Journal of Molecular Graphics and Modelling 28 (2010) 820–827

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Structural study of carboxylesterase from hyperthermophilic bacteria *Geobacillus stearothermophilus* by molecular dynamics simulation

Sangeeta Kundu, Debjani Roy*

Bioinformatics Centre, Bose Institute, Kolkata, India

ARTICLE INFO

Article history:
Received 10 December 2009
Received in revised form 25 February 2010
Accepted 1 March 2010
Available online 6 March 2010

Keywords:
Carboxylesterase
Hyperthermophilic
Molecular dynamics simulation
Radial distribution functions
Principal component analysis

ABSTRACT

Carboxylesterases are ubiquitous enzymes with important physiological, industrial and medical applications such as synthesis and hydrolysis of stereo specific compounds, including the metabolic processing of drugs, and antimicrobial agents. Here, we have performed molecular dynamics simulations of carboxylesterase from hyperthermophilic bacterium *Geobacillus stearothermophilus* (GsEst) for 10 ns each at five different temperatures namely at 300 K, 343 K, 373 K, 473 K and 500 K. Profiles of root mean square fluctuation (RMSF) identify thermostable and thermosensitive regions of GsEst. Unfolding of GsEst initiates at the thermosensitive α-helices and proceeds to the thermostable β-sheets. Five ion-pairs have been identified as critical ion-pairs for thermostability and are maintained stably throughout the higher temperature simulations. A detailed investigation of the active site residues of this enzyme suggests that the geometry of this site is well preserved up to 373 K. Furthermore, the hydrogen bonds between Asp188 and His218 of the active site are stably maintained at higher temperatures imparting stability of this site. Radial distribution functions (RDFs) show similar pattern of solvent ordering and water penetration around active site residues up to 373 K. Principal component analysis suggests that the motion of the entire protein as well as the active site is similar at 300 K, 343 K and 373 K. Our study may help to identify the factors responsible for thermostability of GsEst that may endeavor to design enzymes with enhanced thermostability.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

Biomolecules isolated from thermophilic organisms provide numerous potential biotechnological advantages. Enzymes and proteins isolated from thermophilic microorganisms exhibit a high stability in conditions usually used for denaturing proteins such as high temperature, ionic strength, extreme pH values, elevated concentration of detergents and chaotropic agents [1,2]. As a result they can be used as biocatalysts under harsh environmental conditions [3]. In addition to the potential industrial applications, proteins and enzymes that are stable and active over 100 °C represent good models to shed light on the molecular adaptation of life at high temperature [4,5]. Carboxylesterases are ubiquitous enzymes with important physiological, industrial and medical roles in the synthesis and hydrolysis of stereo specific compounds, including the metabolic processing of drugs, and antimicrobial agents [6–9]. Carboxylesterases (carboxylic ester hydrolase, EC 3.1.1.1) are members of the α/β hydrolase family. They hydrolyze short chain aliphatic and aromatic esters and are inhibited by low concentrations of organophosphorus compounds [6]. The carboxylesterase gene has been isolated from gram-positive thermophilic bacterium *Geobacillus stearothermophilus* (GsEst). This enzyme has been shown to have optimal activity for short chain acyl derivatives of length C4–C8 at a temperature of 70 °C [9]. The crystallization of carboxylesterases of *G. stearothermophilus* has been performed at room temperature and the subsequent three-dimensional structure of GsEst has been elucidated (PDB code: 1TQH) [9]. GsEst containing 242 residues folds into two domains with the catalytic triad located at the interface between the domains. The smaller domain consists of three α-helices B1′ (residues 53–56), D1′ (residues 119–137) and D2′ (residues 141–151) that form a cap over the active site (Fig. 1) The larger domain resembles the classical α/β hydrolase fold. It contains a central seven-stranded β-sheet surrounded by six α-helices αA (residues 25–36), αB (residues 61–78), αC (residues 90–99), αD (residues 159–171), αE (residues 195–202) and αF (residues 226–239) (Fig. 1). The β-sheet of the typical α/β hydrolase fold consists of eight strands with the second β-stand antiparallel to the others. However, the first β-stand of the typical α/β hydrolase fold is absent from this structure. Therefore, the β-sheets of GsEst only have seven strands (β2–β8) with the N-terminal β2 strand antiparallel to the others. Besides, GsEst has two additional 3_{10}-helices (residues 175–177 and residues 220–222).

* Corresponding author at: Bioinformatics Centre, Bose Institute, Acharya J.C. Bose Centenary Building, P-1/12, C.I.T. Scheme – VII M, Kolkata 700054, West Bengal, India. Tel.: +91 33 2355 6626/2816; fax: +91 33 2355 3886.
E-mail addresses: debjani@bic.boseinst.ernet.in, drdebjani@yahoo.com (D. Roy).

1093-3263/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2010.03.001

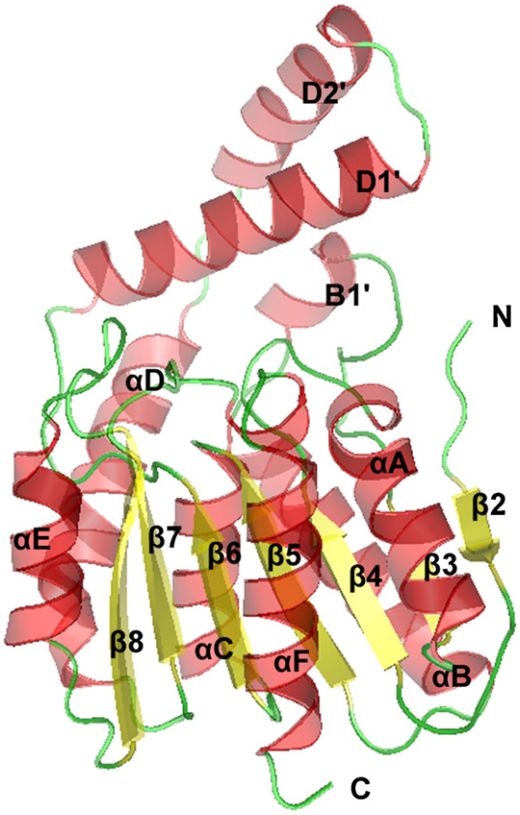

Fig. 1. The native structure of GsEst. Helices are shown as red ribbons, β-strands as yellow arrows and the rest are shown as loops. Six α-helices (A–F), and seven β-strands (2–8) are labeled using the nomenclature for the classical α/β hydrolase fold [38] and the three cap helices are labeled B1′, D1′ and D2′. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

The relationship among rigidity, flexibility, thermostability, and activity has always been of fundamental interest for relating enzyme activity to enzyme stability. The most effective rational way to evaluate the factors governing thermostability of enzymes is molecular dynamics (MD) simulation [10]. Several MD simulation studies have been performed on thermophilic organisms, particularly focusing on the molecular strategies adopted by these enzymes and proteins to work in very harsh environment. For instance, using MD simulations the thermal stable mechanisms of rubredoxin [11], nuclease [12], barnase [13], nitrile hydratase [10], and adenylate kinase [14] have been studied. MD study has also been employed to study the dynamics of lipase [15], xylanase [16] dihydrofolate reductase [17], ribonuclease H1 [18] at high temperatures. So far MD simulation studies of psychrophilic esterase from *Pseudoalteromonas haloplanktis* [19,20], hyperthermophilic esterase from *Archaeoglobus fulgidus* [3] and thermostable para-nitrobenzyl esterase from *Bacillus subtilis* [21] have been reported. However, to date no molecular dynamics simulation study of any hyperthermophilic carboxylesterase has been reported in the literature. In this study MD simulations of GsEst have been carried out at five different temperatures namely 300 K, 343 K, 373 K, 473 K and 500 K in order to explore the key factors governing thermostabilty of this enzyme. By calculating the root mean square deviation (RMSD) and root mean square fluctuation (RMSF) values for backbone and Cα atoms, the thermal sensitive regions have been identified. Salt bridge interactions have also been analyzed. The dynamic properties of GsEst at different temperatures have been compared in terms of secondary structure content, molecular flexibility, intramolecular hydrogen bonds and protein–solvent interactions. Comparison of essential conformational subspaces of this protein at different temperatures has been monitored by principal component analysis. A detailed investigation of the residues of the catalytic triad has also been performed.

2. Methodology

2.1. Molecular dynamics simulation

All molecular dynamics (MD) simulations were performed using the GROMACS 3.3.1 [22,23] package and GROMOS96 [24] 43a1 force field implemented on a LINUX architecture. The starting configuration for hyperthermophilic carboxylesterase was obtained from *G. stearothermophilus* (PDB code: 1TQH [9]) which consisted of 242 residues. We abbreviated 1TQH as GsEst throughout the text. Crystallographic water molecules and heteroatoms were removed from the system. The starting structure was immersed in a triclinic box of SPC water molecules [25]. To neutralize the charge of the system 14 NA^+ ions were added to each system. The resulting system sizes are listed in Table 1. Each system was subjected to energy minimization for 2000 steps by steepest descents. The minimized systems for GsEst were equilibrated for 100 ps each at five different temperatures namely 300 K, 343 K, 373 K, 473 K and 500 K by position restrained molecular dynamics simulation in order to relax the solvent. The equilibrated systems were then subjected to molecular dynamics simulations for 10 ns each at five different temperatures. Periodic boundary conditions combined with minimum image convention were used under isothermal, isobaric conditions using Berendsen coupling algorithm [26] with relaxation times of 0.1 ps and 0.5 ps respectively. The LINCS algorithm [27] was used to constrain bond lengths using a time step of 2 fs for all calculations. Electrostatic interactions were calculated using the Particle Mesh Ewald (PME) [28,29] summation scheme. van der Waals and Coulomb interactions were truncated at 0.9 nm. The non-bonded pair list was updated every 10 steps and conformations were stored every 2 ps. Secondary structure analysis was performed using the program DSSP [30]. Other analyses were performed using scripts included with the Gromacs [31] distribution. The visual analysis of protein structures was carried out using Rasmol [32] and Pymol [33].

2.2. Essential dynamics

Essential degrees of freedom of GsEst were extracted from equilibrated portion of the trajectories according to principal component analysis (PCA) or essential dynamics (ED) method [34,35]. The ED method is based on the construction of the covariance matrix of the coordinate fluctuations. The covariance matrix is diagonalized to obtain the eigenvectors and

Table 1
Setup of MD simulations of GsEst.

Temperature	300 K	343 K	373 K	473 K	500 K
Protein atoms	2481	2481	2481	2481	2481
Counterions (Na^+)	14	14	14	14	14
Water	7723	7723	7723	8518	8518
Total atoms	25664	25664	25664	28049	28049
Box size (nm^3)	5.86 × 6.25 × 7.38	5.86 × 6.25 × 7.38	5.86 × 6.25 × 7.38	6.06 × 6.45 × 7.58	6.06 × 6.45 × 7.58

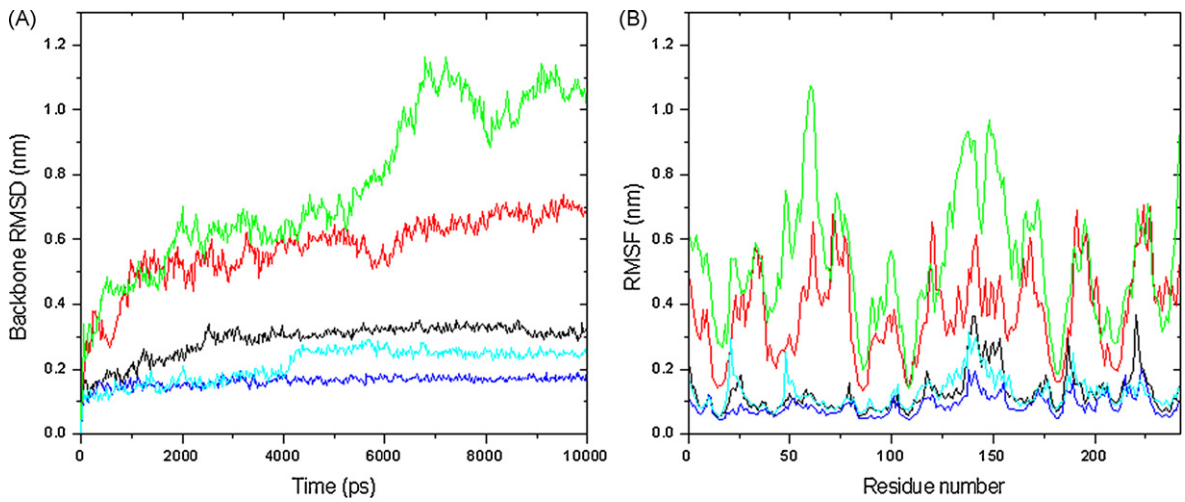

Fig. 2. (A) Time evolution of backbone RMSD of GsEst at different temperatures. (B) RMSFs as a function of residue number of GsEst at different temperatures. The color-coding scheme is as follows: 300 K (blue), 343 K (cyan), 373 K (black), 473 K (red) and 500 K (green). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Table 2
Average values of Rg, SASA and average number of protein–protein (intramolecular) and protein–solvent H-bonds in GsEst at different temperature simulations.

	300 K	343 K	373 K	473 K	500 K
Backbone RMSD (nm)	0.17	0.21	0.31	0.63	0.80
[a]Rg (nm)	1.77	1.76	1.75	1.78	1.81
[a]SASA (nm^2)	120.72	119.16	117.56	126.36	138.06
PSASA (nm^2)	63.99	62.01	60.49	71.13	80.48
ASASA (nm^2)	56.51	57.14	56.66	56.57	57.57
Protein–protein H-bond	189	191	171.5	154	135
Protein–solvent H-bond	459.5	418.5	409.5	315	291

[a] Rg, SASA and H-bonds denote radius of gyration, solvent accessible surface area and hydrogen bonds respectively. PSASA and ASASA denote polar and apolar solvent accessible surface area respectively.

eigenvalues that provide information about correlated motions throughout the protein. The eigenvectors represent the directions of motion, and the eigenvalues represent the amount of motion along each eigenvector. The eigenvectors are then sorted according to their eigenvalues in descending order. Usually, the first 10–20 eigenvectors suffice to describe almost all conformational substates accessible to the protein, only Cα atoms were included in the definition of the covariance matrices for the enzyme.

The root-mean-square inner product (RMSIP) [36,37] was calculated to measure the degree of overlap between the conformational spaces of GsEst explored by the simulation at different temperatures. The RMSIP is defined as follows:

$$\mathrm{RMSIP} = \sqrt{\frac{1}{10}\sum_{i=1}^{10}\sum_{j=1}^{10}(\eta_i^a \cdot \eta_j^b)^2}$$

Table 3
Average secondary structure contents in GsEst trajectories obtained at different temperatures and in the corresponding starting structures.

	GsEst					
	Starting	300 K	343 K	373 K	473 K	500 K
Coil	59[24.38%]	46.5(0.70) [19.21%]	44.5(2.12) [18.38%]	48(4.24) [19.83%]	71(38.18) [29.33%]	70.5(36.06) [29.13%]
β-Sheet	34[14.04%]	36(1.41) [14.87%]	34(0) [14.04%]	34(1.41) [14.04%]	33.5(0.70) [13.84%]	43(12.72) [17.76%]
β-Bridge	0[0%]	1(1.41) [0.41%]	2(2.82) [0.82%]	0(0) [0%]	3.5(4.94) [1.44%]	4(5.65) [1.65%]
Bend	21[8.67%]	20.5(2.12) [8.47%]	24.5(2.12) [10.12%]	26.5(9.19) [10.95%]	35.5(19.09) [14.66%]	38.5(26.16) [15.9%]
Turn	27[11.15%]	22(1.41) [9.09%]	20.5(4.94) [8.47%]	25.5(2.12) [10.53%]	34(11.31) [14.04%]	23.5(4.94) [9.71%]
α-Helix	109[45.02%]	108(4.24) [44.62%]	111.5(9.19) [46.07%]	105(5.65) [43.38%]	61.5(68.58) [25.41%]	56.5(71.41) [23.34%]
3_{10}-Helix	6[2.47%]	5.5(2.1) [2.27%]	5(7.07) [2.06%]	3(4.24) [1.23%]	1.5(2.12) [0.61%]	0(0) [0%]

Standard deviations and percentages are given in parenthesis and square brackets, respectively.

Table 4
Important salt bridge interactions of GsEst at 300 K, 343 K and 373 K simulation and their occupancy time.

Salt bridge	300 K		343 K		373 K	
	Occupancy time (%)	Distance (nm)[a]	Occupancy time (%)	Distance (nm)[a]	Occupancy time (%)	Distance (nm)[a]
D232–K211	100	0.37	99	0.34	98	0.38
E105–K82	94	0.35	84	0.45	78	0.42
E54–K134	87	0.39	83	0.33	73	0.41
E238–R12	81	0.54	73	0.46	62	0.48
E147–K134	85	0.39	79	0.29	51	0.50
E147–K139	88	0.45	73	0.39	–	–

[a] The average distance of MD simulation.
Only ion-pairs with a distance less than 0.5 nm are reported.
Occupancy time is the percentage of structures where the salt bridge distance is less than 0.5 nm. Only ion-pairs with time occupancy ≥50% are reported.

where η_i^a, η_j^b are the ith and jth eigenvectors from two different sets a and b respectively. Sub-matrices of the overall covariance matrices that only involved segments surrounding the active site were also subjected to PCA and RMSIP analyses. The RMSIP was computed from the equilibrated portion of the trajectory. For diagonal elements, the RMSIP was computed by splitting the equilibrated portion of the trajectory into two equal halves.

3. Results and discussion

3.1. Global structural stability

In Fig. 2A, the backbone RMSDs of GsEst from the corresponding starting structure are shown as a function of time at five different temperatures. RMS deviations from initial structure are stable during the entire simulation at 300 K and 343 K and converged to values about 0.17 nm and 0.21 nm respectively showing a close resemblance to the starting structure. In the trajectory run at 373 K backbone RMSD of GsEst reaches equilibrium at about 2.5 ns and attains a value of 0.3 nm, RMSD remains the same afterwards for rest of the simulation. At 473 K and 500 K simulations, backbone RMSDs increase from the beginning of the simulation and attain a very high values of 0.6 nm and 1 nm respectively at the end of the simulation. Thus for GsEst, increasing temperature up to 373 K shows very little effect on the stability of its native structure. Significant structural distortions are observed only at higher temperatures.

Radius of gyration (Rg) does not show significant change at different temperatures for GsEst (Table 2). Solvent accessible surface area (SASA) shows that no significant conformational change occurs during the simulations up to 373 K. At 473 K and 500 K, change of SASA is more pronounced. Rise of SASA is mostly contributed by the hydrophobic residues (Table 2). This is indicative of an increase in solvent penetration into the internal core of the protein.

3.2. Secondary structure

Table 3 reports the average secondary structure content in GsEst at different temperature simulations, as well as in the corresponding starting structure according to DSSP classification. All the β-sheets, α-helices observed within the hydrolase fold of the crystal structure of GsEst are well maintained throughout the whole simulation period at 300 K, 343 K and 373 K. At higher temperatures (473 K and 500 K) α-helices of the hydrolase fold gradually disorder accompanied by denaturation of original β-sheets of the hydrolase fold. Therefore a significant decrease in α-helix content is observed at higher temperatures. Unfolding process generates new sheets, originating from areas that were initially loops and α-helices, so altogether β-sheet content does not alter significantly at higher temperatures. In general it has been observed that the β-sheets forming the core of the hydrolase fold are more stable than α-helices which encompass the core region. 3_{10}-Helices are well maintained up to 343 K, only one 3_{10}-helix (residues 173–175) is left at 373 K. Beyond 343 K simulation 3_{10}-helices are not maintained at all.

The starting structure of GsEst contains 227 intramolecular hydrogen bonds (H-bonds). The average number of intramolecular hydrogen bonds is calculated at different temperature simulations for GsEst (Table 2). The network of hydrogen bonds is well preserved at 300 K and 343 K. With increase in temperature, there is a concomitant decrease in number of hydrogen bonds holding the secondary structures together. Average numbers of protein–solvent interactions also decrease with rise in temperature. Effect of loss of intermolecular hydrogen bonds and protein–solvent interactions is more significant at higher temperature (473 K and 500 K) simulations (Table 2).

3.3. Structural flexibility

The average RMSF values in the MD simulation are considered as the criterion for overall flexibility of the system. Fig. 2B compares the root mean square fluctuations (RMSFs) of Cα atoms on a residue basis for every temperature simulations for GsEst. The RMSFs derived from the simulation at 300 K agree well with the atomic mobility of the crystal structures as derived from B-factors. At 300 K and 343 K, most regions of GsEst show slight fluctuation with increase in temperature, indicating protein is relatively thermostable at these temperatures. At 373 K residues of helices D1′, D2′ and interconnecting loop (residues 129–158) show major RMSF fluctuation. In addition to these residues showing major fluctuations, loops between β3 and αA (residues 20–24), β4 and B1′ (residues 47–52), β7 and αE (residues 186–202), and N-terminus and loop region connecting β8 and αF (residues 217–226) show minor fluctuations. When the temperature is elevated much more (at 473 K and 500 K) most of the residues become highly mobile due to the loss of native secondary structure at high temperatures. At elevated temperatures, residues involving mainly helices of the hydrolase fold, β2, helices of the cap (1–10, 22–38, 50–80, 118–126, 132–155, 162–175, 189–202, and 216–233) show major fluctuation while the regions involving β-sheets of the hydrolase fold (10–21, 39–53, 80–90, 104–114, 178–187 and 203–214) show less fluctuation at these temperatures.

3.4. Profile of salt bridge interaction

Electrostatic interactions of ion-pairs in thermophilic enzyme may contribute to its increased stability at higher temperatures. Among many ion-pairs found in GsEst with 0.5 nm distance cutoff, only five ion-pairs are identified as critical ion-pairs for thermostability of GsEst. These ion-pairs are monitored throughout the simulation at all five temperatures. At 300 K and 343 K, ion-pairs Asp232–Lys211, Glu105–Lys82, Glu54–Lys134, Glu238–Arg12, Glu147–Lys134, and Glu147–Lys139 are well main-

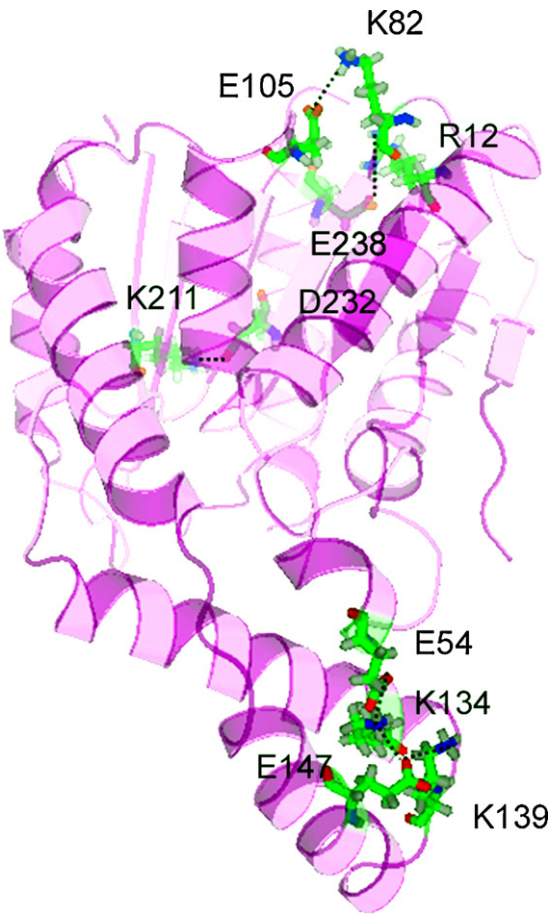

Fig. 3. Important salt bridge interactions involving GsEst. Dashed lines indicate salt bridge interactions. Red, oxygen atom; blue, nitrogen atom; green, carbon atom; white, hydrogen atom.

tained throughout the simulation (Table 4). These ion-pairs of GsEst are shown in Fig. 3. At 373 K salt bridge Glu147–Lys139 gets totally disrupted however salt bridge Glu147–Lys134 is maintained only for 50% of the simulation. These two salt bridges include helices D1′ and D2′ of the cap region which has previously been identified by high RMSFs and is responsible for initial heat denaturation. This cap region adjacent to the active site of the enzyme forms the substrate binding pocket [20]. Stabilizing this region would be effective to improve the global thermal stability.

Table 5
RMSF flexibility of active site residues.

GsEst	300 K	343 K	373 K	473 K	500 K
Ser89	0.06	0.11	0.08	0.25	0.31
Asp188	0.04	0.15	0.20	0.35	0.42
His218	0.06	0.14	0.19	0.42	0.47

3.5. Catalytic triad

In GsEst active site cleft is formed by six loops at the C-terminal ends of the predominantly parallel β-strands. The catalytic triad consists of residues Ser89, Asp188 and His218. Ser89 of GsEst is located in the conserved sequence Gly-X-Ser-X-Gly. A network of hydrogen bonds stabilizes the catalytic triad. In fact, the well-known mechanism of the catalytic triad requires the formation of an interaction between ND1 of His and the caboxylate moiety of Asp to ensure the right charge state of the imidazole ring: if this H-bond is not stable, the enzyme will be less active [20]. In the catalytic triad the OD1 and OD2 atoms of Asp188 of GsEst form bifurcated hydrogen bonds with the ND1 atom in His218. Ser89 forms hydrogen bond with side chain of His218. These hydrogen bonds involving the catalytic residues have been monitored throughout the simulations at five different temperatures. In GsEst Asp188 and His218 hydrogen bond is stably maintained even at very high temperature (up to 500 K). The hydrogen bond between Ser89 and side chain of His218 is maintained only up to 343 K. Cα–Cα distances between the catalytic residues have also been monitored. The Cα–Cα distance between the catalytic residues Asp188–His218, Ser89–Asp188 and His218–Ser89 of the crystal structure are 0.50 nm, 1.09 nm and 0.77 nm respectively. These distances are stably maintained up to 373 K simulation and the geometry of the catalytic site is well preserved up to this temperature. RMSFs of three catalytic residues namely Ser89, Asp188 and His218 show considerably higher flexibility at 343 K and 373 K than 300 K (Table 5). This observation suggests that the gain in flexibility of these residues might favor the entry of the substrates and/or catalysis [20]. Whereas at low temperature (300 K), the active site of the protein shows rigid conformation and the catalysis slows. Similar trend has also been observed for the segments of the substrate binding cap domain.

Radial distribution function (RDF) allows us to obtain insights into the extent of water penetration and solvent ordering around

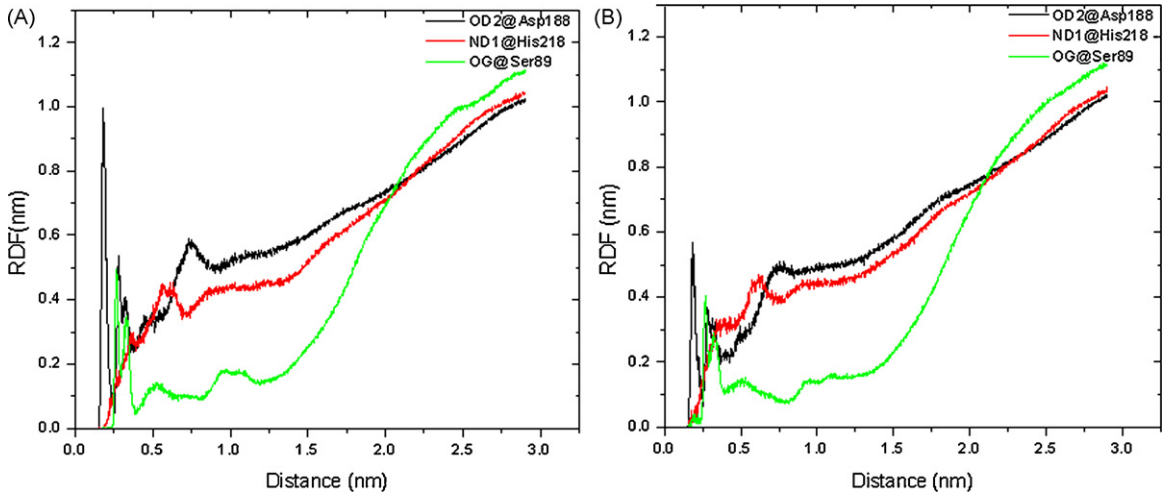

Fig. 4. Plot of the pair distribution function for OD2@Asp188, ND1@His218 and OG@Ser89 to the oxygens in the surrounding water molecules for GsEst at (A) 300 K and (B) 343 K.

Table 6
RMSIP between the first ten eigenvectors for trajectories obtained at different temperatures for GsEst.

GsEst	300 K	343 K	373 K	473 K	500 K
300 K	0.644[a]	0.630	0.584	0.484	0.471
343 K		0.627[a]	0.600	0.469	0.444
373 K			0.583[a]	0.447	0.440
473 K				0.464[a]	0.443
500 K					0.420[a]

[a] These values refer to the RMSIP obtained by comparing two halves of the equilibrated portion of the corresponding trajectories.

the atoms of interest in the active site. Fig. 4 shows RDFs of OG atom of Ser89, OD2 of Asp188 and ND1 of His218 at 300 K and 343 K (other figures are not shown).

3.6. Correlated structure variation

The overall concerted motions of GsEst at all five temperatures are investigated by the essential dynamics analysis. Table 6 reports the comparison of the essential subspaces of the proteins. We have calculated the root-mean-square inner product (RMSIP) between the first 10 eigenvectors of each simulation. This number accounts for the similarity/dissimilarity of the essential motions. For reference values, we have divided each trajectory in two halves and have calculated the inner product between the first 10 essential eigenvectors of each half. The values obtained represent the expected reference values that indicate a similarity in the motions and are taken as thresholds for a comparison among different trajectories. The results show that essential subspaces of the protein are significantly overlapped at 300 K, 343 K and 373 K. On the contrary, the motions of GsEst are significantly different when compared at higher simulated temperature.

The fluctuations of Cα atoms of GsEst projected on first principal component at 300 K, 343 K, 373 K, 473 K and 500 K are shown in Fig. 5. The primary internal motion of the protein is similar to corresponding RMSF plots of Cα atoms shown in Fig. 2B.

The overall flexibility of GsEst has been calculated by the traces of the diagonalized covariance matrix of the atomic positional fluctuations. We have obtained the following values for GsEst: 2.114 nm^2 at 300 K, 4.337 nm^2 at 343 K, 4.740 nm^2 at 373 K, 38.778 nm^2 at 473 K and 76.577 nm^2 at 500 K. This result indicates that GsEst shows similar flexibility at 343 K and 373 K.

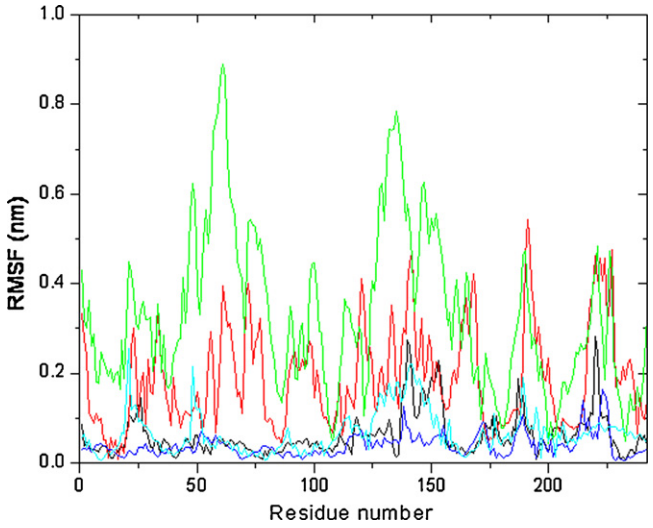

Fig. 5. RMSFs of Cα atoms projected along PC1 of GsEst at different temperatures. The color-coding scheme is as follows: 300 K (blue), 343 K (cyan), 373 K (black), 473 K (red) and 500 K (green). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Fig. 6 illustrates the motion along the first principal component (PC1) of GsEst at 300 K, 343 K and 373 K. PC1 accounts for ~25% internal motion for GsEst at 300 K and ~40% internal motion at 343 K and 373 K. For, GsEst, this mode mainly involves the correlated variations of the region encompassing helices of the cap domain and the loops connecting β8 and αF, and β7 and αE. The secondary structure components forming the core of GsEst are very rigid along this mode. However the rigidity of the core region loses to some extent at 373 K, at higher temperature loss is much severe.

To analyze correlated variations localized within the active site, we have constructed covariance matrices including only regions

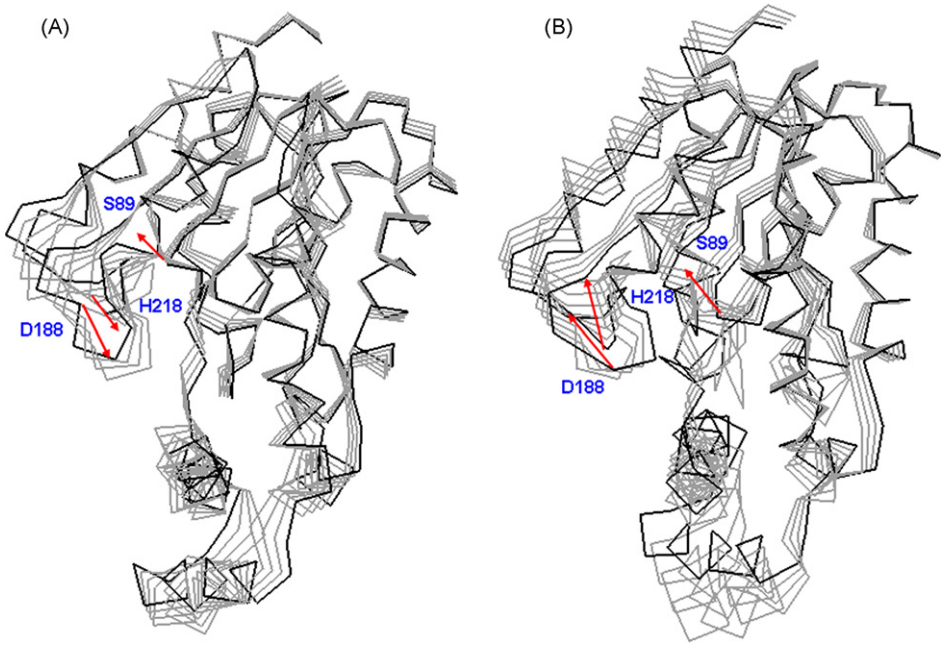

Fig. 6. Superimposition of 4 configurations obtained by projecting the Cα motion onto the first eigenvector at (A) 300 K and (B) 343 K of GsEst. Arrows are used as qualitative indicators of the motion direction.

Table 7
RMSIP between the first ten eigenvectors for trajectories obtained at different temperatures involving regions surrounding the active center.

GsEst	300 K	343 K	373 K	473 K	500 K
300 K	0.8848[a]	0.8910	0.8240	0.7536	0.7602
343 K		0.8689[a]	0.8420	0.7628	0.7674
373 K			0.7739[a]	0.7443	0.7886
473 K				0.7628[a]	0.7803
500 K					0.7293[a]

[a] These values refer to the RMSIP obtained by comparing two halves of the equilibrated portion of the corresponding trajectories.

surrounding the active center (residues 87–91, 186–190, and 216–220) and calculated RMSIPs (Table 7). The resulting principal components have remarkable degrees of overlap with each other suggesting the essential subspaces of the catalytic site of the protein are largely overlapped up to 373 K. The degree of overlap decreases at higher temperatures (473 K onwards). Although the overlap values are higher than the overlap value obtained for the entire protein inflicting distortion of the active site of the enzyme is less compared to the entire protein at higher temperatures.

The internal motion of three catalytic site residues has also been studied. We have found that internal motion of catalytic Ser89 is similar at 300 K, 343 K and 373 K. The relative direction of H218 and D188 of GsEst is also alike at these temperatures (Fig. 6). At higher temperatures (at 473 K), the motion as well as the relative positions of active site fragments are quite altered, the alteration becomes more pronounced at 500 K (figure not shown).

4. Conclusion

In this study, MD simulations have been employed to study thermodynamic behavior of GsEst at different temperatures. At 300 K, 343 K, 373 K GsEst is stable and maintain its global three-dimensional structure throughout the simulation. Although the hydrogen bond between catalytic Ser89–His218 is impaired at 373 K the active site geometry remains unaltered up to this temperature as revealed by measuring Cα–Cα distance between the catalytic residues (Asp188–His218, Ser89–Asp188 and His218–Ser89) throughout the simulation at 300 K, 343 K, 373 K. The internal motion of the catalytic site fragments is also similar at these temperatures.

Five ion-pairs have been identified (D232–K211, E105–K82, E54–K134, E238–R12 and E147–K134) as critical ion-pairs. These ion-pairs are of fundamental importance in maintaining the structural integrity of the protein structure. Among these five pairs, one ion-pair (D232–K211) has been found to be extremely stable throughout the simulation up to 373 K. D232–K211 ion-pair between αF and β8 adds more stability towards thermostable β-sheet core region.

The overall flexibility of the entire protein calculated by the trace of the diagonalized covariance matrix displays similar flexibility of GsEst at 343 K and 373 K. This indicates that protein may have similar activities at these two temperatures. However based on RMSF calculations it has been found that certain regions of GsEst show higher flexibility namely D1′, D2′ and interconnecting loop (residues 129–158) at 373 K. This may be due to the breakdown of Glu147–Lys139 ion-pair of the substrate binding cap domain. This is the most thermosensitive fragment of GsEst and denaturation initiates from this region.

RMSFs of three catalytic residues namely Ser89, Asp188 and His218 show considerably higher flexibility at 343 K and 373 K than 300 K. This observation suggests that the gain in flexibility of these residues might favor the entry of the substrates and/or catalysis. Whereas at low temperature (300 K), the active site of the protein shows rigid conformation and the catalysis slows. Similar trend has also been observed for the segments of the substrate binding cap domain.

This study is the first attempt to observe the dynamic behavior of thermophilic carboxylesterase at high temperatures and to understand the factors conferring thermostability of this protein. The findings of our study may help to design biotechnologically improved thermostable proteins.

Acknowledgement

This study was supported by grants from Department of Biotechnology Govt. of India.

References

[1] P. Herman, M. Staiano, A. Marabotti, A. Varriale, A. Scirè, F. Tanfani, J. Vecer, M. Rossi, S. D'Auria, D-Trehalose/D-maltose-binding protein from the hyperthermophilic archaeon *Thermococcus litoralis*: the binding of trehalose and maltose results in different protein conformational states, Proteins 63 (2006) 754–767.
[2] S. D'Auria, R. Nucci, M. Rossi, I. Gryczynski, Z. Gryczynski, J.R. Lakowicz, The β-glycosidase from the hyperthermophilic archaeon *Sulfolobus solfataricus*: enzyme activity and conformational dynamics at temperatures above 100 °C, Biophys. Chem. 81 (1999) 23–31.
[3] S. D'Auria, P. Herman, J.R. Lakowicz, E. Bertoli, F. Tanfani, M. Rossi, G. Manco, The thermophilic esterase from *Archaeoglobus fulgidus*: structure and conformational dynamics at high temperature, Proteins: Struct. Funct. Genet. 38 (2000) 351–360.
[4] D. Fessas, M. Staiano, A. Barbiroli, A. Marabotti, A. Schiraldi, A. Varriale, M. Rossi, S. D'Auria, Molecular adaptation strategies to high temperature and thermal denaturation mechanism of the D-trehalose/D-maltose-binding protein from the hyperthermophilic archaeon *Thermococcus litoralis*, Proteins 67 (2007) 1002–1009.
[5] S. D'Auria, I.M. Moracc, F. Febbraio, F. Tanfani, R. Nucci, M. Rossi, Structure–function studies on β-glycosidase from *Sulfolobus solfataricus*. Molecular bases of thermostability, Biochimie 80 (1998) 949–957.
[6] K. Krish, Carboxylic ester hydrolysis, in: P.D. Boyer (Ed.), The Enzymes, 5, Academic Press, New York, 1971, pp. 43–49.
[7] T. Satoh, M. Hosokawa, The mammalian carboxylesterases: from molecules to functions, Annu. Rev. Pharmacol. Toxicol. 38 (1998) 257–288.
[8] U.T. Bornscheuer, Methods to increase enantioselectivity of lipases and esterases, Curr. Opin. Biotechnol. 13 (2002) 543–547.
[9] P. Liu, Y.F. Wang, H.E. Ewis, A.T. Abdelal, C.D. Lu, R.W. Harrison, I.T. Weber, Covalent reaction intermediate revealed in crystal structure of the *Geobacillus stearothermophilus* carboxylesterase Est30, J. Mol. Biol. 342 (2004) 551–561.
[10] J. Liu, H. Yu, Z. Shen, Insights into thermal stability of thermophilic nitrile hydratases by molecular dynamics simulation, J. Mol. Graph. Model. 27 (2008) 529–535.
[11] A. Grottesi, M.A. Ceruso, A. Colosimo, A. Di Nola, Molecular dynamics study of a hyperthermophilic and a mesophilic rubredoxin, Proteins: Struct. Funct. Genet. 46 (2002) 287–294.
[12] A.D. Gruia, S. Fischer, J.C. Smith, Molecular dynamics simulation reveals a surface salt bridge forming a kinetic trap in unfolding of truncated *Staphylococcal nuclease*, Proteins: Struct. Funct. Genet. 50 (2003) 507–515.
[13] J. Yin, D. Bowen, W.M. Southerland, Barnase thermal titration via molecular dynamics simulations: detection of early denaturation sites, J. Mol. Graph. Model. 24 (2006) 233–243.
[14] E. Bae, G.N. Phillips Jr., Identifying and engineering ion pairs in adenylate kinases Insights from molecular dynamics simulations of thermophilic and mesophilic homologues, J. Biol. Chem. 280 (2005) 30943–30948.
[15] R. Abedi Karjiban, M.B. Abdul Rahman, M. Basri, A.B. Salleh, D. Jacobs, H. Abdul Wahab, Molecular dynamics study of the structure, flexibility and dynamics of thermostable L1 lipase at high temperatures, Protein J. 28 (2009) 14–23.
[16] M. Purmonen, J. Valjakka, K. Takkinen, T. Laitinen, J. Rouvinen, Molecular dynamics studies on the thermostability of family 11 xylanases, Protein Eng. Des. Sel. 20 (2007) 551–559.
[17] J. Pang, R.K. Allemann, Molecular dynamics simulation of thermal unfolding of *Thermatoga maritima* DHFR, Phys. Chem. Chem. Phys. 9 (2007) 711–718.
[18] L. Tang, H. Liu, A comparative molecular dynamics study of thermophilic and mesophilic ribonuclease HI enzymes, J. Biomol. Struct. Dyn. 24 (2007) 379–392.
[19] S. D'Auria, V. Aurilia, A. Marabotti, M. Gonnelli, G. Strambini, Structure and dynamics of cold-adapted enzymes as investigated by phosphorescence spectroscopy and molecular dynamics studies 2. The case of an esterase from *Pseudoalteromonas haloplanktis*, J. Phys. Chem. B 113 (2009) 13171–13178.
[20] V. Aurilia, J.F. Rioux-Dubé, A. Marabotti, M. Pézolet, S. D'Auria, Structure and dynamics of cold-adapted enzymes as investigated by FT-IR spectroscopy and MD. The case of an esterase from *Pseudoalteromonas haloplanktis*, J. Phys. Chem. B 113 (2009) 7753–7761.
[21] P.L. Wintrode, D. Zhang, N. Vaidehi, F.H. Arnold, W.A. Goddard 3rd, Protein dynamics in a family of laboratory evolved thermophilic enzymes, J. Mol. Biol. 327 (2003) 745–757.

[22] E. Lindahl, B. Hess, D. van der Spoel, Gromacs 3.0: a package for molecular simulation and trajectory analysis, J. Mol. Model. 7 (2001) 306–317.
[23] D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J. Berendsen, GROMACS: fast, flexible, and free, J. Comp. Chem. 26 (2005) 1701–1718.
[24] W.F. van Gunsteren, S.R. Billeter, A.A. Eising, P.H. Hunenberger, P. Kruger, A.E. Mark, W.R.P. Scott, I.G. Tironi, Biomolecular Simulation: The Gromos 96 Manual and User Guide, Hochschulverlag AG an der Zurich, Zurich, Switzerland, 1996.
[25] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, J. Hermans, Interaction models for water in relation to protein hydration, in: B. Pullman (Ed.), Intermolecular Forces, D Reidel Publishing Company, Dordrecht, 1981, pp. 331–342.
[26] H.J.C. Berendsen, J.P.M. Postma, A. DiNola, J.R. Hakk, Molecular dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690.
[27] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: a linear constraint solver for molecular simulations, J. Comp. Chem. 18 (1997) 1463–1472.
[28] T. Darden, D. York, L. Pedersen, Particle mesh ewald: an $N - \log(N)$ method for ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10093.
[29] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pederson, A smooth particle meshes ewald potential, J. Chem. Phys. 103 (1995) 8577–8592.
[30] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (1983) 2577–2637.
[31] D. van der Spoel, E. Lindahl, B. Hess, A.R. van Buuren, E. Apol, P.J. Meulenhoff, D.P. Tieleman, A.L.T.M. Sijbers, K.A. Feenstra, R. van Drunen, H.J.C. Berendsen, Gromacs User Manual Version 3.3, 2005, Internet: www.gromacs.org.
[32] R.A. Sayle, E.J. Millner-White, Rasmol-biomolecular graphics for all, Trends. Biochem. Sci. 20 (1995) 374–376.
[33] W.L. DeLano, The PyMOL Molecular Graphics System, 2002, Internet http://www.pymol.org.
[34] A. Amadei, A.B. Linssen, H.J.C. Berendsen, Essential dynamics of proteins, Proteins 17 (1993) 412–425.
[35] A. Amadei, M.A. Ceruso, A. Di Nola, On the convergence of the conformational coordinates basis set obtained by the essential dynamics analysis of proteins' molecular dynamics simulations, Proteins 36 (1999) 419–424.
[36] B.L. de Groot, D.M.F. van Aalten, A. Amadei, H.J.C. Berendsen, The consistency of large concerted motions in proteins in molecular dynamics simulations, Biophys. J. 71 (1996) 1707–1713.
[37] D.M.F. van Aalten, B.L. de Groot, J.B.C. Findlay, H.J.C. Berendsen, A. Amadei, A comparison of techniques for calculating protein essential dynamics, J. Comp. Chem. 18 (1997) 169–181.
[38] D.L. Ollis, E. Cheah, M. Cygler, B. Dijkstra, F. Frolow, S.M. Franken, The alpha/beta hydrolase fold, Protein Eng. 5 (1992) 197–211.