J Mol Model (2013) 19:2647–2656
DOI 10.1007/s00894-013-1816-x

ORIGINAL PAPER

Influence of C-terminal tail deletion on structure and stability of hyperthermophile Sulfolobus tokodaii RNase HI

Lin Chen · Ji-Long Zhang · Qing-Chuan Zheng · Wen-Ting Chu · Qiao Xue · Hong-Xing Zhang · Chia-Chung Sun

Received: 17 January 2013 / Accepted: 4 March 2013 / Published online: 26 March 2013

Abstract The C-terminus tail (G144-T149) of the hyperthermophile *Sulfolobus tokodaii* (Sto-RNase HI) plays an important role in this protein's hyperstabilization and may therefore be a good protein stability tag. Detailed understanding of the structural and dynamic effects of C-terminus tail deletion is required for gaining insights into the thermal stability mechanism of Sto-RNase HI. Focused on *Sulfolobus tokodaii* RNase HI (Sto-RNase HI) and its derivative lacking the C-terminal tail (ΔC6 Sto-RNase HI) (PDB codes: 2EHG and 3ALY), we applied molecular dynamics (MD) simulations at four different temperatures (300, 375, 475, and 500 K) to examine the effect of the C-terminal tail on the hyperstabilization of Sto-RNase HI and to investigate the unfolding process of Sto-RNase HI and ΔC6 Sto-RNase HI. The simulations suggest that the C-terminal tail has significant impact in hyperstabilization of Sto-RNase HI and the unfolding of these two proteins evolves along dissimilar pathways. Essential dynamics analysis indicates that the essential subspaces of the two proteins at different temperatures are non-overlapping within the trajectories and they exhibit different directions of motion. Our work can give important information to understand the three-state folding mechanism of Sto-RNase HI and to offer alternative strategies to improve the protein stability.

Keywords Essential dynamics · Hyperthermophilic protein · Molecular dynamics · Thermal stability · Unfolding pathway

L. Chen · J.-L. Zhang · Q.-C. Zheng (✉) · W.-T. Chu · Q. Xue · H.-X. Zhang (✉) · C.-C. Sun
State Key Laboratory of Theoretical and Computational Chemistry, Institute of Theoretical Chemistry, Jilin University, Changchun 130023, People's Republic of China
e-mail: zhengqc@jlu.edu.cn
e-mail: zhanghx@mail.jlu.edu.cn

Abbreviations

MD	Molecular dynamics
Sto-RNase HI	Sulfolobus tokodaii RNase HI
ED	Essential dynamics
SD	Steepest decent
PME	Partical mesh Ewald
RMSIP	Root-mean-square inner product
RMSD	Root mean-square deviations
RMSF	Root-mean-square fluctuations

Introduction

Many hyperthermophilic proteins from different organisms living in constantly hot environments have been identified and characterized during the last few decades [1]. Thanks to their unique structure-function properties of high thermostability and optimal activity at high temperatures, these proteins have attracted much interest not only for academic research but also for the biotechnological market [2, 3]. Understanding the molecular basis of thermostability of hyperthermophilic proteins plays an important role in designing efficient proteins with characteristics for a particular application [4]. *Sulfolobus tokodaii* (Sto) strain 7 is an aerobic thermoacidophilic archaeon with an optimum growth temperature and pH of 80 °C and 2.5, respectively [5]. The Sto genome sequence was reported in 2001 and encoded a variety of hyperthermophilic proteins [6]. These proteins are shown to exhibit high thermostability and are capable of performing their biological functions such as DNA replication, repair, and transcription [7, 8]. Many investigators have made great efforts to understand the structure, hyperthermophilic mechanism, and protein thermostability [5, 9–12].

Ribonuclease HI from the hyperthermophile *Sulfolobus tokodaii* (Sto-RNase HI) is a monomeric protein of 149

amino acids [5]. Sto-RNase HI is characterized by its ability to degrade not only an RNA/DNA hybrid but also a double stranded RNA. Recently, a truncated variant of Sto-RNase HI, lacking six residues at the C-terminus, called ΔC6 Sto-RNase HI, has been produced [11]. The backbone root-mean-square deviation (RMSD) of Sto-RNase HI and ΔC6 Sto-RNase HI is 0.57 Å, implying that C-terminal tail deletion does not significantly affect folding, which is in agreement with the earlier experimental studies [11]. Their overall structure (Fig. 1) consists of a central five-stranded mixed β sheet with three antiparallel and two parallel strands and four α-helices. The C-terminal tail of Sto-RNase HI, which is responsible for its hyperstability, is anchored to the core region through a disulfide bond (Cys58-Cys145), several hydrogen bonds, and hydrophobic interactions [9]. Heat-induced unfolding experiments [11] showed that Sto-RNase HI is a hyperthermostable protein with a denaturation temperature of 102 °C around pH3, 24 °C higher than that of ΔC6 Sto-RNase HI. More importantly, the flexible C-terminal tail can be extensively modified, and this tail has been shown to play a critical role in the designing of variants with enhanced protein stability [11]. Many investigators have made great efforts to understand the role of C-terminal tail in Sto-RNase HI [9, 11, 12], however, the effect of C-terminal tail deletion on the kinetics of thermal denaturation of Sto-RNase HI at elevated temperatures is still not fully understood.

Molecular dynamics (MD) simulations, a powerful tool to complement experimental results with detailed dynamics behavior of biomolecules [13–16], have been frequently used to study the structure function relationships of biological macromolecules [17–23]. This technique can provide insight into the atomic level dynamic structural effects of C-terminus tail deletion of Sto-RNase HI that are not easily observable with conventional experimental tools. However, theoretical studies investigating the structural and dynamic effects of C-terminus tail deletion of Sto-RNase HI are mostly lacking in the currently existing literature. Therefore, in our work, MD simulations have been carried out for Sto-RNase HI and ΔC6 Sto-RNase HI at four different temperatures (300, 375, 475, and 500 K) to elucidate the structural and dynamic consequences caused by the deletion of the C-terminal tail and to evaluate its impact on the unfolding kinetic. Then, the essential dynamics (ED) method was applied to measure the degree of overlap between the conformational spaces of Sto-RNase HI and ΔC6 Sto-RNase HI, as well as to understand the detailed dynamics properties. These characteristics can help us to better understand the effect of C-terminal tail on structure and stability of Sto-RNase HI. During this study, we have specifically focused on the following: (1) characterizing the structural and stable effects of C-terminal tail deletion in Sto-RNase HI, (2) assessing quantitatively the essential subspace in Sto-RNase HI and ΔC6 Sto-RNase HI by calculating the root mean square inner product, (3) elucidating the thermal unfolding pathways of the two proteins.

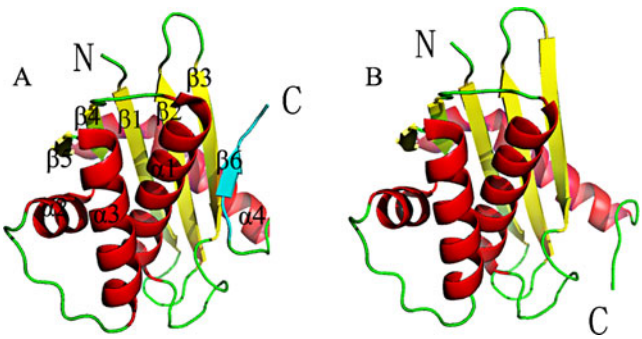

Fig. 1 (**a**) Crystal structure of wild-type Sto-RNase HI. The C-terminal residues are in cyan. (**b**) Crystal structure of ΔC6 Sto-RNase HI

Materials and methods

Molecular dynamics simulations

The crystal structures of wild-type Sto-RNase HI (PDB code: 2EHG [9]) and ΔC6 Sto-RNase HI (PDB code: 3ALY [11]) were obtained from the Protein Data Bank. The protonation states of ionizable residues were chosen on the basis of pKa values of isolated side chain groups at pH3. All starting structures were placed in a rectangular box with a 1.0 nm buffer. The box was filled with SPC water molecules [24] with periodic boundary conditions. The overall charge of the two systems was neutralized by adding the appropriate number of chloride ions. The resulting systems were subjected to energy minimization for 2000 steps of steepest decent (SD) followed by 5000 steps of conjugate gradient minimization. Subsequently, the minimized systems for both Sto-RNase HI and ΔC6 Sto-RNase HI were heated gradually from 100 K to target temperatures (300, 375, 475, and 500 K), and then equilibrated for 300 ps. Finally, production MD simulations for the equilibrated systems were then carried out for 100 ns.

All MD simulations were performed using the GROMACS software package version 4.5.3 [25, 26] and the Amber ff03 force field [27]. All bonds involving hydrogen atoms were constrained by use of the SHAKE algorithm [28]. The time step in all MD simulations was set to 2 fs and the temperature was maintained by the Berendsen coupling algorithm [29] with a 0.1 ps time constant. Long-range electrostatic interactions were calculated using the partical mesh Ewald (PME) [30, 31] summation scheme. Van der Waals and Coulomb interactions were truncated at 1.2 nm. Secondary structure analysis was performed using the program DSSP [32]. Other analyses were performed using script included with the GROMACS [33] distribution. The data given in the figures and tables were obtained from the last 40 ns of the MD simulations unless otherwise stated. The PyMOL [34] and Chimera

[35] software were used to visualize the trajectories and to produce the pictures of the molecular structures.

Essential dynamics analysis

Essential degrees of freedom of both Sto-RNase HI and ΔC6 Sto-RNase HI were extracted from the equilibrated portion of the trajectories on the basis of the ED method [36, 37]. The ED method relies on the construction of the covariance matrix of the atomic coordinate fluctuations. After removal of the translational and rotational degrees of freedom the covariance matrix was calculated. Then this matrix was diagonalized to obtain the eigenvalues and eigenvectors, which characterize information about correlated motions throughout the protein. The eigenvectors indicate the directions of the motion, and the eigenvalues represent the amount of motion along their eigenvector. Finally, the eigenvectors are sorted according to their eigenvalues in descending order. Generally speaking, the first 10 eigenvectors can be applied to describe almost all conformational substates accessible to the protein.

For the simulations of both Sto-RNase HI and ΔC6 Sto-RNase HI, only Cα atoms were included in the definition of the covariance matrices. The root-mean-square inner product (RMSIP) [38–40] was calculated to measure the degree of overlap between the essential subspaces (first 10 eigenvectors) of the two simulated systems at both the same and different temperatures [41]. The RMSIP can be quantified as follows:

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

where η_i^a and η_j^b are the *i*th and *j*th eigenvectors for system a and system b, respectively; 10 means the first 10 eigenvectors. The RMSIP value was calculated from the equilibrated portion of the trajectory. For diagonal elements the RMSIP value was calculated from two equal halves of the trajectory. The ED analysis was performed with the tool available in the GROMACS software package.

Results and discussion

Structural stability of Sto-RNase HI and ΔC6 Sto-RNase HI

To explore the impact of the deletion C-terminal on Sto-RNase HI, MD simulation studies were undertaken on the Sto-RNase HI and ΔC6 Sto-RNase HI at four different temperatures (300, 375, 475, and 500 K). The backbone RMSD referred to the crystal structure for the four parallel simulations were calculated to characterize the structural changes in Sto-RNase HI and ΔC6 Sto-RNase HI (Fig. 2). As can be seen, the structures from the 300 K simulations are in good agreement with the crystal structures exhibiting average RMSD deviation values of 0.09 and 0.12 nm for Sto-RNase HI and ΔC6 Sto-RNase HI respectively (Table 1). In the trajectory run at 375 K the RMSD increase up to an equilibrium value of 0.12 and 0.18 nm respectively for Sto-RNase HI and ΔC6 Sto-RNase HI, showing slight structural changes for ΔC6 Sto-RNase HI. Thus for ΔC6 Sto-RNase HI, changing temperature up to 375 K shows little effect on the stability of its structure according to the MD simulations. At 475 and 500 K simulations, the RMSD of both Sto-RNase HI and ΔC6 Sto-RNase HI increase gradually during the 100 ns simulation. However,

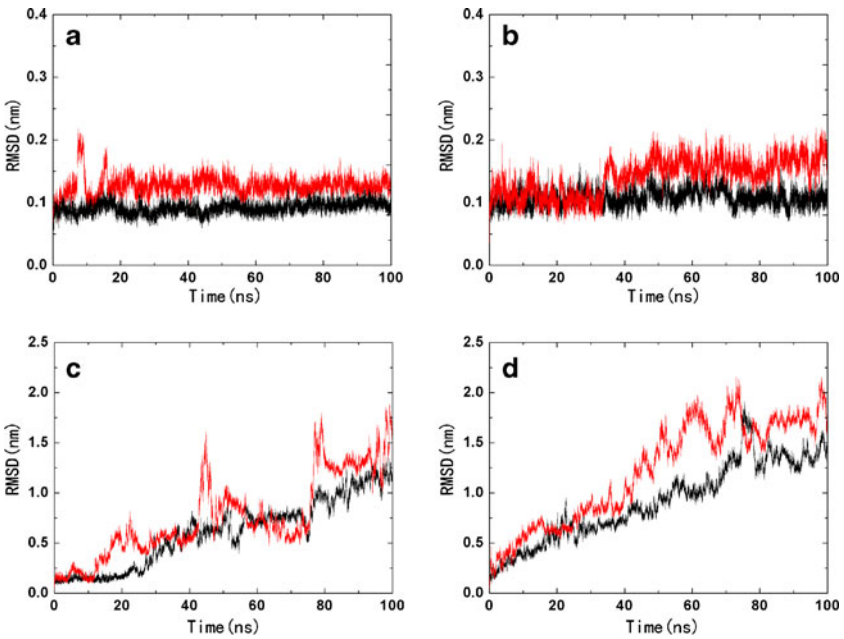

Fig. 2 Root mean square deviations of the backbone atoms of the wild-type Sto-RNase HI (*black*) and ΔC6 Sto-RNase HI (*red*) at four temperatures (**a** 300 K, **b** 375 K, **c** 475 K, and **d** 500 K)

Table 1 Average values of RMSD, Rg, hydrophobic SASA, secondary structure contents, and main chain-main chain H-bonds in wild-type Sto-RNase HI and ΔC6 Sto-RNase HI at home temperature simulations

	300 K	375 K	475 K	500 K
Sto-RNase HI				
RMSD (nm)	0.091±0.008	0.118±0.008	0.88±0.19	1.249±0.21
Rg (nm)	1.48±0.007	1.49±0.007	1.63±0.09	1.69±0.069
Hydrophobic SASA	46.30±0.97	48.41±1.29	57.69±2.01	64.96±2.27
# of residues in β-sheets	34.56±0.78	36.64±1.04	29.9±1.19	16.28±2.69
# of residues in α-helix	54.96±0.97	54.78±1.38	49.3±1.71	40.32±2.10
Main chain-main chain H-bonds	81.16±3.06	78.54±3.66	64.9±4.23	50.46±4.89
ΔC6 Sto-RNase HI				
RMSD (nm)	0.124±0.012	0.183±0.02	1.33±0.14	1.671±0.16
Rg (nm)	1.47±0.004	1.48±0.006	1.66±0.09	1.81±0.18
Hydrophobic SASA	44.35±0.98	46.41±1.18	57.03±1.79	67.11±2.15
# of residues in β-sheets	32.68±0.65	32.71±1.34	25.3±1.59	2.88±0.77
# of residues in α-helix	55.47±0.84	54.16±1.69	43.2±1.97	27.69±2.34
Main chain-main chain H-bonds	77.98±3.22	74.69±3.54	56.2±3.97	39.09±4.21

Sto-RNase HI displays lower RMSD values than that of ΔC6 Sto-RNase HI, implying the unfolding transitions of the Sto-RNase HI occurred much later than the ΔC6 Sto-RNase HI. This is in agreement with the observed reduced stability of the ΔC6 Sto-RNase HI [9, 11]. The calculated radius of gyration also shows similar trends (Table 1). At 300 and 375 K simulations the proteins structures are compact similar to the crystal structures of wild-type Sto-RNase HI and ΔC6 Sto-RNase HI, while the structures from the 475 and 500 K simulations lose their compactness. The above results indicate that the structures of Sto-RNase HI and ΔC6 Sto-RNase HI are in the folded state and are closer to the crystal structures at 300 and 375 K, while the structures undergo significant structural changes leading to unfolded states at 475 and 500 K. This leads us to deduce that low-temperature simulations (300 and 375 K) do not have enough energy to overcome the barrier around the starting conformation in our limited time simulation. However, high-temperature unfolding simulations do show the rapid denaturation of the native structure due to the C-tail deletion.

In addition, the solvent accessible surface area (SASA) for the two proteins was calculated to investigate the impact of the deletion of six residues in the C-terminal on the structure of Sto-RNase HI (Table 1). As can be seen, Sto-RNase HI has higher hydrophobic SASA compared to ΔC6 Sto-RNase HI at 300 and 375 K simulations. These differences in the MD trajectories of Sto-RNase HI and ΔC6 Sto-RNase HI mainly result from the deletion of six residues in the C-terminal of Sto-RNase HI, which contributes to the

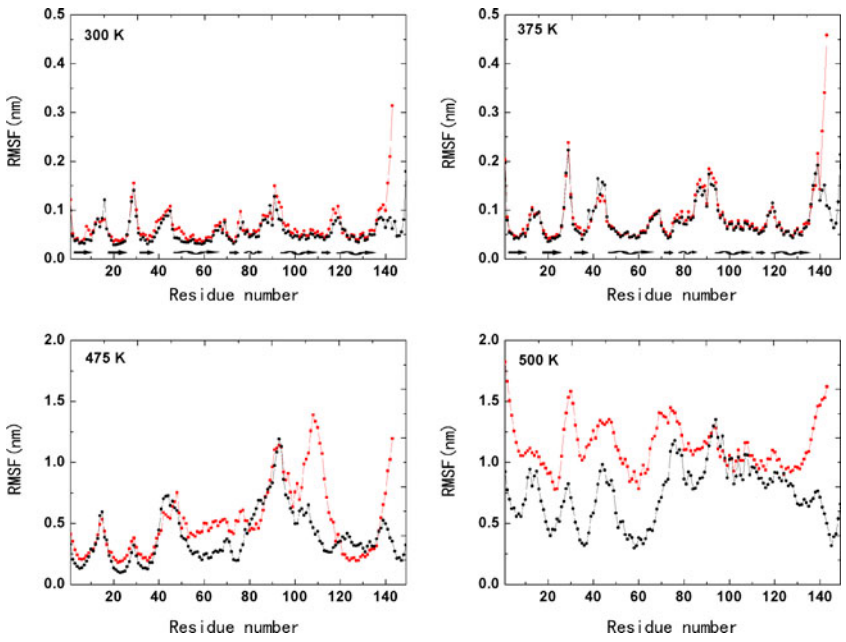

Fig. 3 Root mean square fluctuations of the residues of the wild-type Sto-RNase HI (*black*) and ΔC6 Sto-RNase HI (*red*) at 300, 375, 475, and 500 K. The β-strands and the α-helix are represented by block arrows and wavy arrows, respectively

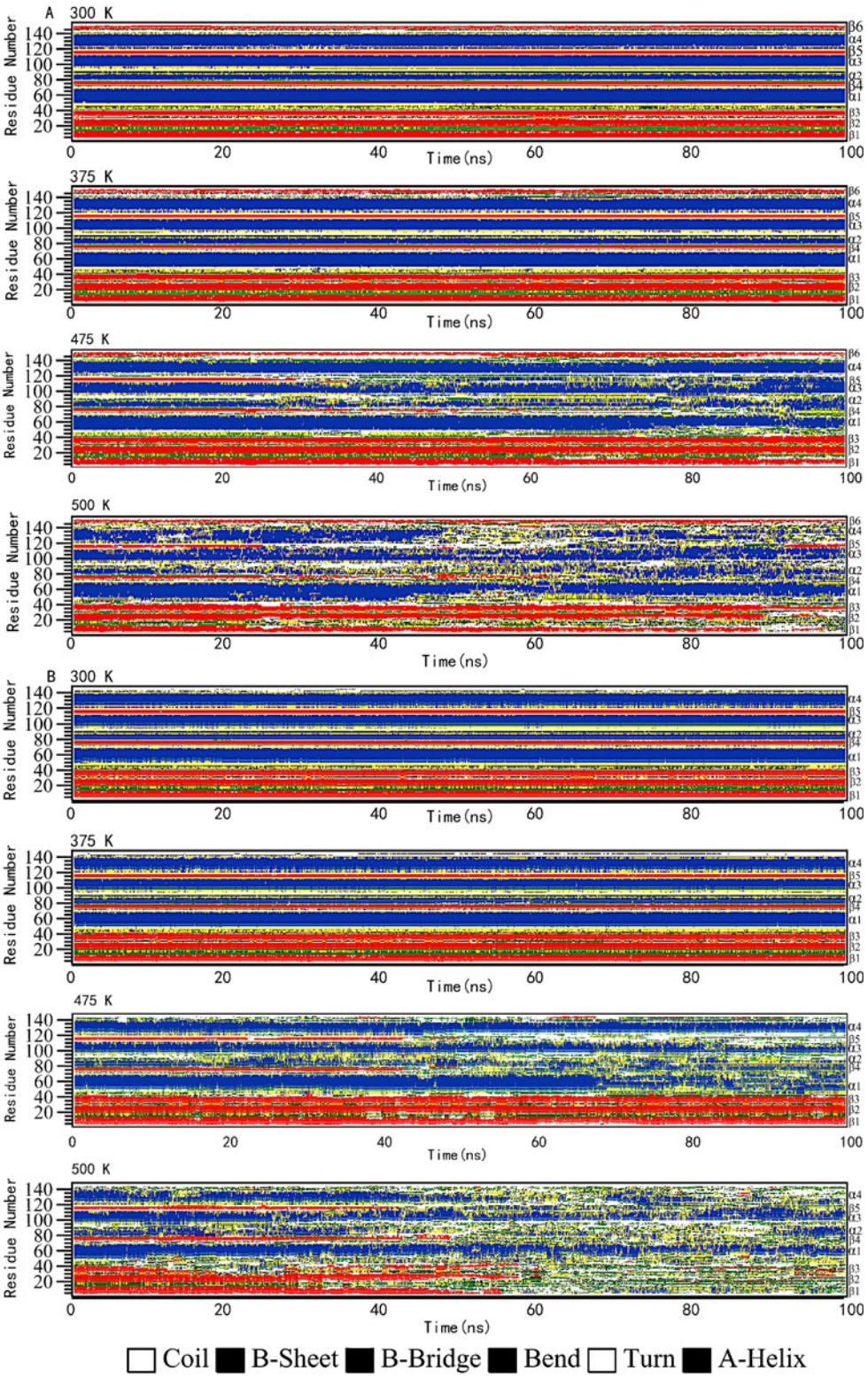

Fig. 4 Time evolution of the secondary structure elements of wild-type Sto-RNase HI (**a**) and ΔC6 Sto-RNase HI (**b**) at 300, 375, 475, and 500 K. The β-strands and the α-helix are represented by β and α, respectively

high stability of the wild type. Such high stability of Sto-RNase HI is consistent with our previous analysis that the unfolding transitions of the Sto-RNase HI occurred much later than the ΔC6 Sto-RNase HI at high temperature simulations.

Structural flexibility

Average root-mean-square fluctuations (RMSF) values in the MD simulation were usually considered as the criterion for the overall flexibility of the system. Thus the RMSF of

Cα atoms on a residue basis for four different temperature simulations were calculated for both Sto-RNase HI and ΔC6 Sto-RNase HI (Fig. 3). The overall fluctuations obtained at 300 K are similar to Sto-RNase HI and ΔC6 Sto-RNase HI except for the terminal regions had sharp peaks in the RMSF plot. Obviously, the RMSF values of these two proteins increase marginally at 375 K. However, at 475 and 500 K simulations most of the residues become highly mobile therefore the curve shows a lot more fluctuations. This is due to the loss of secondary structure at these high temperatures. As expected, the loop and terminal regions are more flexible compared to the rest of the protein since they have no large network of tertiary interactions. The C-terminal residues 141, 142, and 143 of ΔC6 Sto-RNase HI showed significantly different mobility compared to that of Sto-RNase HI, indicating that deletion of the previous C-terminal residues made the new C-terminus flexible. In general, the RMSF of ΔC6 Sto-RNase HI showed greater fluctuations than Sto-RNase HI at the same temperature simulations. These results are in accordance with our finding obtained from the essential dynamics analysis of simulations which will be discussed in a later section.

Secondary structure evolution

A close analysis of the time evolution of the secondary structure of Sto-RNase HI and it derivative ΔC6 Sto-RNase HI can present further information about their structural flexibilities. As a result, the dynamics consequences of C-terminal tail deletion in Sto-RNase HI can be thus addressed. The nature of the structural transitions during the simulations was investigated by examining the evolution of the secondary structure as a function of time (Fig. 4). Figure 4 shows that at both 300 and 375 K the secondary structure is quite conserved. The average number of residues that are conserved in β-sheets and the α-helix is given (Table 1) for comparison between the different simulations. As can be seen, the average number of residues in β-sheets and the α-helix at 300 and 375 K are similar, confirming minor structural differences. On the contrary, both Sto-RNase HI and ΔC6 Sto-RNase HI lose their original secondary structure elements considerably at 475 and 500 K (Fig. 4). Compared to Sto-RNase HI, the unfolding of ΔC6 Sto-RNase HI is much faster and already starts after 2–4 ns at 500 K. After 60 ns simulations, the secondary and, thus, the tertiary structure are almost completely lost in ΔC6 Sto-RNase HI whereas these of Sto-RNase HI are still partially structured. Moreover, the orders in which they lose their secondary elements are dissimilar to each other. The triple-stranded antiparallel β-sheet at N-terminal in ΔC6 Sto-RNase HI undergoes unfolding very quickly, while in Sto-RNase HI, the β-sheet begins unfolding after 80 ns and the β3-sheet seems to be stable within the time scale of the simulation retaining its structure to a larger extent. Helixes 2 and 4 exhibit a pronounced conformational flexibility in both Sto-RNase HI and ΔC6 Sto-RNase HI. For helixes 1 and 3, a lower stability

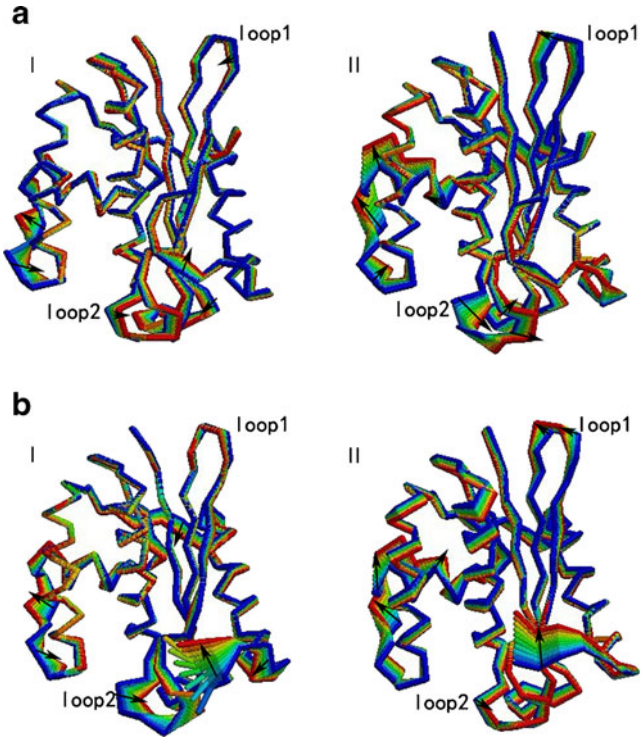

Fig. 5 Superimposition of ten configurations obtained by projecting the Cα motion onto the first eigenvector at 300 (I) and 375 K (II) of wild-type Sto-RNase HI (**a**) and ΔC6 Sto-RNase HI (**b**). Arrows are used as qualitative indicators of the motion direction

Table 2 Root mean square inner products (RMSIP) between the first 10 eigenvectors for MD simulation trajectories at different temperatures

	W300	W375	W475	W500	C300	C375	C475	C500
W300	0.689	0.689	0.399	0.323	0.452	0.412	0.341	0.289
W375		0.65	0.409	0.382	0.412	0.506	0.418	0.362
W475			0.420	0.426	0.435	0.334	0.394	0.365
W500				0.461	0.375	0.362	0.412	0.317
C300					0.774	0.698	0.543	0.335
C375						0.79	0.541	0.336
C475							0.508	0.337
C500								0.553

W stands for Sto-RNase HI and C for ΔC6 Sto-RNase HI

is observed for ΔC6 Sto-RNase HI compared to the wild-type protein. Remarkably, strands 4 and 5 are completely gone after 50 ns in ΔC6 Sto-RNase HI, while in the wild type they disappear partially after 60 ns but reform later. Most importantly, the β5-sheet in Sto-RNase HI seems to be stable within the time scale of the simulation, which is in agreement with the earlier experimental studies that the C-terminal tail plays a critical role in hyperstabilization of Sto-RNase HI [11]. In generally, the sequence of unfolding events is different in these two proteins. These different melting processes were also in qualitative agreement with the earlier experimental studies [11] which indicate lower thermal stability of ΔC6 Sto-RNase HI than Sto-RNase HI.

Hydrogen bonding is one of the dominant interactions in maintaining an intact secondary structure, which plays an important role in protein folding and stabilization [32]. The

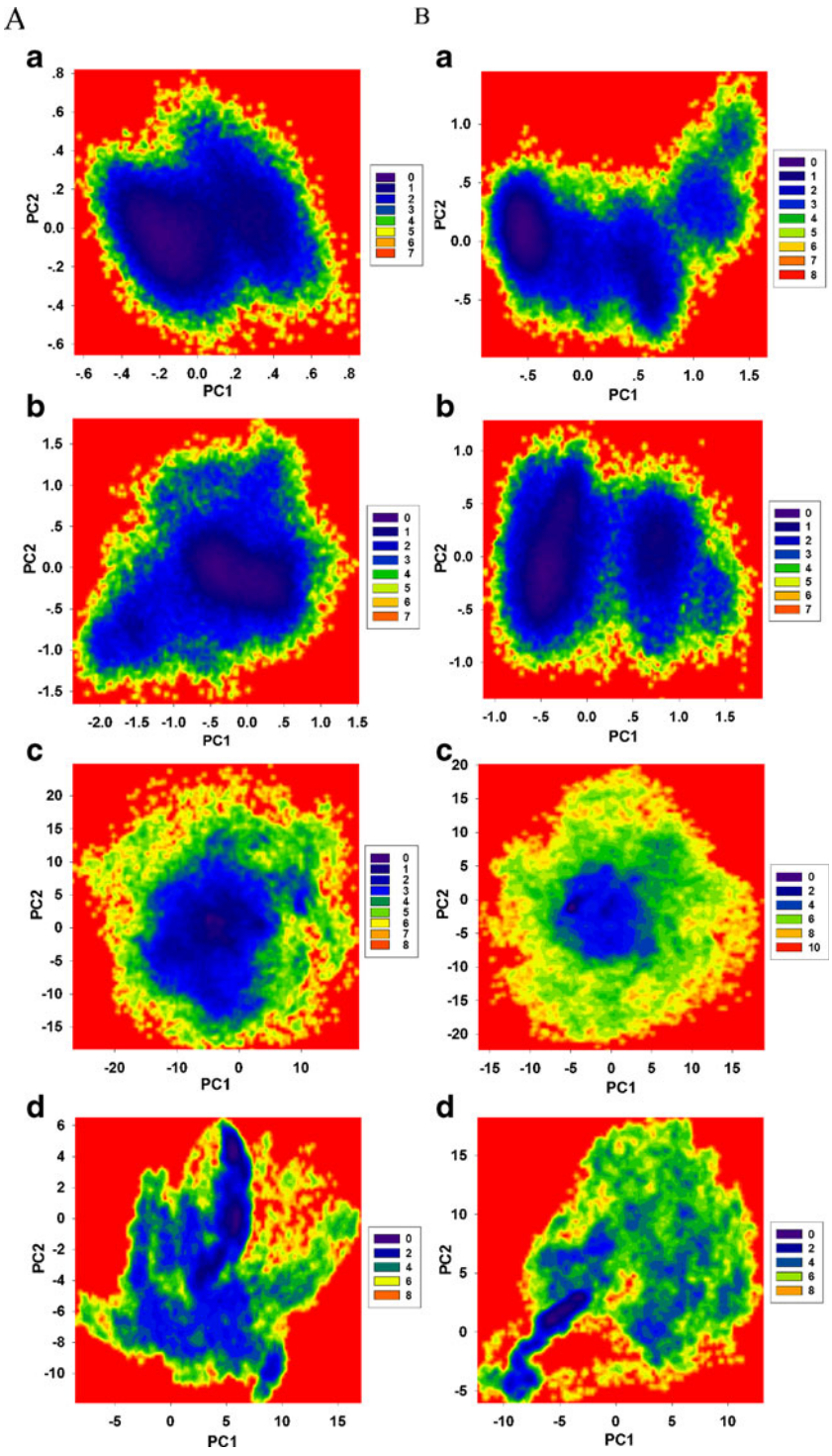

Fig. 6 The free energy landscapes of wild-type Sto-RNase HI (**A**) and ΔC6 Sto-RNase HI (**B**) versus the principal components PC1 and PC2 at 300 (**a**), 375 (**b**), 475 (**c**), and 500 (**d**) K, with deeper color indicating lower energy

analysis of the hydrogen bonds shows that there is a concomitant decrease in number of intact hydrogen bonds with temperature increasing for Sto-RNase HI and ΔC6 Sto-RNase HI, and Sto-RNase HI has more hydrogen bonds than ΔC6 Sto-RNase HI at the same temperature simulations (Table 1). On the basis of the above analysis, this is reasonable as the structure of ΔC6 Sto-RNase HI becomes more distorted as the simulation temperature is raised.

Essential dynamics analysis

The ED analysis on the Cα atoms was carried out to evaluate the dynamic properties of Sto-RNase HI and ΔC6 Sto-RNase HI in the simulations. The overall flexibility of Sto-RNase HI and ΔC6 Sto-RNase HI were evaluated by calculating the trace of the diagonalized covariance matrix of the atomic positional fluctuations. We have obtained the following values for Sto-RNase HI: 0.411 nm^2 at 300 K, 0.849 nm^2 at 375 K, 8.781 nm^2 at 475 K, and 10.57 nm^2 at 500 K. Similarly for ΔC6 Sto-RNase HI we have 0.544 nm^2 at 300 K, 0.955 nm^2 at 375 K, 9.267 nm^2 at 475 K, and 13.668 nm^2 at 500 K, thus confirming the overall increased flexibility of ΔC6 Sto-RNase HI than Sto-RNase HI with temperature increasing. The tendency of this kind of change can also be seen from the evolution of the secondary structure elements of Sto-RNase HI and ΔC6 Sto-RNase HI. In addition, the overall flexibility of ΔC6 Sto-RNase HI is larger than the corresponding flexibility of wild-type Sto-RNase HI at the same temperature. It is consistent with our previous analysis that the rate of unfolding is faster in ΔC6 Sto-RNase HI compared to Sto-RNase HI.

To estimate quantitatively the essential subspaces in Sto-RNase HI and ΔC6 Sto-RNase HI, we calculated RMSIP between the first 10 eigenvectors of each simulation (Table 2). Between the same proteins, the RMSIP values are high at lower temperatures (300 and 375 K) while they are significantly low at higher temperatures (475 and 500 K), indicating that the highest degree of overlap occurs within the individual trajectories obtained at lower temperatures (300 and 375 K). However, the RMSIP values of the two different proteins at lower as well as at higher temperatures are significantly low, implying that essential subspaces explored from the simulations of the two different proteins overlap insignificantly. In light of this, it may be hypothesized that these two proteins display different direction of motions at different simulated temperatures. To further characterize and visualize the dynamical differences between Sto-RNase HI and ΔC6 Sto-RNase HI, the first three eigenvector was analyzed at 300 and at 375 K (Fig. 5). Since the partial unfolding of Sto-RNase HI and ΔC6 Sto-RNase HI at high temperatures, the data are not analyzed at 475 and 500 K. As can been seen from Fig. 5, the dominant motion mainly involves the loop and both the C-terminal tails for Sto-RNase HI and ΔC6 Sto-RNase HI. The arrows indicate the dynamical behaviors of the three-dimensional correlated motion of the involved regions. Notably, the motion of loop1 and loop2 is different at 300 and 375 K in these two proteins. However, the amplitude of the fluctuations is higher in ΔC6 Sto-RNase HI. These results further verify RMSIP data. In summary, these results demonstrate that Sto-RNase HI and ΔC6 Sto-RNase HI show rather different principal directions of motions with different amplitudes.

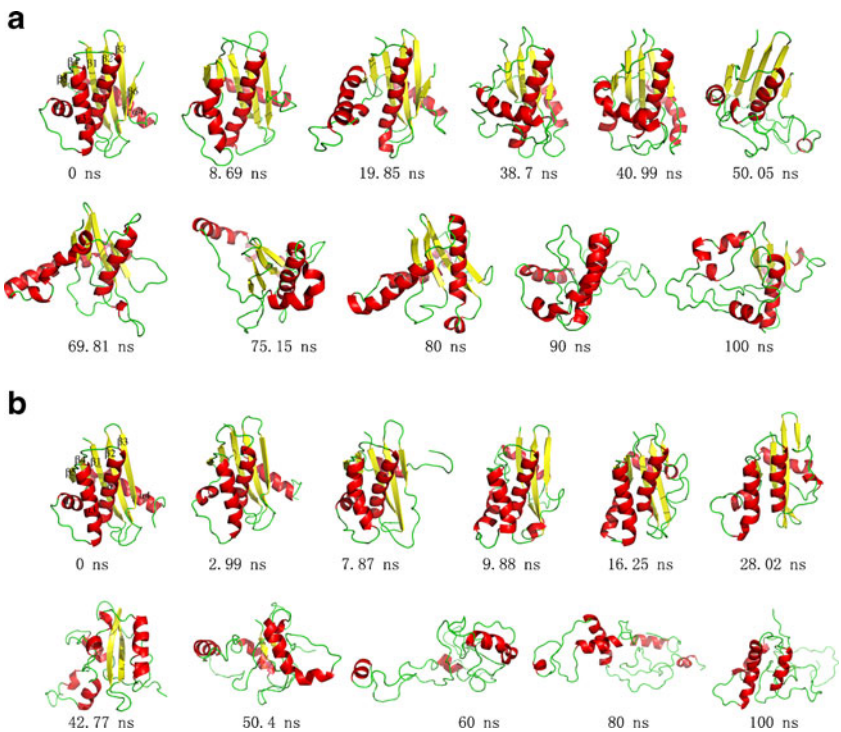

Fig. 7 Several snapshots of Sto-RNase HI (**a**) and ΔC6 Sto-RNase HI (**b**) during the unfolding simulations at 500 K. The time of each representative conformation are indicated

Unfolding pathway

The structural changes of Sto-RNase HI and ΔC6 Sto-RNase HI is investigated via characterization of the free energy landscapes of eight trajectories at different temperatures (Fig. 6). We can see that Sto-RNase HI and ΔC6 Sto-RNase HI have different free energy changes at four temperatures. The essential subspace and the structural fluctuations in these two proteins can be extremely different with increasing temperature. The deeper color regions at 475 and 500 K are smaller than that at other lower temperatures, indicating Sto-RNase HI and ΔC6 Sto-RNase HI undergo remarkable structural changes. Such large structural fluctuations at higher temperature may be associated with an initial unfolding process that changes the essential coordinate definition. The earlier experimental studies have also shown that the deletion of C-terminal tail leads to lower stability of ΔC6 Sto-RNase HI compared to Sto-RNase HI [9, 11]. We want to further investigate the differences in the stability of Sto-RNase HI and ΔC6 Sto-RNase HI using thermal unfolding studies by carrying out MD simulations at elevated temperatures. Thereby several representative snapshots at different times along the unfolding processes of the studied proteins are displayed (Fig. 7). As shown in Fig. 7a, the unfolding processes of the Sto-RNase HI began with the melting of β-strand at the C-terminal. This was followed by the unfolding of β5 around 19.85 ns. The distortion of the α4 happened at around 50 ns. The unfolding of the entire triple-stranded β-sheet happened at around 69.81 ns. The β3 strand firstly separated from the other two strands, which is followed by a complete collapse of the entire triple-strand β-sheet at about 90 ns. In fact, the distorted β3 strand did not induce the unfolding of the whole protein. From the snapshot at 100 ns, it is found that the distorted triple-stranded β-sheet was refolded. Compared to the wild-type protein, the unfolding of the ΔC6 Sto-RNase HI evolved much more quickly which was also indicated by the change of RMSD. Different to the Sto-RNase HI, the β3 strand was firstly melted in the mutant as shown in Fig. 7b and then a distortion of the C-terminal α-helix was observed around 7.87 ns. Similar to the Sto-RNase HI, the N-terminal strands were relatively stable and stayed until 50.4 ns. In brief, the most prominent differences between Sto-RNase HI and ΔC6 Sto-RNase HI are detected for the stability of the C-terminal strands 1, 2, and 3. There are several elements of secondary structure (helixes 1 and 3, strands 3 and 6) that are retained until the end of the simulation in Sto-RNase HI. The unfolding pathway analysis leads us to conclude that the rate of unfolding is faster in ΔC6 Sto-RNase HI compared to Sto-RNase HI and the sequence of unfolding events is different in these two proteins.

Conclusions

Focused on the hyperthermophilic protein Sto-RNase HI and its derivative lacking the C-terminal tail form (ΔC6 Sto-RNase HI), we performed MD simulations at four different temperatures (300, 375, 475, and 500 K) to investigate the effect of C-terminal tail on the structure and stability of Sto-RNase HI.

Detailed analysis of these MD trajectories in terms of secondary structure content, solvent accessibility, intramolecular hydrogen bonds indicate that the deletion of the C-terminal tail in Sto-RNase HI substantially changes the structure and overall fluctuations with respect to the wild type Sto-RNase HI structure. The unfolding studies of Sto-RNase HI and ΔC6 Sto-RNase HI suggest that the rate of unfolding is faster in ΔC6 Sto-RNase HI compared to Sto-RNase HI and the sequence of unfolding events is different in these two proteins. The overall flexibility calculated by the trace of the diagonalized covariance matrix displays higher flexibility of ΔC6 Sto-RNase HI than Sto-RNase HI at the same temperature. The ED analysis also indicates that Sto-RNase HI and ΔC6 Sto-RNase HI exhibit different principal directions of motion.

MD simulations and ED analysis provide the above structural and dynamic properties, which are inaccessible in the static crystal structure. This study can provide argument for further experimental and theoretical investigation into Sto-RNase HI.

Acknowledgments This work is supported by Natural Science Foundation of China and Specialized Research Fund for the Doctoral Program of Higher Education (Grant Nos. 21273095, 20903045, 21203072, and 20070183046).

References

1. Vieille C, Zeikus GJ (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. Microbiol Mol Biol Rev 65(1):1–43
2. Turner P, Mamo G, Karlsson EN (2007) Potential and utilization of thermophiles and thermostable enzymes in biorefining. Microb Cell Fact 6(9):1–23
3. Bruins ME, Janssen AEM, Boom RM (2001) Thermozymes and their applications. Appl Biochem Biotech 90(2):155–186
4. Zhou HX (2004) Polymer models of protein stability, folding, and interactions. Biochemistry 43(8):2141–2154
5. Suzuki T, Iwasaki T, Uzawa T, Hara K, Nemoto N, Kon T, Ueki T, Yamagishi A, Oshima T (2002) Sulfolobus tokodaii sp. nov. (f. Sulfolobus sp. strain 7), a new member of the genus Sulfolobus isolated from Beppu Hot Springs, Japan. Extremophiles 6(1):39–44
6. Kawarabayasi Y, Hino Y, Horikawa H, Jin-No K, Takahashi M, Tanaka T, Kikuchi H (2001) Complete genome sequence of an aerobic thermoacidophilic. DNA Res 8(4):123–140
7. Arudchandran A, Cerritelli S, Narimatsu S, Itaya M, Shin D, Shimada Y, Crouch R (2000) The absence of ribonuclease H1 or H2 alters the sensitivity of Saccharomyces cerevisiae to hydroxyurea,

caffeine and ethyl methanesulphonate: implications for roles of RNases H in DNA replication and repair. Genes Cells 5:789–802
8. Qiu J, Qian Y, Frank P, Wintersberger U, Shen B (1999) Saccharomyces cerevisiae RNase H (35) functions in RNA primer removal during lagging-strand DNA synthesis, most efficiently in cooperation with Rad27 nuclease. Mol Cell Biol 19(12):8361–8371
9. You DJ, Chon H, Koga Y, Takano K, Kanaya S (2007) Crystal structure of type 1 ribonuclease H from hyperthermophilic archaeon Sulfolobus tokodaii: role of arginine 118 and C-terminal anchoring. Biochemistry 46(41):11494–11503
10. Ohtani N, Yanagawa H, Tomita M, Itaya M (2004) Cleavage of double-stranded RNA by RNase HI from a thermoacidophilic archaeon, Sulfolobus tokodaii 7. Nucleic Acids Res 32(19):5809–5819
11. Takano K, Okamoto T, Okada J, Tanaka S, Angkawidjaja C, Koga Y, Kanaya S (2011) Stabilization by fusion to the C-terminus of hyperthermophile Sulfolobus tokodaii RNase HI: a possibility of protein stabilization tag. PLoS One 6(1):e16226
12. Nguyen TN, Angkawidjaja C, Kanaya E, Koga Y, Takano K, Kanaya S (2012) Activity, stability, and structure of metagenome-derived LC11-RNase H1, a homolog of Sulfolobus tokodaii RNase H1. Protein Sci 21:553–561
13. Li L, Li Y, Zhang L, Hou T (2012) Theoretical studies on the susceptibility of oseltamivir against variants of 2009 A/H1N1 influenza neuraminidase. J Chem Inf Model 52(10):2715–2729
14. Ren Y, Gao J, Ge W, Li J (2008) Thermal unfolding of a double-domain protein: molecular dynamics simulation of Rhodanese. Ind Eng Chem Res 48(19):8865–8871
15. Nasiri R, Bahrami H, Zahedi M, Moosavi-Movahedi AA, Sattarahmady N (2010) A theoretical elucidation of glucose interaction with HSA's domains. J Biomol Struct Dyn 28(2):211–226
16. Wiesner J, Kříž Z, Kuča K, Jun D, Koča J (2010) Influence of the acetylcholinesterase active site protonation on omega loop and active site dynamics. J Biomol Struct Dyn 28(3):393–403
17. Guo J, Ren H, Ning L, Liu H, Yao X (2012) Exploring structural and thermodynamic stabilities of human prion protein pathogenic mutants D202N, E211Q and Q217R. J Struct Biol 178(3):225–232
18. Bohr HG (2009) Handbook of molecular biophysics: methods and applications. Wiley-VCH, Weinheim
19. Papaleo E, Pasi M, Tiberti M, De Gioia L (2011) Molecular dynamics of mesophilic-like mutants of a cold-adapted enzyme: insights into distal effects induced by the mutations. PLoS One 6(9):e24214
20. Garofoli S, Falconi M, Desideri A (2004) Thermophilicity of wild type and mutant cold shock proteins by molecular dynamics simulation. J Biomol Struct Dyn 21(6):771–780
21. Merlino A, Graziano G, Mazzarella L (2004) Structural and dynamic effects of α–Helix deletion in Sso7d: Implications for protein thermal stability. Proteins 57(4):692–701
22. Sham YY, Ma B, Tsai CJ, Nussinov R (2008) Molecular dynamics simulation of Escherichia coli dihydrofolate reductase and its protein fragments: relative stabilities in experiment and simulations. Protein Sci 10(1):135–148
23. Chu WT, Wu YJ, Zhang JL, Zheng QC, Chen L, Xue Q, Zhang HX (2012) Constant pH Molecular Dynamics (CpHMD) and mutation studies: insights into AaegOBP1 pH-induced ligand releasing mechanism. BBA-Proteins Proteom 1824:913–918
24. Van der Spoel D, van Maaren PJ, Berendsen HJC (1998) A systematic study of water models for molecular simulation: derivation of water models optimized for use with a reaction field. J Chem Phys 108(24):10220–10230
25. Berendsen HJC, van der Spoel D, van Drunen R (1995) GROMACS: a message-passing parallel molecular dynamics implementation. Comput Phys Commun 91(1):43–56
26. Hess B, Kutzner C, Van Der Spoel D, Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput 4(3):435–447
27. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T (2003) A point–charge force field for molecular mechanics simulations of proteins based on condensed–phase quantum mechanical calculations. J Comput Chem 24(16):1999–2012
28. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 23(3):327–341
29. Berendsen HJC, Postma JPM, Van Gunsteren WF, DiNola A, Haak J (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81:3684–3690
30. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: An N·log (N) method for Ewald sums in large systems. J Chem Phys 98:10089–10092
31. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A smooth particle mesh Ewald method. J Chem Phys 103:8577–8593
32. Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen–bonded and geometrical features. Biopolymers 22(12):2577–2637
33. Van der Spoel D, Lindahl E, Hess B, Van Buuren A, Apol E, Meulenhoff P, Tieleman D, Sijbers A, Feenstra K, van Drunen R (2008) Gromacs user manual version 3.3. www.gromacs.org
34. DeLano WL (2002) The PyMOL user's manual. DeLano Scientific, San Carlos, p 382
35. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera—a visualization system for exploratory research and analysis. J Comput Chem 25(13):1605–1612
36. Amadei A, Linssen A, Berendsen HJC (1993) Essential dynamics of proteins. Proteins 17(4):412–425
37. Amadei A, Ceruso MA, Di Nola A (1999) On the convergence of the conformational coordinates basis set obtained by the essential dynamics analysis of proteins' molecular dynamics simulations. Proteins 36(4):419–424
38. Ceruso MA, Amadei A, Nola AD (1999) Mechanics and dynamics of B1 domain of protein G: role of packing and surface hydrophobic residues. Protein Sci 8(1):147–160
39. Chan TF, Mulet P (1999) On the convergence of the lagged diffusivity fixed point method in total variation image restoration. SIAM J Numer Anal 36(2):354–367
40. Van Aalten D, Conn D, De Groot B, Berendsen H, Findlay J, Amadei A (1997) Protein dynamics derived from clusters of crystal structures. Biophys J 73(6):2891–2896
41. Daidone I, Amadei A (2012) Essential dynamics: foundation and applications. WIREs Comput Mol Sci 2:762–770