on the ratio of buried and exposed residues. Larger the number of darker regions higher is the hydrophobicity of protein, the patterns seen in Fig. 3 showed a strong similarity to that of the obtained SASA values. The increase in hydrophobic residues constitute for compactness of a protein which had been directly related to their secondary structures (Lobanov et al., 2008; Lou et al., 2009). Thus, secondary structure data of protein was obtained by STRIDE that clarified the number of amino acids involved in formation of α helix, β sheets and coils as shown in Fig. 4.

On that account, native showed 13% α helix, whereas mutants showed an increase between 14% and 16% except 2M which showed a decrease of 2% than native comparatively. Similarly, the β sheets contributed 12% in native while mutants showed a range between 15% and 16% rise except 9M and 2M which counted 12% and 13% respectively. Of note, the coil was found to be high in native (17%) than that of mutants (13% –15%) as shown in Fig. 4. The distribution of secondary structures was plotted graphically in Fig. 5, illustrating α helix, β sheets and coils. The data showed a high variance of secondary structures in four mutants, 3M, 4M, 6M and 12M to that of native. The distribution of secondary structures of a protein, in specific, α helix and β

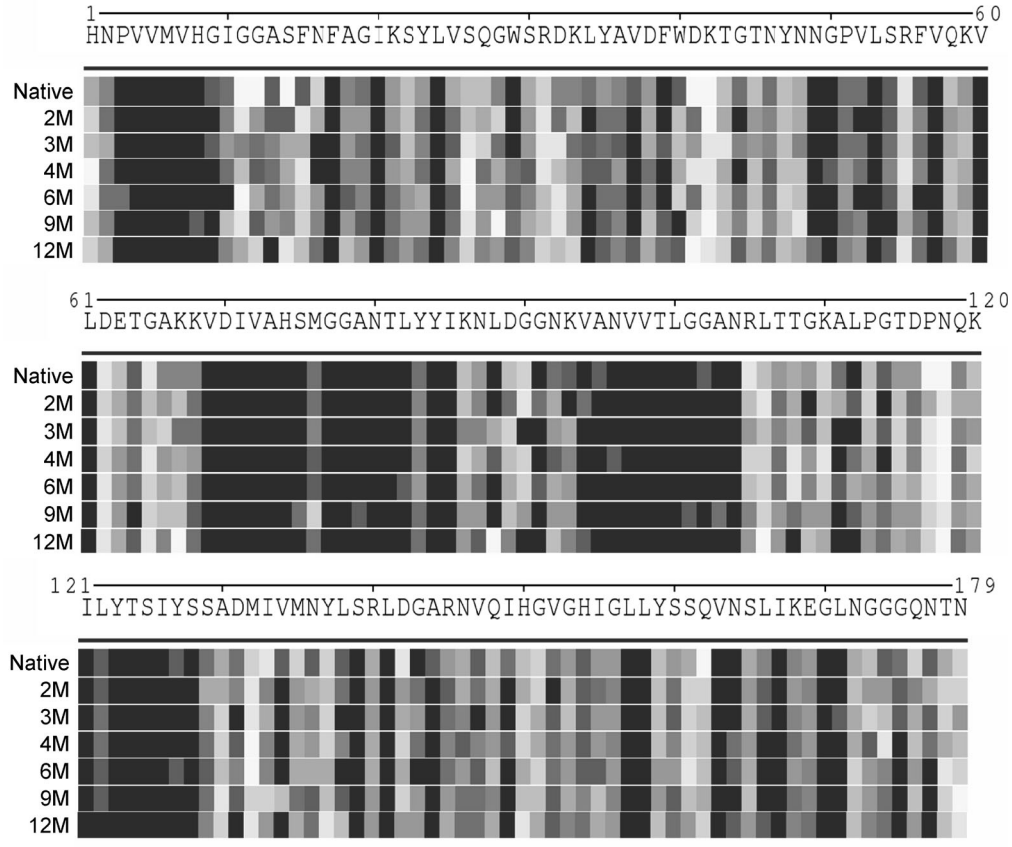

Figure 3 Two dimensional view of variable relative solvent solubility of native LipA and its mutants.

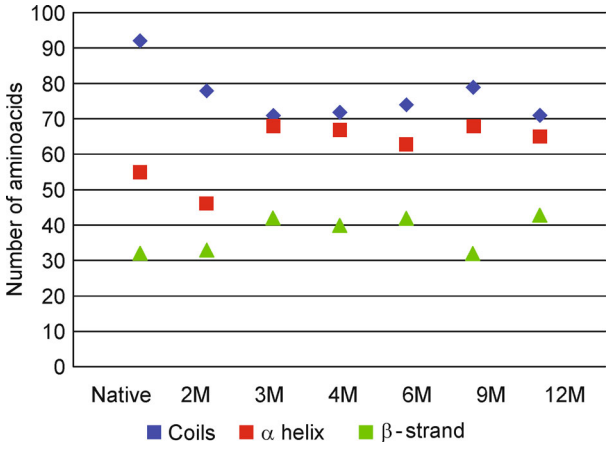

Figure 4 Scatter plot showing the distribution of secondary structures in the native LipA and its mutants.

sheets depend on hydrogen bonds accounting for stability of protein (Fersht et al., 1991). Thus, in depth analysis of hydrogen bonds and other features pertaining to stability was inevitable.

Hydrogen bond pattern and dilution of LipA

The secondary and tertiary structure of a protein is related directly to the number of hydrogen bonds (Bandyopadhyay et al., 2005; Luo et al., 2013). The secondary and tertiary structural distribution of hydrogen bonds was studied using HB plot. An increase in the number of Class II bonding was noted in two mutants, 9M and 12M demonstrating the possibility of high number of hydrogen bonds in LipA mutants. On the other hand, stability of proteins was said to increase deliberately with increased side chain hydrogen bonding (M-S, S-M and S-S) and low main chain-main chain (M-M) interactions (Ji and Liu, 2011; Pace et al., 2014). According to Table 2, our study demonstrated the decrease of main chain-main chain interactions in all mutants (59.2%–62.9%) comparatively than native (64.9%). While, an increase in interactions involving side chain (M-S, S-M and S-S) was seen among mutants to that of native. The comparisons relating to side chain-side chain (S-S) interactions of native (26.5%) and mutants (31.25%– 37%) showed a discernible difference. Thus, the three mutants, viz., 4M, 9M and 12M were found to be significant comparatively.

Presence of a single pair of acceptor and donor residues constitute a normal hydrogen bond. Whereas, if the pattern of hydrogen bonding includes more than one acceptor or donor, they are named as bifurcated hydrogen bonds whose functionality involves in the bending of helices and side chain rotamer (Feldblum and Arkin, 2014). Moreover, the allocation of bifurcated hydrogen bonds obtained by HB plot was simultaneously studied as shown in Table 3. Apart from 4M and 12M, all other mutants showed a decrease in

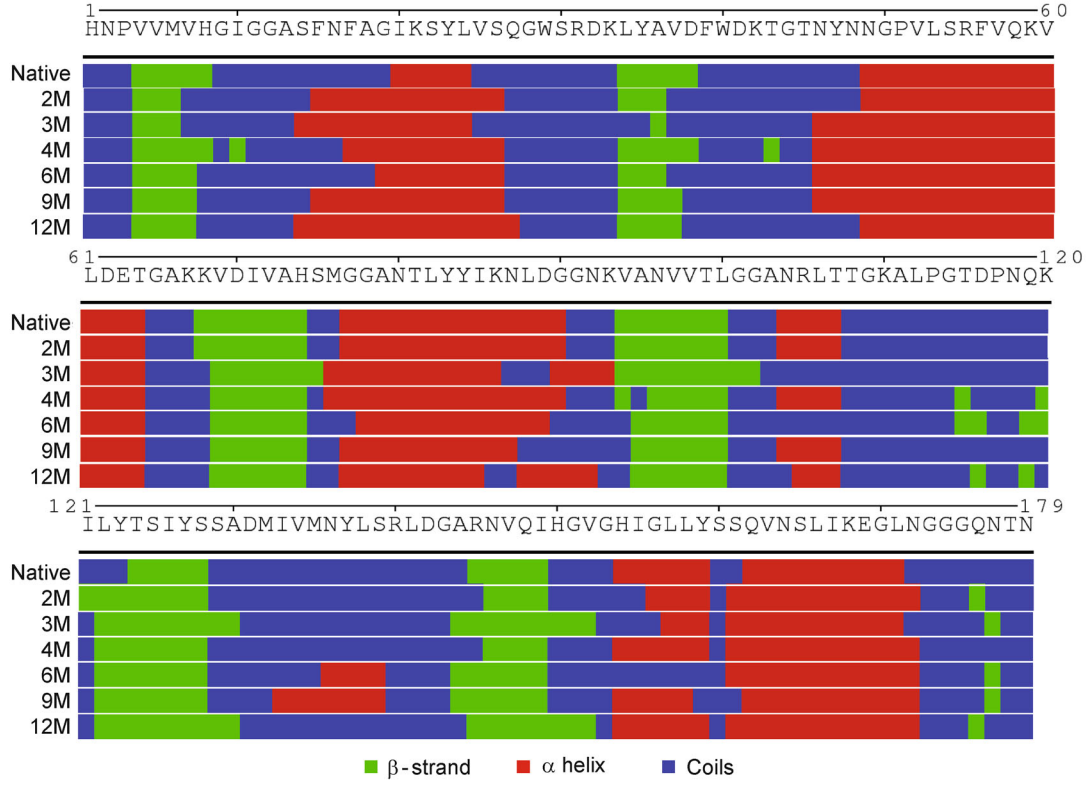

Figure 5 Two dimensional representation of secondary structures for native LipA and its mutants.

Table 2 Structural distribution of hydrogen bonds in native LipA and its mutants

Parameters	Range and type of H-bonding	Native	2M	3M	4M	6M	9M	12M
		Number of hydrogen bonds						
Range	Class I	0	0	0	0	0	0	0
	Class II	189	188	187	185	186	192	197
	Class III	5	1	3	4	4	5	2
Type	M – M	126	119	114	112	114	118	119
	M – S	17	13	14	15	13	16	21
	S – M	33	33	36	39	39	35	34
	S – S	18	27	26	23	24	28	25

Table 3 Structural distribution of bifurcated hydrogen bonds for native LipA and its mutants

Parameters	Range and type of H-bonding	Native	2M	3M	4M	6M	9M	12M
		Number of hydrogen bonds						
Class	Class I	0	0	0	0	0	0	0
	Class II	23	20	16	26	22	23	30
	Class III	1	0	0	0	0	1	0
Type	M – M	11	4	2	5	4	5	10
	M – S	1	0	0	1	0	1	2
	S – M	11	9	9	14	11	11	13
	S – S	1	7	5	6	7	7	5

Class II bonding compared to native, except 9M which showed similar numbers to that of native comparatively. Subsequently, the Table 3 exhibited a decrease in main chain-main chain interactions in all mutants (12.5%–33.3%) than that of native (45.8%). On the other hand, an increased interaction of side chain (M-S, S-M and S-S) of mutants (66.7%–87.5%) to that of native (54.2%) was observed. In addition, side chain-side chain interactions revealed a large difference between native (7.7%) and mutants (25%–43.75%) comparatively. The results were analogous to Table 2, thus showing an increase in thermostability of LipA mutants, viz., 3M, 6M, 9M and 12M.

Further, these data illustrated the presence of side chain interactions (M-S, S-M and S-S) and their contribution to increased structural stability in LipA. The study involving dilution pattern of hydrogen bonds using Stonehinge can aid in determining its rigidity in depth. The transition of folded rigid protein to a denatured state is called unfolding. At the process of unfolding, the hydrogen bonds and salt bridges are broken resulting in deformation of secondary and tertiary structures. This can be achieved by thermal denaturation of the protein which results in bond dilution. The weaker bonds are broken first followed by stronger ones and this aids in computation of flexible and rigid regions of a given protein structurally (Keating et al., 2009). The varied energy levels required to break hydrogen bonds were evaluated and thus, higher the energy required stronger is the bonding. Henceforth, the number of bonds that remained to be broken after each energy level depicted the presence of flexible and rigid regions in the given protein. The number of stable hydrogen bonds that remained to be broken till the end point was higher among mutants (15–23) than that of native (9), except mutant 4M which showed similarity with that of native (Fig. 6) comparatively. Of note, a large number of bonds were broken relatively both in native (26) and mutants (16–55) at energy levels between −1 and −2.9 kcal/mol. Besides, the evaluated percent rigidity of LipA native (42.9%) and its mutants (47.3%–73%) showed a large variation, except mutant 3M (34.9%) which showed a lower percent. Similar result was obtained while calculating the number of cluster residues constituting rigid regions in LipA, where native (84) and mutants (91–137) showed large variations, while the mutant 3M (72) registered a lower number. Analyzing these data, four mutants viz., 2M, 6M, 9M and 12M were said to be more rigid among the given mutants as seen in Table 4. The results of previous study (Vogt and Argos, 1997) and present substantiated the fact that hydrogen bonding increases stability and rigidity of a protein, in this case LipA of *Bacillus subtilis*.

Average binding affinity of lipase A

The functional analysis of LipA was calculated by docking triglyceride with native and mutant models using Patch dock. The program Patch dock renders the information regarding enzymatic efficiency and binding capacity of the enzyme

Table 4 Percent rigidity and cluster residues that constitute *Bacillus subtilis* Lip A native and its mutants

SI No.	PDB ID	Cluster positions	Number of cluster residues	Percent rigidity (%)
1	Native	2-9, 28-39, 48-64, 68-86, 93-100, 153-172	84	42.9
2	2M	4-9, 14-42, 47-103, 123-130, 133-136, 147-179	137	73
3	3M	5-9, 16-28, 51-66, 71-76, 79-89, 98-102, 124-128, 164-174	72	34.9
4	4M	4-11, 18-31, 36-41, 50-67, 72-77, 78-90, 98-103, 124-129, 162-175	91	47.3
5	6M	4-31, 35-41, 51-67, 70-91, 96-102, 106-110, 123-130, 147-176	124	65.8
6	9M	4-10, 16-41, 50-90, 96-103, 123-131, 134-141, 146-177	131	70.9
7	12M	4-10, 15-42, 47-51, 51-67, 70-89, 96-103, 124-130, 146-151, 162-175	112	58.7

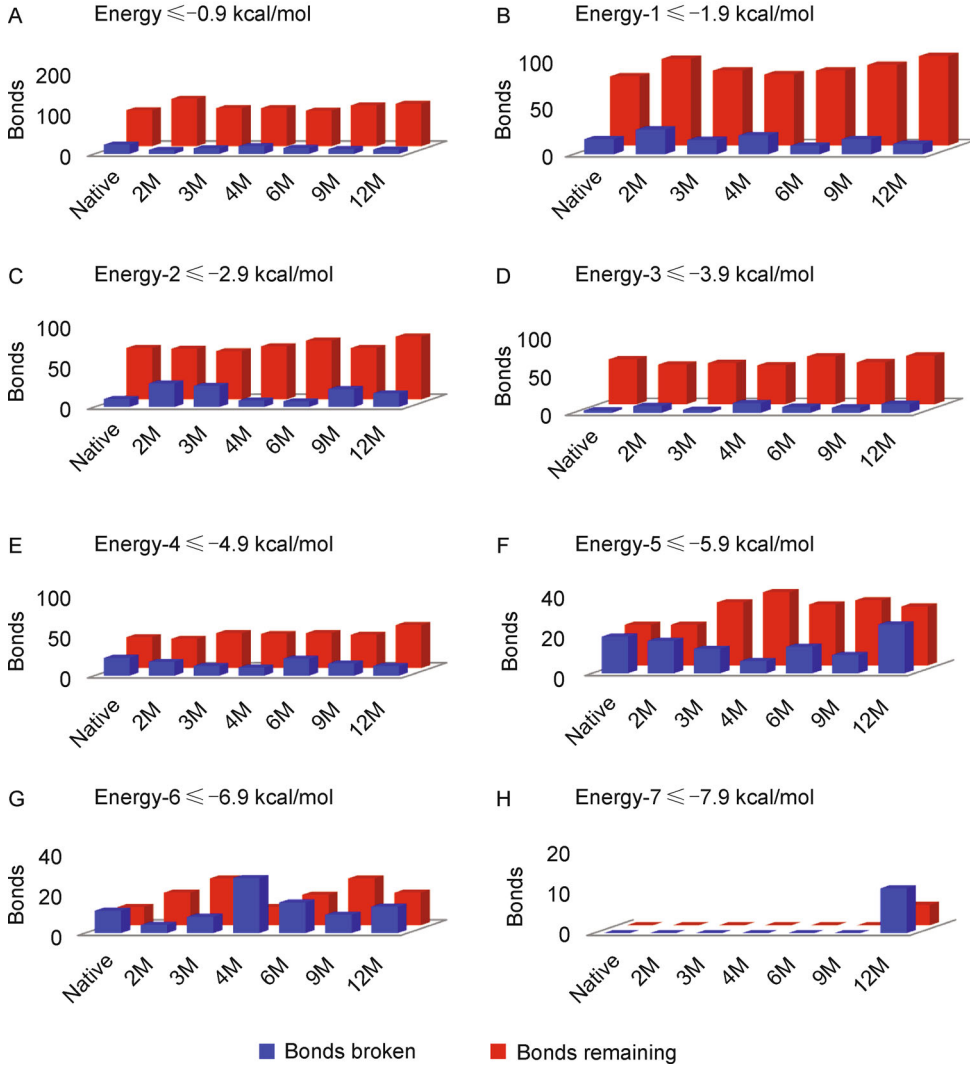

Figure 6 A comparative graphical representation of hydrogen bond dilution, showing the number of hydrogen bonds broken and the number of hydrogen bonds remained to be broken, at different energy levels (A-H).

LipA with its corresponding substrate, triglyceride. The two major parameters that determine the binding free energy of a protein–ligand complex in water environment are binding enthalpy and solvation/desolvation free energy. Desolvation energy aids in evaluating the binding interaction of a true ligand. As per the details in Fig. 7A, the models mean desolvation energy (Atomic Contact Energy/ACE) for the following mutants 2M, 3M, 4M, 6M, 9M and 12M was –37.94 kcal/mol, –57.66 kcal/mol, –43.82 kcal/mol, –52.02 kcal/mol, –38.6 kcal/mol and –67.17 kcal/mol respectively with native showing –35.88kcal/mol. Thus, the results showed an increase in binding efficiency of mutants along with their thermostability and mutations. The Fig. 7B illustrated the superimposition of best scored docked

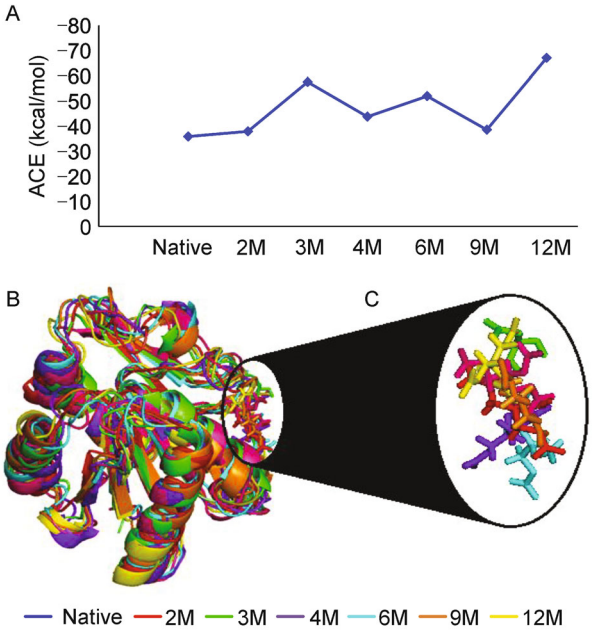

Figure 7 Diagramatic view of the docked native Lip A and its mutants with triglyceride (A) mean ACE (B) superimposed structure of the docked LipA native and mutants and (C) the orientation of triglyceride at the active site.

complexes of native and mutants, where the models en route the position at which *Bacillus subtilis* LipA was appended with the substrate triglyceride. Besides, binding of substrate triglyceride (Fig. 7C) was seen with slight variations in their spatial orientation which may be caused due to deviation in conformational orientation of LipA mutants. Moreover, structural change in the orientation of mutants compared to that of native might be a said reason for the improved functionality and thermostability in LipA mutants. Thereby, the three mutants, viz., 3M, 6M and 12M are found to have best docked complex energy along with functionality, an important parameter for industrial firm.

Conclusion

Native and six thermophilic mutants of LipA were analyzed *in silico*, where the lack in differences among RMSD and RMSF values failed to depict structural variation between native and mutants. Whereas, Rg and PSA showed a considerable difference in mutants than native comparatively, supporting stability. Thus, leading the search toward hydrogen bonds and other physiochemical parameters, that determines structural and functional stability among LipA mutants. Hydrophobicity of native (55.9%) was comparatively less than that of mutants (4M- 58.4%, 9M- 57.5% and 12M- 57.3%). Subsequently, an increase in complex secondary structural formation of mutants (3M- 16%, 4M- 15%, 6M- 15%, 9M- 14% and 12M- 16%) than native (13%) was noted. Hydrogen bonding pattern and its thermal dilution showed a significant raise in side chain interactions among the three mutants 4M (40.8%), 9M (40.2%) and 12M (40.3%) compared to native (35%), thus supporting mutant stability. In addition, the docking analysis explained an augmentation in desolvation energy of three mutants 6M, 3M and 12M with −52.02kcal/mol, −57.66kcal/mol and −67.19kcal/mol respectively, representing their enzyme activity to be higher than native (−35.88kcal/mol). Though, the three LipA mutants 4M, 9M and 12M showed a consistent results pertaining to stability analysis in docking studies, 12M showed increased functionality. An industrial enzyme is selected based on its thermostability and functional activity; hence the mutant 12M has been selected as an effective industrial biocatalyst among LipA mutants. The study demonstrated the presence of hydrogen bonds as one of the solitary reasons for survival of *Bacillus subtilis* LipA as a thermophilic enzyme computationally. Thus, in future thermostable lipases can be designed based on its hydrogen bond numbers with 12M as a stable template. Through this study, we also like to register the use of conformation sampling as an alternate for classical molecular dynamics (MD), which is laborious and time consuming. The comparison of results obtained via conformational sampling and classical MD showed a lucid level of similarity, thus pronouncing our researches to take a leap into technological improvisation.

Acknowledgements

The authors thank the management of VIT University for providing the facilities and encouragement to carry out this research work.

Compliance with ethics guidelines

B. Senthilkumar, D. Meshachpaul, Rao Sethumadhavan and R. Rajasekaran declare that they have no conflict of interest.

This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Acharya P, Rajakumara E, Sankaranarayanan R, Rao N M (2004). Structural basis of selection and thermostability of laboratory evolved *Bacillus subtilis* lipase. J Mol Biol, 341(5): 1271–1281

Ahmad S, Kamal M Z, Sankaranarayanan R, Rao N M (2008). Thermostable *Bacillus subtilis* lipases: in vitro evolution and structural insight. J Mol Biol, 381(2): 324–340

Ahmed A, Rippmann F, Barnickel G, Gohlke H (2011). A normal mode-based geometric simulation approach for exploring biologically relevant conformational transitions in proteins. J Chem Inf Model, 51 (7): 1604–1622

Annenkov G A, Klepikov N N, Martynova L P, Puzanov V A (2004). Wide range of the use of natural lipases and esterases to inhibit *Mycobacterium tuberculosis.* Probl Tuberk Bolezn Legk, (6): 52–56

Bandyopadhyay S, Chakraborty S, Bagchi B (2005). Secondary structure sensitivity of hydrogen bond lifetime dynamics in the

protein hydration layer. J Am Chem Soc, 127(47): 16660–16667

Berman H M, Westbrook J, Feng Z, Gilliland G, Bhat T N, Weissig H, Shindyalov I N, Bourne P E (2000). The Protein Data Bank. Nucleic Acids Res, 28(1): 235–242

Bikadi Z, Demko L, Hazai E (2007). Functional and structural characterization of a protein based on analysis of its hydrogen bonding network by hydrogen bonding plot. Arch Biochem Biophys, 461(2): 225–234

Bruins M E, Janssen A E, Boom R M (2001). Thermozymes and their applications: a review of recent literature and patents. Appl Biochem Biotechnol, 90(2): 155–186

Cavallo L, Kleinjung J, Fraternali F (2003). POPS: A fast algorithm for solvent accessible surface areas at atomic and residue level. Nucleic Acids Res, 31(13): 3364–3366

Chou C C, Rajasekaran M, Chen C (2010). An effective approach for generating a three-Cys2His2 zinc-finger-DNA complex model by docking. BMC Bioinformatics, 11(1): 334

Cossio P, Granata D, Laio A, Seno F, Trovato A (2012). A simple and efficient statistical potential for scoring ensembles of protein structures. Sci Rep, 2(14):Available at: http://www.nature.com/doifinder/10.1038/srep00351

Duhovny D, Nussinov R, Wolfson H J (2002). Efficient unbound docking of rigid molecules. In Algorithms in bioinformatics. Springer, pp. 185–200

Eggert T, van Pouderoyen G, Dijkstra B W, Jaeger K E (2001). Lipolytic enzymes LipA and LipB from *Bacillus subtilis* differ in regulation of gene expression, biochemical properties, and three-dimensional structure. FEBS Lett, 502(3): 89–92

Feldblum E S, Arkin I T (2014). Strength of a bifurcated H bond. Proc Natl Acad Sci USA, 111(11): 4085–4090

Fersht A R, Bycroft M, Horovitz A, Kellis J T Jr, Matouschek A, Serrano L (1991). Pathway and stability of protein folding. Philos Trans R Soc Lond B Biol Sci, 332(1263): 171–176

Gaillard P, Carrupt P A, Testa B, Boudon A (1994). Molecular lipophilicity potential, a tool in 3D QSAR: method and applications. J Comput Aided Mol Des, 8(2): 83–96

Gasteiger J, Rudolph C, Sadowski J (1990). Automatic generation of 3D-atomic coordinates for organic molecules. Tetrahedron Comput. Methodol., 3(6): 537–547

Goodenough P W, Jenkins J A (1991). Protein engineering to change thermal stability for food enzymes. Biochem Soc Trans, 19(3): 655–662

Gupta R, Gupta N, Rathi P (2004). Bacterial lipases: an overview of production, purification and biochemical properties. Appl Microbiol Biotechnol, 64(6): 763–781

Hasan F, Shah A A, Hameed A (2006). Industrial applications of microbial lipases. Enzyme Microb Technol, 39(2): 235–251

Heinig M, Frishman D (2004). STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. Nucleic Acids Res, 32(Web Server issue): W500-2

Houde A, Kademi A, Leblanc D (2004). Lipases and their industrial applications: an overview. Appl Biochem Biotechnol, 118(1-3): 155–170

Ito S, Kobayashi T, Ara K, Ozaki K, Kawai S, Hatada Y (1998). Alkaline detergent enzymes from alkaliphiles: enzymatic properties, genetics, and structures. Extremophiles, 2(3): 185–190

Jaeger K E, Reetz M T (1998). Microbial lipases form versatile tools for biotechnology. Trends Biotechnol, 16(9): 396–403

Jamroz M, Kolinski A, Kmiecik S (2013). CABS-flex: Server for fast simulation of protein structure fluctuations. Nucleic Acids Res, 41 (Web Server issue): W427-31

Ji X L, Liu S Q (2011). Is stoichiometry-driven protein folding getting out of thermodynamic control? J Biomol Struct Dyn, 28(4): 621–623, discussion 669–674

Kamal M Z, Ahmad S, Molugu T R, Vijayalakshmi A, Deshmukh M V, Sankaranarayanan R, Rao N M (2011). In vitro evolved non-aggregating and thermostable lipase: structural and thermodynamic investigation. J Mol Biol, 413(3): 726–741

Keating K S, Flores S C, Gerstein M B, Kuhn L A (2009). StoneHinge: hinge prediction by network analysis of individual protein structures. Protein Sci, 18(2): 359–371

Klibanov A M (1990). Asymmetric transformations catalyzed by enzymes in organic solvents. Acc Chem Res, 23(4): 114–120

Klibanov A M (1997). Why are enzymes less active in organic solvents than in water? Trends Biotechnol, 15(3): 97–101

Krüger D M, Ahmed A, Gohlke H (2012). NMSim web server: integrated approach for normal mode-based geometric simulations of biologically relevant conformational transitions in proteins. Nucleic Acids Res, 40(Web Server issue): W310-6

Kruskal W H (1952). A nonparametric test for the several sample problem. Ann Math Stat, 23(4): 525–540

Kynclova E, Hartig A, Schalkhammer T (1995). Oligonucleotide labelled lipase as a new sensitive hybridization probe and its use in bio-assays and biosensors. J Mol Recognit, 8(1-2): 139–145

Larios A, García H S, Oliart R M, Valerio-Alfaro G (2004). Synthesis of flavor and fragrance esters using *Candida antarctica* lipase. Appl Microbiol Biotechnol, 65(4): 373–376

Linko Y Y, Lämsä M, Wu X, Uosukainen E, Seppälä J, Linko P (1998). Biodegradable products by lipase biocatalysis. J Biotechnol, 66(1): 41–50

Lobanov M Y, Bogatyreva N S, Galzitskaya O V (2008). Radius of gyration as an indicator of protein structure compactness. Mol Biol, 42(4): 623–628

Lou Y C, Wang I, Rajasekaran M, Kao Y F, Ho M R, Hsu S T D, Chou S H, Wu S H, Chen C (2014). Solution structure and tandem DNA recognition of the C-terminal effector domain of PmrA from *Klebsiella pneumoniae.* Nucleic Acids Res, 42(6): 4080–4093

Lou Y C, Wei S Y, Rajasekaran M, Chou C C, Hsu H M, Tai J H, Chen C (2009). NMR structural analysis of DNA recognition by a novel Myb1 DNA-binding domain in the protozoan parasite *Trichomonas vaginalis.* Nucleic Acids Res, 37(7): 2381–2394

Luo S C, Lou Y C, Rajasekaran M, Chang Y W, Hsiao C D, Chen C (2013). Structural basis of a physical blockage mechanism for the interaction of response regulator PmrA with connector protein PmrD from *Klebsiella pneumoniae.* J Biol Chem, 288(35): 25551–25561

Ma J, Zhang Z, Wang B, Kong X, Wang Y, Cao S, Feng Y (2006). Overexpression and characterization of a lipase from *Bacillus subtilis.* Protein Expr Purif, 45(1): 22–29

Mahalingam R, Peng H P, Yang A S (2014). Prediction of fatty acid-binding residues on protein surfaces with three-dimensional probability distributions of interacting atoms. Biophys Chem, 192: 10–19

Mahalingam R, Peng H P, Yang A S (2014). Prediction of FMN-binding residues with three-dimensional probability distributions of interacting atoms on protein surfaces. J Theor Biol, 343: 154–161

Munoz A, Katerndahl D A (2000). Diagnosis and management of acute pancreatitis. Am Fam Physician, 62(1): 164–174

Noureddini H, Gao X, Philkana R S (2005). Immobilized *Pseudomonas cepacia* lipase for biodiesel fuel production from soybean oil. Bioresour Technol, 96(7): 769–777

Pace C N, Fu H, Fryar K L, Landua J, Trevino S R, Shirley B A, Hendricks M M, Iimura S, Gajiwala K, Scholtz J M, Grimsley G R (2011). Contribution of hydrophobic interactions to protein stability. J Mol Biol, 408(3): 514–528

Pace C N, Fu H, Lee Fryar K, Landua J, Trevino S R, Schell D, Thurlkill R L, Imura S, Scholtz J M, Gajiwala K, Sevcik J, Urbanikova L, Myers J K, Takano K, Hebert E J, Shirley B A, Grimsley G R (2014). Contribution of hydrogen bonds to protein stability. Protein Sci, 23 (5): 652–661

Pace C N, Shirley B A, McNutt M, Gajiwala K (1996). Forces contributing to the conformational stability of proteins. FASEB J, 10 (1): 75–83

Pedretti A, Villa L, Vistoli G (2004). VEGA—an open platform to develop chemo-bio-informatics applications, using plug-in architecture and script programming. J Comput Aided Mol Des, 18(3): 167–173

Porollo A, Meller J (2010). POLYVIEW-MM: web-based platform for animation and analysis of molecular simulations. Nucleic Acids Res, 38(Web Server issue): W662-6

Rajasekaran M, Abirami S, Chen C (2011). Effects of single nucleotide polymorphisms on human N-acetyltransferase 2 structure and dynamics by molecular dynamics simulation. PLoS ONE, 6(9): e25801

Rajasekaran M, Chen C (2012). Structural effect of the L16Q, K50E, and R53P mutations on homeodomain of pituitary homeobox protein 2. Int J Biol Macromol, 51(3): 305–313

Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson H J (2005). PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic Acids Res, 33(Web Server issue): W363-7

Sharma R, Chisti Y, Banerjee U C (2001). Production, purification, characterization, and applications of lipases. Biotechnol Adv, 19(8): 627–662

Singh B, Bulusu G, Mitra A (2015). Understanding the thermostability and activity of *Bacillus subtilis* lipase mutants: insights from molecular dynamics simulations. J Phys Chem B, 119(2): 392–409

Srivastava A, Sinha S (2014). Thermostability of in vitro evolved *Bacillus subtilis* lipase A: a network and dynamics perspective. PLoS ONE, 9(8): e102856

Suhre K, Sanejouand Y H (2004). ElNemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement. Nucleic Acids Res, 32(Web Server issue): W610-4

Tian F, Yang C, Wang C, Guo T, Zhou P (2014). Mutatomics analysis of the systematic thermostability profile of *Bacillus subtilis* lipase A. J Mol Model, 20(6): 2257

Tina K G, Bhadra R, Srinivasan N (2007). PIC. Nucleic Acids Res, 35 (Web Server issue): W473-6

Unsworth L D, van der Oost J, Koutsopoulos S (2007). Hyperthermophilic enzymes—stability, activity and implementation strategies for high temperature applications. FEBS J, 274(16): 4044–4056

van Pouderoyen G, Eggert T, Jaeger K E, Dijkstra B W (2001). The crystal structure of *Bacillus subtilis* lipase: a minimal α/β hydrolase fold enzyme. J Mol Biol, 309(1): 215–226

Vogt G, Argos P (1997). Protein thermal stability: hydrogen bonds or internal packing? Fold Des, 2(4): S40–S46

Zeikus J G, Vieille C, Savchenko A (1998). Thermozymes: biotechnology and structure-function relationships. Extremophiles, 2(3): 179–183