Journal of Molecular Graphics and Modelling 27 (2009) 871–880

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

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

Comparative structural studies of psychrophilic and mesophilic protein homologues by molecular dynamics simulation

Sangeeta Kundu, Debjani Roy *

Bioinformatics Centre, Bose Institute, Kolkata, India

ARTICLE INFO

Article history:
Received 27 November 2008
Received in revised form 14 January 2009
Accepted 15 January 2009
Available online 23 January 2009

Keywords:
Psychrophilic
Mesophilic
Molecular dynamics simulation
Protein unfolding
Principal component analysis

ABSTRACT

Comparative molecular dynamics simulations of psychrophilic type III antifreeze protein from the North-Atlantic ocean-pout *Macrozoarces americanus* and its corresponding mesophilic counterpart, the antifreeze-like domain of human sialic acid synthase, have been performed for 10 ns each at five different temperatures. Analyses of trajectories in terms of secondary structure content, solvent accessibility, intramolecular hydrogen bonds and protein–solvent interactions indicate distinct differences in these two proteins. The two proteins also follow dissimilar unfolding pathways. The overall flexibility calculated by the trace of the diagonalized covariance matrix displays similar flexibility of both the proteins near their growth temperatures. However at higher temperatures psychrophilic protein shows increased overall flexibility than its mesophilic counterpart. Principal component analysis also indicates that the essential subspaces explored by the simulations of two proteins at different temperatures are non-overlapping and they show significantly different directions of motion. However, there are significant overlaps within the trajectories and similar directions of motion of each protein especially at 298 K, 310 K and 373 K. Overall, the psychrophilic protein leads to increased conformational sampling of the phase space than its mesophilic counterpart.

Our study may help in elucidating the molecular basis of thermostability of homologous proteins from two organisms living at different temperature conditions. Such an understanding is required for designing efficient proteins with characteristics for a particular application at desired working temperatures.

1. Introduction

There has been a growing interest in understanding the molecular basis of thermostability of proteins from the organisms living at different temperature conditions. Such an understanding is critical for designing efficient enzymes with characteristics for a particular application. Homologous proteins can be found in organisms that live in very different environments. Such homologous proteins may be highly similar in their sequences, structures but drastically different in the temperature dependencies of their activity and stability. Psychrophilic organisms live at low temperature, where most other organisms cannot grow. Type III antifreeze proteins (AFP III) are very small (~65 amino acids long) psychrophilic proteins that are seasonally found at high concentrations in the blood of some fishes living in polar or high latitude temperate seas [1]. AFPs depress the freezing point of blood and body fluids below that of the surrounding seawater by binding to and inhibiting the growth of seed ice crystals [2,3]. Because of their lack of enzymatic activity, study of AFPs may allow to differentiate between features of cold-adapted proteins that may have arisen as consequence of structural features, such as flexibility, that are needed for catalysis [1]. The mesophilic homologue of AFP III is the C-terminal domain of human sialic acid synthase, called the antifreeze like (AFL) domain [4], since it is similar to those observed in a variety of functional type III antifreeze proteins. It has been proposed that the AFL domain is also involved in sugar binding [4], but the details of its function have remained elusive. Recently the crystal structure of AFP III containing 66 residues from *Macrozoarces americanus* [5] has been solved. The AFL domain of human sialic acid synthase consisting of 79 residues has also been determined by NMR spectroscopy [4]. These two proteins show 38% sequence identity (Fig. 1A) and the root-mean-square deviation of Cα atoms of superimposed structures is 2.0 Å. Their overall fold contains one α-helix (residues 37–40), two 3_{10} helices (residues 19–21 and 57–59) and two β-strands (residues 4–7 and 22–25) (Fig. 1B). In addition the AFP III

* 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 2334 3886.
E-mail addresses: debjani@bic.boseinst.ernet.in, drdebjani@yahoo.com (D. Roy).

1093-3263/$ – see front matter © 2009 Published by Elsevier Inc.
doi:10.1016/j.jmgm.2009.01.004

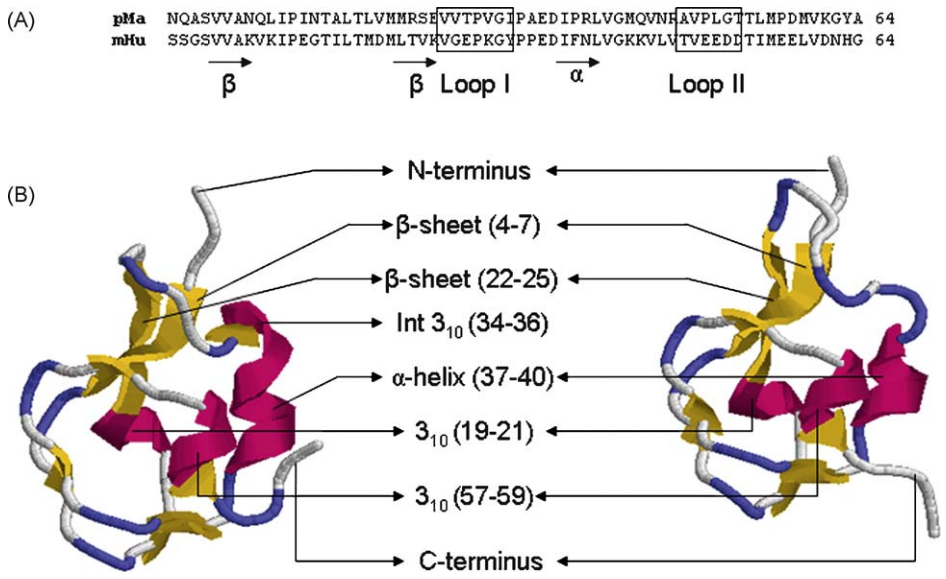

Fig. 1. (A) Pairwise sequence alignment of pMa and mHu. Sequences of loops I and II are shown in box. (B) The native structure of (I) pMa and (II) mHu. Helices are shown as magenta ribbons, β-strands as yellow arrows and the rest are shown as loops.

has an intermediate 3_{10} helix (residues 34–36). The rest of the structures comprise mostly of β-bridges and loops. Despite of very similar structure, they have very different temperature dependencies in stability and activity. The working temperature of AFP III protein is around 273 K (0 °C) while the working temperature of mesophilic proteins is in the range of 293–310 K (20–37 °C) [1]. Psychrophilic proteins have been found to be much more temperature liable than mesophilic counterpart [6] indicating that psychrophilic proteins are generally less stable compared to mesophilic homologue [6]. Until recently, the absence of available crystal structures of cold-adapted proteins limited the comparison to their homologues from meso and thermophilic organisms. Molecular dynamics (MD) simulation is a suitable tool to evaluate the comparative basis of protein thermostability between homologous psychrophilic and mesophilic proteins. However only a few comparative molecular dynamics studies aimed at evaluating temperature dependencies of the dynamics of psychrophilic and mesophilic proteins at different temperatures have been done [6–10]. In fact to our knowledge no comparative molecular dynamics study dealing with wide range of temperature has been reported. In this paper for the first time we have performed molecular dynamics simulation of AFP III and its mesophilic homologue at five different temperatures namely 298 K, 310 K, 373 K, 423 K and 473 K. In addition to five temperatures we have carried out 277 K (4 °C) simulation for AFP III, which is close to the growth temperature of this protein. The dynamic properties of two proteins at different temperatures have been compared in terms of secondary structure content, molecular flexibility, intramolecular hydrogen bonds and protein–solvent interactions. The thermal unfolding pathways of two proteins have also been investigated. Comparisons of essential conformational subspaces of these two proteins at different temperatures have been monitored by principal component analysis.

2. Methodology

2.1. Molecular dynamics simulation

All MD simulations were performed using the GROMACS 3.3.1 [11,12] package and GROMOS96 [13] 43a1 force field implemented on LINUX architecture. The coordinates for starting configurations AFP III and its mesophilic homologue were obtained from the Protein Data Bank (PDB entry codes 9AME [5] and 1WVO [4]), which consisted of 66 and 79 residues, respectively. Pairwise sequence alignment by ClustalW [14] was performed to find out the correspondence between residues of two proteins. In order to make equal number of residues for both the proteins four residues from the N-terminus and eleven residues from the C-terminus of 1WVO and one residue each from N and C termini of 9AME were deleted in such a way that original fold retained in both the structures. The starting structure of MD simulation of both the proteins thus contained 64 residues. We abbreviated the truncated 9AME as pMa and 1WVO as mHu throughout the text. Crystallographic water molecules and heteroatoms were removed from the systems. All starting structures were immersed in a triclinic box of SPC water molecules [15]. Box dimensions for pMa and mHu were 5.503 nm × 4.821 nm × 4.554 nm with 3652 SPC water molecules and 4.781 nm × 4.923 nm × 4.783 nm with 3407 SPC water molecules, respectively. All protein atoms were at a distance equal to 1.0 nm from the box edges. No counter ions were added to pMa because the system was already neutral, whereas 6 NA^+ ions were added to neutralize the charge of mHu. Each system was subjected to energy minimization for 2000 steps by steepest descents. The minimized systems for both mHu and pMa were equilibrated for 50 ps each at five different temperatures namely 298 K, 310 K, 373 K, 423 K and 473 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. An additional MD simulation was also performed at 277 K for pMa only. In sum, trajectories of 110 ns (60 ns for pMa, 50 ns for mHu) were collected for two proteins investigated. Periodic boundary conditions combined with minimum image convention were used under isothermal, isobaric conditions using Berendsen coupling algorithm [16] with relaxation times of 0.1 ps and 0.5 ps, respectively. The LINCS algorithm [17] 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) [18,19] 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 [20]. Other analyses were performed using scripts included with the Gromacs [21] distribution. The visual analysis of protein structures was carried out using Rasmol [22].

2.2. Essential dynamics

Essential degrees of freedom of both pMa and mHu were extracted from equilibrated portion of the trajectories according to principal component analysis (PCA) or essential dynamics (ED) method [23,24]. 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 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. For the simulation of both pMa and mHu only Cα atoms were included in the definition of the covariance matrices for the protein.

The root-mean-square inner product (RMSIP) [25,26] was calculated to measure the degree of overlap between the conformational spaces of the two proteins explored by the simulation at both the same and different temperatures. The RMSIP is defined as follows:

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

where η_i^a, h_j^b are the ith and jth eigenvectors from two different sets a and b, respectively. 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.

In order to estimate how much of the major essential modes from the covariance analysis resemble random diffusion, its cosine content [27,28] c_i of the eigenvector i was calculated as follows:

$$c_i = \frac{2}{T}\left(\int_0^T \cos(i\pi t)\, p_i(t)\, dt\right)^2 \left(\int_0^T p_i^2(t)\, dt\right)^{-1}$$

where $pi(t)$ is the amplitude of the motion along eigenvector i at time t. The cosine content can take values between 0 (no cosine) and 1 (a perfect cosine), and provides an indicator of the extent of sampling.

3. Results and discussion

3.1. Global structural stability

Fig. 2A.I and A.II show the backbone RMSDs of pMa and mHu from the corresponding starting structures as a function of time at different temperatures. At 277 K the average backbone RMSD of pMa is ~0.1 nm showing close resemblance to the starting crystal structure. At 298 K and 310 K, the RMSDs increase up to an equilibrium value of 0.12 nm and 0.24 nm, respectively for pMa and 0.20 nm and 0.21 nm, respectively for mHu showing very small difference in RMSD values of the two proteins. In the trajectory run at 373 K backbone RMSDs of pMa and mHu attain a value of 0.3 nm and 0.19 nm, respectively indicating significant structural changes for pMa. At 423 K and 473 K simulations, backbone RMSDs of both pMa and mHu increase from the beginning of the simulation. However, pMa displays higher RMSD

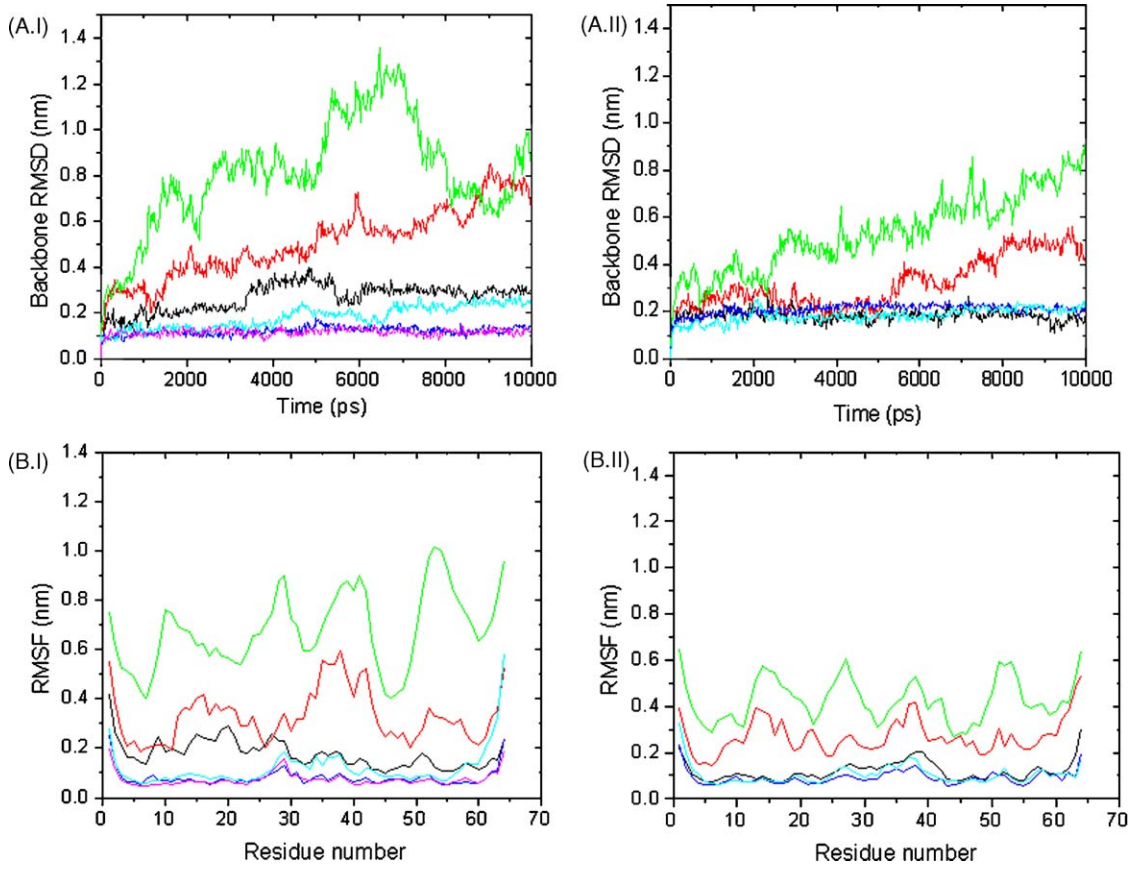

Fig. 2. Time evolution of (A) Backbone RMSD of (I) pMa and (II) mHu at different temperatures. (B) RMSFs as a function of residue number of (I) pMa and (II) mHu at different temperatures. The color-coding scheme is as follows: 277 K (pink), 298 K (blue), 310 K (cyan), 373 K (black), 423 K (red) and 473 K (green).

Table 1
Average values of Rg, SASA and average number of protein–protein (intramolecular) and protein–solvent H-bonds in pMa and mHu at different temperature simulations.

	298 K	310 K	373 K	423 K	473 K
pMa					
[a]Rg	1.03	1.05	1.05	1.06	1.14
[a]SASA	35.85	38.20	39.60	40.90	42.06
Protein–protein H-bond	39.50	31.0	30.50	25.50	22.0
Protein–solvent H-bond	98.50	109.50	89.50	85.50	70.0
mHu					
[a]Rg	1.06	1.07	1.08	1.09	1.12
[a]SASA	41.16	40.84	41.70	42.25	44.50
Protein–protein H-bond	33.50	33.0	32.0	30.50	25.50
Protein–solvent H-bond	159.50	147.0	138.50	120.50	119.50

[a] Rg, SASA and H-bonds denote radius of gyration, solvent accessible surface area and hydrogen bonds, respectively. The values of Rg and SASA are given in nm.

values than that of mHu implying rapid denaturation of pMa than mHu. Thus for mHu, changing temperature up to 373 K shows little effect on the stability of its native structure according to the MD simulations.

Despite these changes in RMSD between pMa and mHu, the radius of gyration (Rg) does not show significant change at different temperatures for both the proteins (Table 1, Supplementary Fig. 1). However the change of solvent accessible surface area (SASA) is considerably more in pMa than mHu (Table 1).

3.2. Secondary structure

Table 2 reports the average number of residues in a given secondary structure and in the corresponding X-ray and NMR structures calculated by DSSP [16]. Time evolution of secondary structural elements of both the proteins is also shown in Supplementary Fig. 2. When the average secondary structure content over time is considered differences between pMa and mHu are evident from 310 K onwards simulations. At 310 K pMa exhibits lower content of β-sheets, 3_{10} helices and greater number of residues in coil-like conformation than mHu. The same pattern is maintained at 373 K. It is worth mentioning that percentage of β-sheet content of the simulated conformations deviates much from the same of experimental structures (crystal and NMR) for both the proteins. This may be due to the fact that several β-bridges have transformed into β-sheets during simulation. It is more so in pMa than mHu. In addition to the original β-sheets these additional β-sheets are maintained throughout the simulation for both the proteins at lower temperatures (up to 373 K). On the contrary the percentage of α-helix content remains close to experimental structures and it is maintained stably up to 373 K for both the proteins. However only in mHu α-helix is stably maintained up to 423 K.

The average number of intramolecular hydrogen bonds (H-bonds) is calculated at different temperature simulations for pMa and mHu (Table 1). Significant differences in protein–solvent interactions are also evident in pMa and mHu. At all temperatures the average number of protein–solvent interactions is greater for mHu (Table 1). It has been noted that mHu contains greater number of hydrophilic residues (28) than pMa (23). The solvent exposed accessible surface area of these hydrophilic residues is also greater in mHu than pMa. Higher number of interactions with the solvent is taking place through these hydrophilic residues in mHu that may be responsible for higher solvent interactions of mHu than pMa.

3.3. Structural flexibility

Fig. 2B compares the root-mean-square fluctuations (RMSF) of Cα atoms on a residue basis for every temperature simulations for pMa and mHu. Protein regions characterized by large RMSF values are generally located at corresponding positions in pMa and mHu even though the intensity of the fluctuations is often different. Regular secondary structure regions show small fluctuation during the simulation whereas significant fluctuations are observed for loop and terminal regions. In order to identify regions characterized by different flexibility in pMa and mHu, the RMSF profile of mesophilic protein is subtracted from the corresponding profile obtained for psychrophilic one (rmsf-diff). The correspondence between residues of pMa and mHu is determined from a sequence alignment of the two proteins (Fig. 1A). The regions characterized by a positive rmsf-diff value indicate greater flexibility in the psychrophilic protein whereas a negative rmsf-diff value indicates greater flexibility in the mesophilic protein. Notably, there are regions characterized by different flexibility in psychrophilic and mesophilic protein. At 298 K, amino acids 26–32 (loop I) show larger flexibility in pMa than mHu (Fig. 3A). At 310 K, in addition to loop I, the protein portion including amino acids 34–42 and 56–64 show greater flexibility in pMa (Fig. 3B). At 373 K, the flexibility of most of the residues in pMa especially amino acids 8–11, 14–29, 33–37 and 48–53 (loop II) increase considerably than mHu (Fig. 3C). In addition to the above-mentioned residues at 373 K, loop I also exhibits greater flexibility in pMa compared to mHu. At 423 K and 473 K both the proteins exhibit enhanced flexibility, but

Table 2
Average secondary structure contents in pMa and mHu trajectories obtained at different temperatures and in the corresponding starting structures.

	Starting	298 K	310 K	373 K	423 K	473 K
pMa						
Coil	20 [31.25%]	18.5 (3.5) [28.9%]	23.5 (2.1) [36.7%]	31 (2.8) [48.4%]	24 (7.0) [37.5%]	29 (8.4) [45.3%]
β-Sheet	8 [12.5%]	15 (1.4) [23.4%]	14 (2.8) [21.8%]	11.5 (0.7) [17.9%]	11 (1.4) [17.1%]	10.5 (3.5) [16.4%]
β-Bridge	8 [12.5%]	3 (2.8) [4.6%]	3.5 (2.1) [5.4%]	3 (1.4) [4.6%]	4.5 (2.1) [7.0%]	2 (0) [3.1%]
Bend	7 [10.93%]	9.5 (2.1) [14.8%]	9 (1.4) [14.0%]	7.5 (2.1) [11.7%]	20 (0) [31.2%]	17.5 (2.1) [27.3%]
Turn	8 [12.5%]	8 (2.8) [12.5%]	10 (0) [15.6%]	6.5 (3.5) [10.1%]	3 (4.2) [4.6%]	5 (7.0) [7.8%]
α-Helix	4 [6.25%]	4 (0) [6.25%]	4 (0) [6.25%]	4.5 (0.7) [7.0%]	0 (0) [0%]	0 (0) [0%]
3_{10}-Helix	9 [14.06%]	7.5 (2.1) [11.7%]	0 (0) [0%]	0 (0) [0%]	0 (0) [0%]	0 (0) [0%]
mHu						
Coil	22 [34.37%]	17.5 (0.7) [27.3%]	15.5 (0.7) [24.2%]	20 (4.2) [31.2%]	22 (2.8) [34.3%]	38.5 (0.7) [60.1%]
β-Sheet	8 [12.5%]	15 (0) [23.4%]	17 (2.8) [26.5%]	18 (4.2) [28.1%]	17 (4.2) [26.51%]	5 (0.7) [7.8%]
β-Bridge	7 [10.93%]	5 (0) [7.8%]	4 (0) [5.4%]	3 (0) [4.6%]	2 (2.8) [3.1%]	3 (1.4) [4.6%]
Bend	9 [14.06%]	10.5 (0.7) [16.4%]	13 (1.4) [20.3%]	13.5 (2.1) [21.0%]	13 (4.2) [20.3%]	12.5 (2.1) [19.5%]
Turn	8 [12.5%]	10.5 (2.1) [16.4%]	10 (2.8) [15.6%]	6 (5.6) [9.3%]	8.5 (6.3) [13.2%]	0 (0.7) [0%]
α-Helix	4 [6.25%]	4 (0) [6.25%]	4 (0) [6.25%]	4.5 (0.7) [7.0%]	2.5 (3.5) [3.9%]	0 (0) [0%]
3_{10}-Helix	6 [9.37%]	3 (0) [4.6%]	4.5 (6.3) [7.0%]	1.5 (0) [0%]	0 (0) [0%]	0 (0) [0%]

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

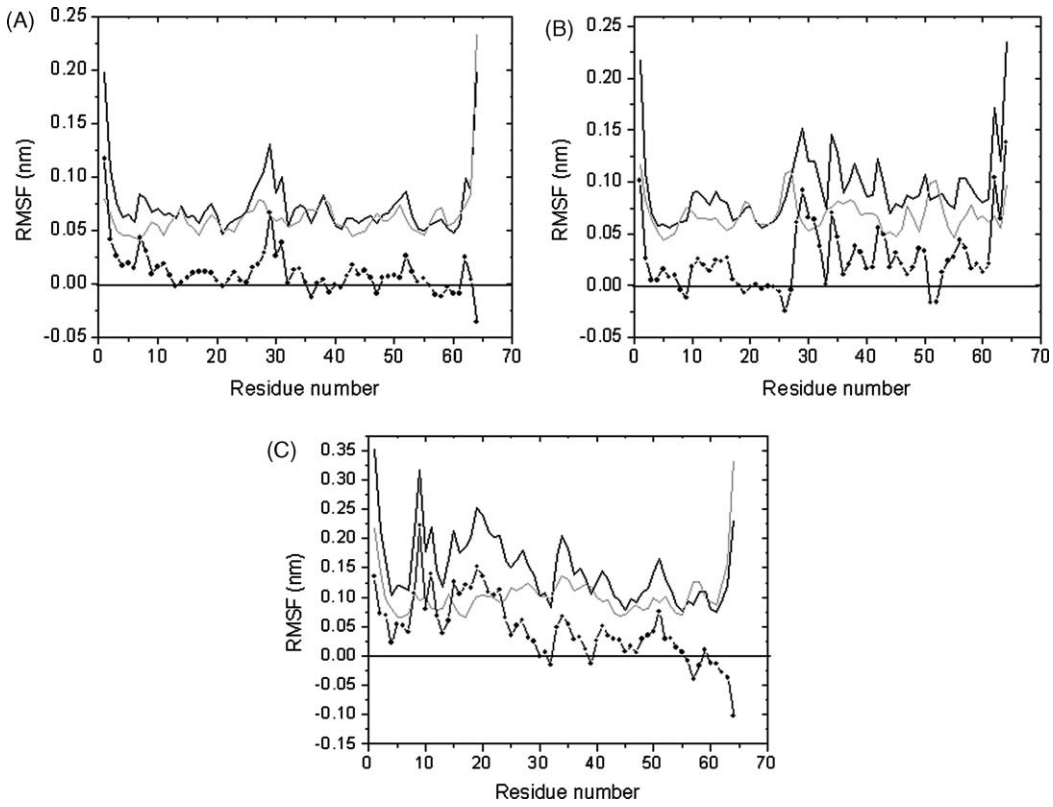

Fig. 3. RMSF and rmsf-difference as a function of residue number. The RMSF values of pMa (black), mHu (grey) and rmsf-difference (black dotted line) at (A) 298 K, (B) 310 K and (C) 373 K.

the extent of flexibility is higher in pMa than mHu. Thus at 277 K and 298 K pMa shows localized enhanced flexibility (in the loop region) than mHu. Analogous finding has been reported from comparative molecular dynamics studies of psychrophilic and mesophilic uracil DNA glycosylase (UDG) where investigators have shown that at 300 K enhanced flexibility lies locally in the loop region involved in DNA recognition [29]. Comparative molecular dynamics studies of psychrophilic and mesophilic trypsin reported that the active site of psychrophilic trypsin shows an apparent higher flexibility and deformability at 300 K [10]. Recently it has been reported that at 300 K psychrophilic trypsin and elastase show enhanced localized flexibility in the loop regions, which confers larger flexibility to these enzymes [7]. Earlier report suggests that enhanced protein flexibility could be strongly related to cold-adaptation [6,9]. Thus improved flexibility of loop regions in pMa (at 277 K and 298 K) could be a positive factor for cold-adaptation.

3.4. Essential dynamics

The essential dynamics analysis has been performed to monitor the overall concerted motions of the two proteins at different temperatures. In order to examine the efficiency of the sampling of conformational space we have calculated the cosine content of first three principal components (PC1, PC2 and PC3) up to 373 K for both the proteins (Table 3). Hess suggested [27,28] that the sampling of a simulation is far from converged when the first few PCs resemble cosines. Insufficient sampling can lead to high cosine values, representative of random motions. The low cosine values of Table 3 indicate that the diffusive contents of three eigenvectors of different simulation groups are relatively low and thus they reveal good simulation convergence.

Table 4 reports the comparison of the essential subspaces of pMa and mHu spanned by the first 10 eigenvectors at five different temperatures by RMSIP method. Root-mean-square inner product measures the degree of overlap or similarity of eigenvector sets. RMSIP data in Table 4 can be best described by splitting into two groups, those between simulations of the same protein and those between different proteins. The diagonal elements of Table 4 represent the overlap within the individual trajectories. As expected the highest degree of overlap occurs within the individual trajectories obtained at different temperatures (diagonal elements of the table). Table 4 also suggests that the RMSIPs between the same proteins at different temperatures are high. On the contrary, the RMSIPs of two proteins at different temperatures are significantly low. Thus these data indicate that the essential subspaces explored by the simulations of the individual protein (but at different temperatures up to 373 K) overlap significantly. Although at higher temperatures, the overlap within the same protein decreases but these values are higher than the overlap values between two different proteins at higher temperatures (373 K and 423 K). Therefore at lower as well as at higher temperatures essential subspaces of the two proteins are not at all overlapped. These results also suggest that each protein has similar direction of motion at different simulated temperatures. However

Table 3
Cosine-contents of pMa and mHu at different temperature simulations.

	298 K	310 K	373 K
pMa			
PC1	0.067	0.005	0.290
PC2	0.184	0.088	0.034
PC3	0.008	0.000	0.026
mHu			
PC1	0.019	0.009	0.141
PC2	0.060	0.098	0.025
PC3	0.044	0.020	0.005

Table 4
RMSIP between the first ten eigenvectors for trajectories obtained at different temperatures.

	p298K	p310K	p373K	p423K	p473K	m298K	m310K	m373K	m423K	m473K
p298K	0.796*	0.712	0.637	0.500	0.453	0.435	0.424	0.426	0.398	0.415
p310K		0.725*	0.626	0.510	0.492	0.472	0.425	0.459	0.397	0.427
p373K			0.715*	0.577	0.461	0.437	0.424	0.454	0.367	0.447
p423K				0.595*	0.498	0.417	0.460	0.424	0.408	0.475
p473K					0.584*	0.448	0.453	0.411	0.441	0.501
m298K						0.716*	0.740	0.675	0.550	0.477
m310K							0.744*	0.670	0.544	0.474
m373K								0.806*	0.537	0.477
m423K									0.615*	0.500
m473K										0.583*

p stands for pMa and m for mHu.
* These values refer to the RMSIP obtained by comparing two halves of the equilibrated portion of the corresponding trajectories.

two proteins show different direction of motions at different simulated temperatures.

The fluctuations of Cα atoms of pMa and mHu projected on first principal component at 298 K, 310 K and 373 K are shown in Fig. 4. The primary internal motions of both the proteins are similar to corresponding RMSF plots of Cα atoms shown in Fig. 3.

The overall flexibility of two proteins has been calculated by the trace of the diagonalized covariance matrix of the atomic positional fluctuations. We have obtained the following values for pMa: 0.26 nm^2 at 277 K, 0.38 nm^2 at 298 K, 0.63 nm^2 at 310 K, 1.63 nm^2 at 373 K, 7.66 nm^2 at 423 K and 31.15 nm^2 at 473 K. Similarly for mHu we have, 0.28 nm^2 at 298 K, 0.31 nm^2 at 310 K, 0.77 nm^2 at 373 K, 5.04 nm^2 at 423 K and 12.35 nm^2 at 473 K again confirming the overall increased flexibility of pMa than mHu at higher temperatures. It should be pointed out that the overall flexibility of both the proteins near their growth temperature is almost the same: 0.26 nm^2 (at 277 K) for pMa, 0.28 nm^2 (at 298 K) and 0.31 nm^2 (at 310 K) for mHu. Moreover, the correlation coefficient of residue dependent RMSF profiles of the two proteins near their growth temperatures (277 K of pMa with 298 K and 310 K of mHu) is also high (0.8). This is in accordance with our finding obtained from the trace of the diagonalized covariance matrix, i.e. the overall flexibility of both the proteins near their growth temperature is almost similar.

Dynamical differences between pMa and mHu can be appreciated by looking at the Cα projections onto the first principal component (PC1) (Fig. 5A and B). PC1 accounts for ∼40% and ∼35% of the internal motion for pMa and mHu, respectively, at 310 K and 373 K. The width of the ribbon indicates the amplitude of the backbone motion whilst the arrows evidence the directions. Notably, the motion of loops I and II is different at 310 K and 373 K in two different proteins. However, the direction of motion of loops I and II is quite similar between the same proteins at 310 K and 373 K. These results strengthen RMSIP data (Table 4) and further suggest that internal motions of two proteins are significantly different at same temperature.

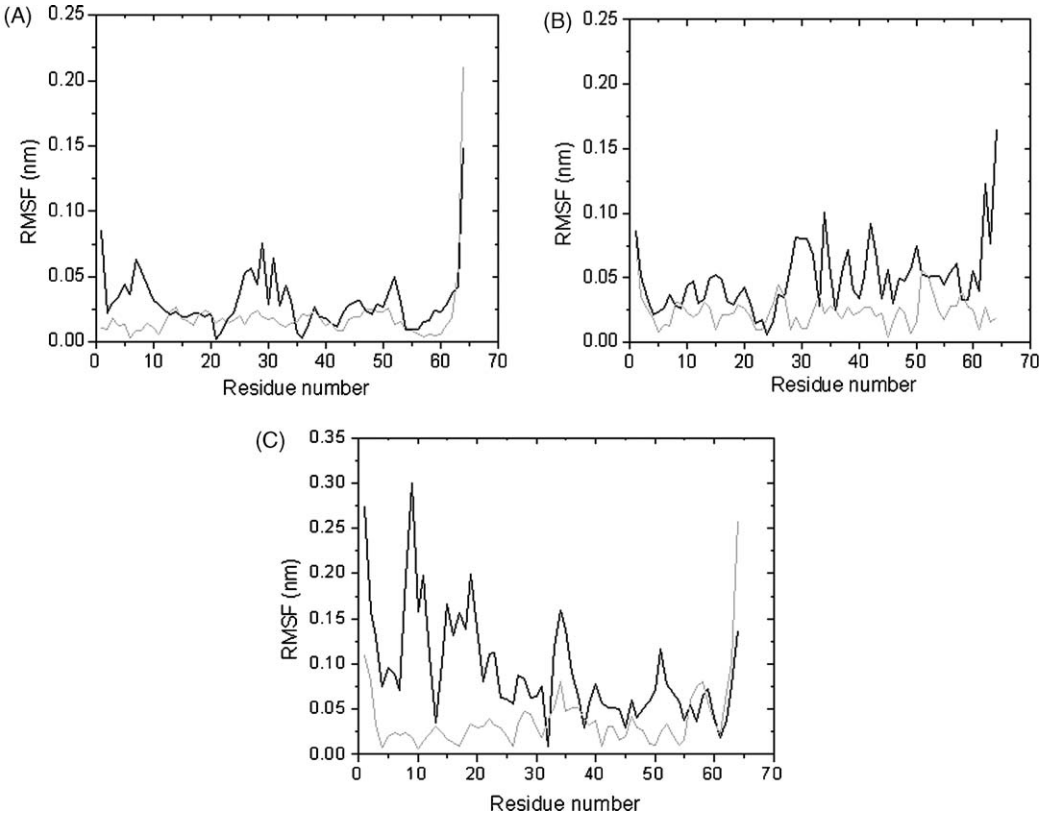

Fig. 4. RMSFs of Cα atoms projected along PC1 of pMa (black) and mHu (grey) at (A) 298 K (B) 310 K and (C) 373 K.

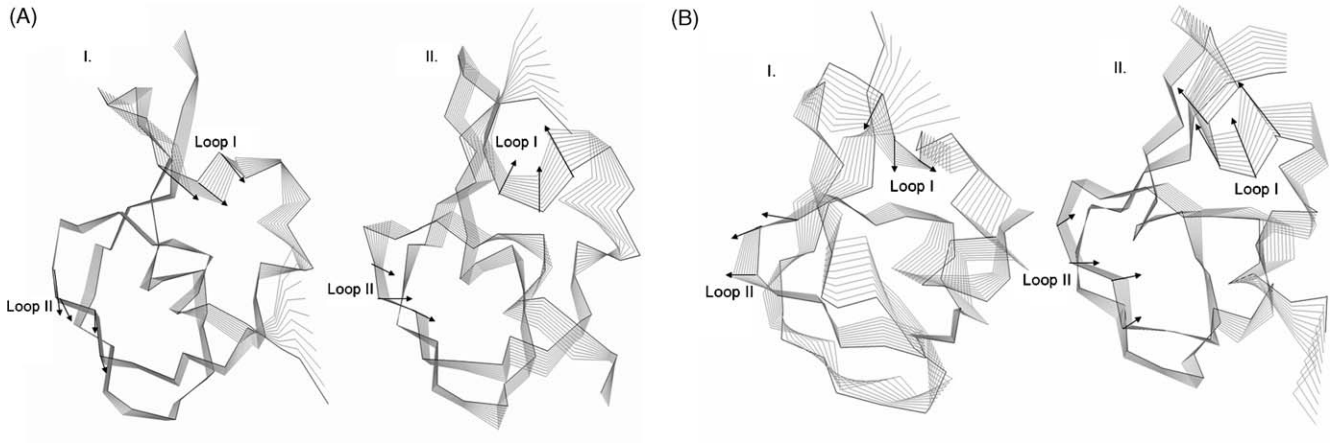

Fig. 5. Superimposition of 10 configurations obtained by projecting the Cα motion onto the first eigenvector at (A) 310 K and (B) 373 K of (I) pMa and (II) mHu. Arrows are used as qualitative indicators of the motion direction.

3.5. Conformational space sampling

The dynamics of two proteins is best achieved via characterization of its phase space behavior. The projection of trajectories obtained at 310 K and 373 K onto the first two principal components (PC1, PC2) shows the motion of two proteins in phase space (Fig. 6) (298 K and higher temperature data not shown). On these projections, we see clusters of stable states. Two features are very apparent from these plots. Firstly the clusters are well defined in mHu than pMa. Secondly, pMa covers a larger region of phase space particularly along PC1 plane compared to that covered by mHu. At high temperature particularly at 373 K mHu tends to show drastically reduced coverage than pMa. Our observation thus corroborates with the idea of higher flexibility of pMa than mHu at 310 K and onwards higher temperatures.

3.6. Unfolding pathway

The sequence of unfolding events has been monitored for both the proteins at all temperatures (298 K, 310 K, 373 K, 423 K and 473 K). At 298 K, the structures of both the proteins are stable and remained close to the native structure during the whole simulation (Fig. 7A and

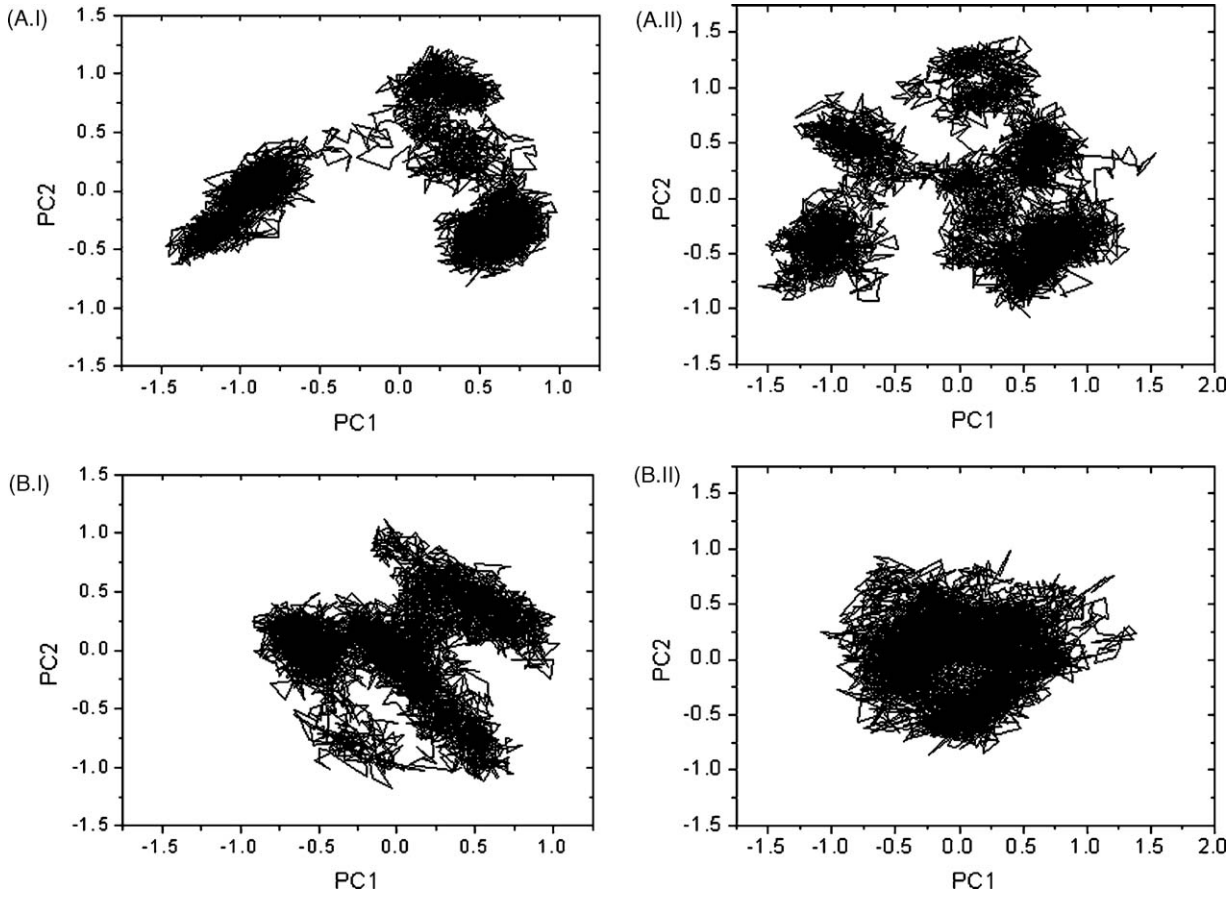

Fig. 6. Projection of the motion of (A) pMa and (B) mHu in phase space along the first two principal eigenvectors at (I) 310 K and (II) 373 K.

B). At 310 K mHu maintains its native structure (Fig. 7B) whereas structural changes start to appear in pMa at this temperature (Fig. 7A). At 310 K intermediate, C- and N-terminal 3_{10} helices denature at around 2.5 ns, 3.9 ns and 4.1 ns, respectively. At 373 K pMa loses all 3_{10} helices (intermediate, C- and N-terminal) within 1 ns, original β-sheets denature at around 4 ns and rearranged β-sheets appears in the structure, α-helix is however stably maintained throughout the simulation. Since intermediate 3_{10} helix is absent in

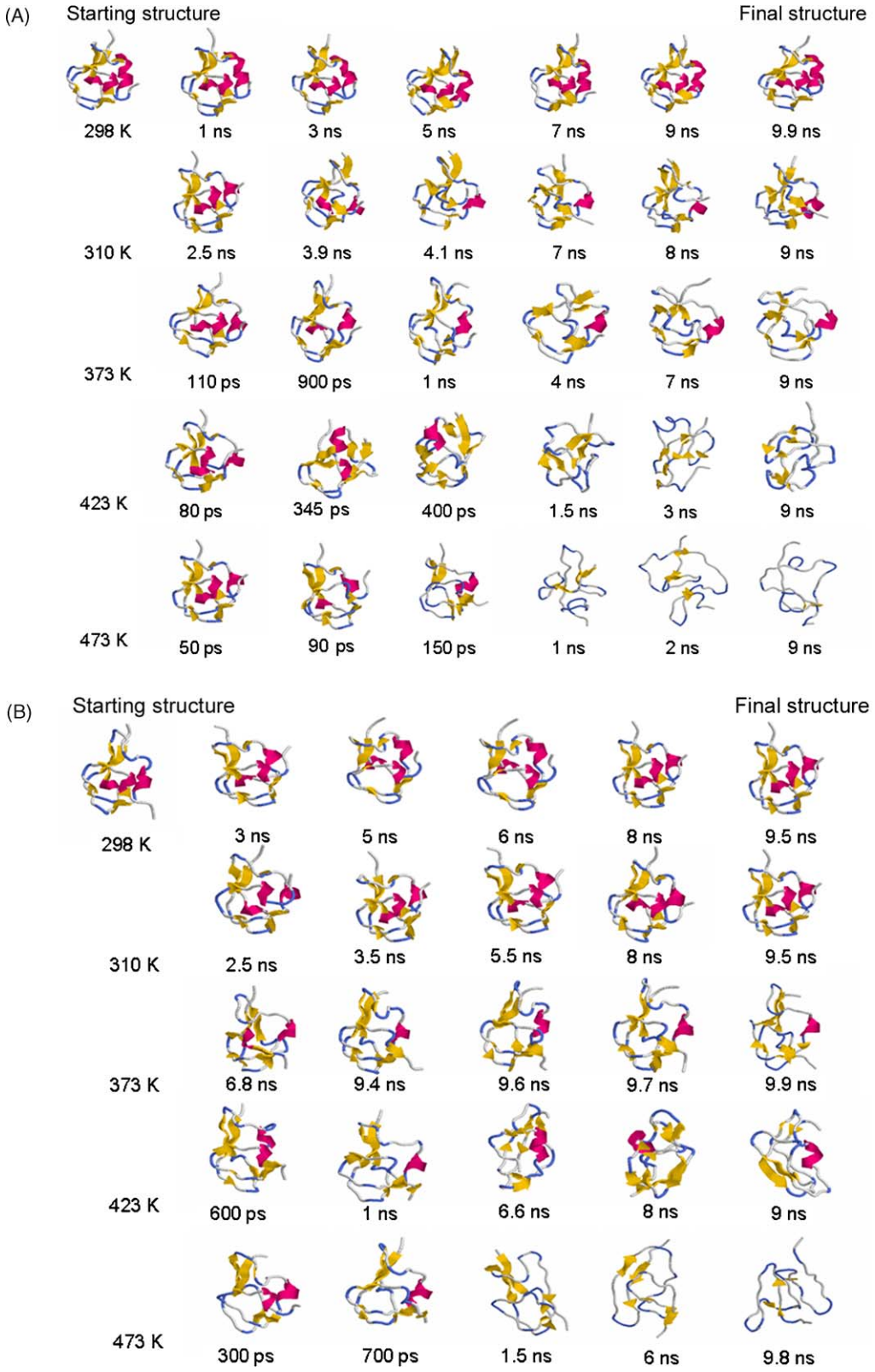

Fig. 7. Snapshots of (A) pMa and (B) mHu evolved at different time points at different temperatures. The structures having similar secondary structures but generated at different temperatures have been put in a same column. For (A) the first column shows initial disruption of the intermediate 3_{10} helix, followed by loss of C- and then N-terminal 3_{10} helices, fourth column shows loss of α-helix followed by rearrangement of β-sheets. For (B) the first column shows disruption of the N-terminal 3_{10} helix followed by the unfolding of C-terminal 3_{10} helix, followed by loss of α-helix and rearrangement of β-sheets.

mHu the unfolding event starts at 373 K by denaturation of N- and C-terminal 3_{10} helices around 6.8 ns and 9.4 ns, respectively. However at 373 K the rest of the structure remains unaltered in mHu throughout the simulation. At 423 K pMa loses all the 3_{10} helices even at earlier time point (within 400 ps), β-sheets disappear after 1.5 ns and α-helix is maintained only for 1.4 ns. At 423 K mHu loses all 3_{10} helices within 1 ns, β-sheets disorganize at 6.6 ns which however reappears as rearranged β-sheets (which is different from the original β-sheets) at the end of the simulation. At 423 K the α-helix is maintained in mHu throughout the simulation. At 473 K, both the proteins become highly coiled early in the simulation only a few rearranged β-sheets remain in the structure.

Thus the study of unfolding pathways of pMa and mHu reveal that, the rate of unfolding is faster in pMa than mHu. With increasing temperature each protein however shows analogous unfolding events at earlier time points. However the sequence of unfolding events is different in two proteins. In pMa, the intermediate 3_{10} helix unfolds first followed by the denaturation of C-terminal 3_{10} helix, the N-terminal 3_{10} helix is denatured after C-terminal helix. On the contrary, in mHu unfolding is initiated at the N-terminal 3_{10} helix followed by the unfolding of C-terminal 3_{10} helix. The unfolding pathway of pMa is similar to the pathway of another type III antifreeze protein from Antarctic eelpout studied previously by molecular dynamics simulation [3].

Comparative unfolding studies of psychrophilic and mesophilic uracil DNA glycosylase by molecular dynamics simulations have been reported previously at three different temperatures namely 375 K, 400 K and 425 K [6]. According to this study, the protein unfolding proceeds slowly in the 375 K simulations. The unfolding rate is faster when the temperature is raised to 400 K. However, at 375 K and 400 K simulations psychrophilic enzyme shows higher level of unfolding compared to their mesophilic counterpart [6]. In our simulation the lowest temperature at which unfolding events first observed is 373 K. The protein unfolding however proceeds slowly at this temperature with the effect of more pronounced unfolding in pMa than mHu. At 423 K the unfolding process accelerates for both the proteins with again pMa shows higher level of unfolding compared to mHu. At 473 K, pMa and mHu lose the secondary structure very early in the simulation afterwards unfolding proceeds at a similar rate for both the proteins.

4. Conclusion

Comparative molecular dynamics simulations of pMa and mHu at five different temperatures have been performed. Analyses of trajectories in terms of secondary structure content, solvent accessibility, intramolecular hydrogen bonds and protein–solvent interactions indicate distinct differences in these two proteins. The study of unfolding pathways of pMa and mHu reveal that the rate of unfolding is faster in pMa than mHu. Moreover the sequence of unfolding events is different in two proteins. The overall flexibility calculated by the trace of the diagonalized covariance matrix displays similar flexibility of both the proteins near their growth temperatures. However at higher temperatures psychrophilic protein shows increased overall flexibility than its mesophilic counterpart. Principal component analysis also indicates that the essential subspaces explored by the simulations of two proteins at different temperatures are non-overlapping and they show significantly different directions of motion especially in the loop regions. However, there are significant overlaps within the trajectories and similar directions of motion of each protein especially at 298 K, 310 K and 373 K. Overall, the psychrophilic protein leads to increased conformational sampling of the phase space than its mesophilic counterpart.

Our study may help in elucidating the molecular basis of thermostability of homologous proteins from two organisms living at different temperature conditions. Such an understanding is required for designing efficient proteins with characteristics for a particular application at desired working temperatures.

Acknowledgement

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

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2009.01.004.

References

[1] O. Garcia-Arribas, R. Mateo, M.M. Tomczak, P.L. Davies, M.G. Mateu, Thermodynamic stability of a cold-adapted protein, type III antifreeze protein, and energetic contribution of salt bridges, Protein Sci. 16 (2007) 227–238.
[2] G. Chen, Z. Jia, Ice-binding surface of fish type III antifreeze, Biophys. J. 77 (1999) 1602–1608.
[3] S. Kundu, D. Roy, Temperature-induced unfolding pathway of a type III antifreeze protein: insight from molecular dynamics simulation, J. Mol. Graph. Model. 27 (2008) 88–94.
[4] T. Hamada, Y. Ito, T. Abe, F. Hayashi, P. Güntert, M. Inoue, T. Kigawa, T. Terada, M. Shirouzu, M. Yoshida, A. Tanaka, S. Sugano, S. Yokoyama, H. Hirota, Solution structure of the antifreeze-like domain of human sialic acid synthase, Protein Sci. 15 (2006) 1010–1016.
[5] S.P. Graether, C.I. DeLuca, J. Baardsnes, G.A. Hill, P.L. Davies, Z. Jia, Quantitative and qualitative analysis of type III antifreeze protein structure and function, J. Biol. Chem. 274 (1999) 11842–11847.
[6] M. Olufsen, B.O. Brandsdal, A.O. Smalås, Comparative unfolding studies of psychrophilic and mesophilic uracil DNA glycosylase: MD simulations show reduced thermal stability of the cold-adapted enzyme, J. Mol. Graph. Model. 26 (2007) 124–134.
[7] E. Papaleo, M. Pasi, L. Riccardi, I. Sambi, P. Fantucci, L. De Gioia, Protein flexibility in psychrophilic and mesophilic trypsins. Evidence of evolutionary conservation of protein dynamics in trypsin-like serine-proteases, FEBS Lett. 582 (2008) 1008–1018.
[8] E. Papaleo, M. Olufsen, L.De Gioia, B.O. Brandsdal, Optimization of electrostatics as a strategy for cold-adaptation: a case study of cold- and warm-active elastases, J. Mol. Graph. Model. 6 (2007) 93–103.
[9] E. Papaleo, L. Riccardi, C. Villa, P. Fantucci, L.De Gioia, Flexibility and enzymatic cold-adaptation: a comparative molecular dynamics investigation of the elastase family, Biochim. Biophys. Acta 2006 (1764) 1397–1406.
[10] B.O. Brandsdal, E.S. Heimstad, I. Sylte, A.O. Smalås, Comparative molecular dynamics mesophilic and psychrophilic protein homologues studied by 1.2 ns simulations, J. Biomol. Struct. Dyn. 17 (1999) 493–506.
[11] 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.
[12] D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J. Berendsen, GROMACS: fast, flexible, and free, J. Comput. Chem. 26 (2005) 1701–1718.
[13] 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.
[14] J.D. Thompson, D.G. Higgins, T.J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Res. 22 (1994) 4673–4680.
[15] 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.
[16] 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.
[17] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: a linear constraint solver for molecular simulations, J. Comput. Chem. 18 (1997) 1463–1472.
[18] 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.
[19] 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.
[20] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (1983) 2577–2637.
[21] 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, http://www.gromacs.org.
[22] R.A. Sayle, E.J. Millner-White, Rasmol-biomolecular graphics for all, Trends Biochem. Sci. 20 (1995) 374–376.
[23] A. Amadei, A.B. Linssen, H.J.C. Berendsen, Essential dynamics of proteins, Proteins 17 (1993) 412–425.
[24] 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.
[25] 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.
[26] 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. Comput. Chem. 18 (1997) 169–181.
[27] B. Hess, Convergence of sampling in protein simulations, Phys. Rev. E 65 (2002) 031910.
[28] B. Hess, Similarities between principal components of protein dynamics and random diffusion, Phys. Rev. E 62 (2000) 8438–8448.
[29] M. Olufsen, A.O. Smalås, E. Moe, B.O. Brandsdal, Increased flexibility as a strategy for cold-adaptation: a comparative molecular dynamics study of cold- and warm-active uracil DNA glycosylase, J. Biol. Chem. 280 (2005) 18042–18048.