Thermal stability and unfolding pathways of hyperthermophilic and mesophilic periplasmic binding proteins studied by molecular dynamics simulation

Lin Chen, Xue Li, Ruige Wang, Fengqin Fang, Wanli Yang and Wei Kan

aCollege of Chemistry and Chemical Engineering, Qiqihar University, Qiqihar 161006, P.R. China; bCollege of Chemistry and Biology, Beihua University, Jilin 132000, P.R. China

Communicated by Ramaswamy H. Sarma

(Received 20 May 2015; accepted 15 August 2015)

The ribose binding protein (RBP), a sugar-binding periplasmic protein, is involved in the transport and signaling processes in both prokaryotes and eukaryotes. Although several cellular and structural studies have been reported, a description of the thermostability of RBP at the molecular level remains elusive. Focused on the hyperthermophilic Thermoytoga maritima RBP (tmRBP) and mesophilic Escherichia coli homolog (ecRBP), we applied molecular dynamics simulations at four different temperatures (300, 380, 450, and 500 K) to obtain a deeper insight into the structural features responsible for the reduced thermostability of the ecRBP. The simulations results indicate that there are distinct structural differences in the unfolding pathway between the two homologs and the ecRBP unfolds faster than the hyperthermophilic homologs at certain temperatures in accordance with the lower thermal stability found experimentally. Essential dynamics analysis uncovers that the essential subspaces of ecRBP and tmRBP are non-overlapping and these two proteins show different directions of motion within the simulations trajectories. Such an understanding is required for designing efficient proteins with characteristics for a particular application.

Keywords: hyperthermophilic; mesophilic; molecular dynamics simulations; unfolding pathway; essential dynamics analysis

Introduction

Proteins from hyperthermophilic organisms are very similar to their homologs proteins from mesophiles in terms of both their sequences and structures, but the hyperthermophilic proteins are more resistant to thermal denaturation at high temperatures (Chakravarty & Varadarajan, 2002; Sadeghi, Naderi-Manesh, Zarrabi, & Ranjbar, 2006). This paradoxical situation offers a unique opportunity for gaining an insight into the thermal stability mechanism of hyperthermophilic proteins. Thus, these differences in amino acid composition, structure, and thermal stability between hyperthermophilic and mesophilic proteins have been the subject of intensive experimental and theoretical studies (Bruins, Janssen, & Boom, 2001; Chakravarty & Varadarajan, 2002; Chen et al., 2013; Grottesi, Ceruso, Colosimo, & Di Nola, 2002; Khechinashvili, Kabanov, Kondratyev, & Polozov, 2014; Kumar, Ma, Tsai, & Nussinov, 2000; Mahdavi, Sajedi, Asghari, Taghdir, & Rassa, 2011; Tang & Liu, 2007). Understanding the molecular basic of thermostability of hyperthermophilic proteins plays a critical role in designing efficient enzymes with enhanced performance for industrial catalytic process (Bruins et al., 2001; Maugini, Tronelli, Bossa, & Pascarella, 2009; Turner, Mamo, & Karlsson, 2007). However, the structural similarity between hyperthermophilic and their mesophilic analogs makes the identification of the determinants responsible for the thermal stability more difficult (Vieille & Zeikus, 2001). Molecular dynamics (MD) simulation is a powerful tool to complement experimental results with atomics and instantaneous details. Hence, atomistic explorations of the structural and dynamic properties of these proteins are of vital importance for the understanding of the factors involved in the stabilization of proteins in general (Sadeghi et al., 2006). Compared to their mesophilic homologs, hyperthermophilic proteins have uncovered many pronounced features including increased hydrogen bonds, salt bridges, hydrophobicity in proteins’ cores, decreased number and volume of internal cavities, and hydrophobic interactions (Chakravarty & Varadarajan, 2002; Kumar et al., 2000; Kumar, Tsai, & Nussinov, 2000; Priyakumar, 2012; Rathi, Höffken, & Gohlke, 2014; Roche et al., 2012; Sadeghi et al., 2006; Vogt, Woell, & Argos, 1997). However, there are no general factors for determining thermostability and it is obvious that different proteins gain their thermostability owing to different combinations of the above factors. To attain a general
understanding, many experiments and theoretical researches have concentrated on the folding-unfolding process of various proteins depending on temperature (Day, Paschek, & Garcia, 2010; Grottesi et al., 2002; Ilizaliturri-Flores, Correa-Basurto, Benítez-Cardoza, & Zamorano-Carrillo, 2014; Meharenna & Poulos, 2010; Priyakumar, 2012; Priyakumar, Ramakrishna, Nagarjuna, & Reddy, 2010; Tang & Liu, 2007). However, microscopic structures of proteins and trapping of dynamic intermediates in this process have always been intractable because of the tiny tempo-spatial scales. Due to computational limitations, most investigations are limited to single-domain proteins, or isolated domains of larger proteins, or even polypeptide fragments. However, most of the proteins comprise more than one independently folding domains. Thus, the analysis of a multidomain protein can reveal important information on the stability of domains as well as the secondary structural changes within the domains.

The ribose binding protein (RBP) is a sugar-binding protein found in the periplasmic space of some bacteria (Björkman & Mowbray, 1998; Boos & Shuman, 1998). These proteins are receptors for extracellular solutes in chemotaxis and intercellular communication processes. RBP from the hyperthermophilic *Thermoytoga maritima* (tmRBP) determined by X-ray diffraction is a 291-amino acid protein (Cuneo, Beese, & Hellenga, 2008), consisting of two structurally homologous domains linked by a three-stranded hinge. The three strands that connect the two domains serve as a flexible hinge that allows opening and closing of the cleft between the two domains, which also form the ribose-binding site. The thermostable homolog of RBP from mesophilic *Escherichia coli* (ecRBP) consisting of 271 residues has also been determined by X-ray diffraction (Björkman & Mowbray, 1998). These two proteins show 39% sequence identity (Figure 1) and the root mean square deviation of Ca atoms of superimposed structures is 1.0 Å. The main difference between ecRBP and tmRBP is at the C-terminus where tmRBP is 13 residues longer than ecRBP. Despite a very similar structure between tmRBP and ecRBP, they have very different temperature dependencies in the stability and activity. The thermal denaturation experiment using circular dichroism exhibited that the tmRBP is a hyperthermostable protein with a denaturation temperature of 108, 52°C higher than that of ecRBP. Although several cellular and structural studies have been reported, a description of the thermal stability of RBP at the molecular level remains elusive. Therefore, the stability and unfolding pathways of these two homologous proteins have been studied based on MD simulations done at higher temperatures in this paper.

In the current paper, two homologs proteins, tmRBP and ecRBP, are studied by MD simulations at 300, 380, 450, and 500 K to examine the structural features responsible for the thermostability and unfolding pathways. Then, the essential dynamics (ED) method was applied to understand the detailed dynamics properties. These characteristics can help us to better understand the thermostability as well as secondary structural changes within the multidomain. During this study, we have specifically attempted to answer the following by examining the structural and dynamic aspects: (a) characterizing the structural and thermal stability for the two homologous proteins: tmRBP and ecRBP, (b) evaluating quantitatively the essential subspace in tmRBP and ecRBP by computing the root mean square inner product, and (c) elucidating the differences in the thermal unfolding pathways of tmRBP and ecRBP.

**Materials and methods**

**MD simulations**

The starting structures for tmRBP and ecRBP were obtained from the Protein Data Bank entries 2FN9 (Cuneo et al., 2008) and 1URP (Björkman & Mowbray, 1998), respectively. The crystallographic water molecules were removed from the starting structures. The proteins were immersed in a rectangular box with a 1.0-nm buffer. The box was filled with SPC water molecules (van der Spoel, van Maaren, & Berendsen, 1995) with periodic boundary conditions. The appropriate number of sodium ions was added to neutralize the proteins’ charges. The resulting systems were subjected to energy minimization using the descent method plus the conjugate gradients method to remove bad hydrogen contacts. Subsequently, each was heated gradually from 100 K to the desired temperatures (300, 380, 450, and 500 K) during 500 ps. The thermostat algorithm used is the Berendsen thermostat (Berendsen, Postma, van Gunsteren, DiNola, & Haak, 1984) algorithm in constant NVT ensembles. And then equilibration was done in NPT ensembles under Berendsen thermostat for 500 ps. Finally, production MD simulations for the equilibrated systems were performed for 100 ns in an NPT ensemble under Berendsen thermostat.

MD simulations were carried out by the GROMACS program (Berendsen, van der Spoel, & van Drunen, 1995; Hess, Kutzner, van der Spoel, & Lindahl, 2008) with the AMBER99SB-LIDN force field (Ryckaert, Ciccotti, & Berendsen, 1977). All bonds involving hydrogen atoms were constrained by the use of the LINCS algorithm (Ryckaert et al., 1977). The time step of 2 fs was used and the temperature was maintained by the Berendsen coupling algorithm (Berendsen et al., 1984) at 1 atm using a Parrinello–Rahman barostat with a .1 ps time constant in all eight MD simulations. The long-range electrostatic interactions were calculated by the particle mesh Ewald (Darden, York, & Pedersen, 1993; Essmann et al., 1995) summation scheme. Both
the van der Waals and Coulomb interactions were truncated at 1.2 nm. Secondary structure analysis was performed using the program DSSP (Kabsch & Sander, 1983). Other analyses were performed using script included with the GROMACS (van der Spoel et al., 2008) distribution. Sequence alignments were carried out using the sequence of the RBP proteins from the hyper- thermophilic *Thermoytoga maritima* and mesophilic *E. coli* using the ClustalW sever (Thompson et al., 1994). The PyMOL (DeLano, 2002) and Chimera (Pettersen et al., 2004) software were used to visualize the trajectories and to produce the pictures of the molecular structures.

**ED analysis**

According to ED approach (Amadei, Ceruso, & Di Nola, 1999; Amadei, Linssen, & Berendsen, 1993; Chan & Mulet, 1999; Daidone & Amadei, 2012), essential degrees of freedom of both tmRBP and ecRBP were extracted from the equilibrated portion of the MD trajectories. The ED analysis was based on the construction of the covariance matrix of the coordinate fluctuations. The covariance matrix was calculated by removing the translational and rotational degrees of freedom. After that this matrix was diagonalized to extract the eigenvectors and eigenvalues, which give information about correlated motions throughout the protein. The eigenvectors show the directions of the motion in the protein, while the eigenvalues show the amount of motion along their eigenvector. Thereafter, the eigenvectors were sorted according to their eigenvalues in descending order. In general, the first 10 eigenvectors describe almost all conformational substrates accessible to the protein.

For the simulations of both tmRBP and ecRBP, only Cα atoms were comprised in the computation of the covariance matrices. The root mean square inner product (RMSIP) (Ceruso, Amadei, & Nola, 1999) was computed to estimate the degree of overlap between the essential subspaces of the two simulated systems at both the same and different temperatures. In this method, the RMSIP is computed by the following equations:

\[
\text{RMSIP} = \sqrt{\frac{1}{10} \sum_{i=1}^{10} \sum_{j=1}^{10} (\eta^a_i \eta^b_j)}
\]

Here, \(\eta^a_i\) and \(\eta^b_j\) are the \(i\)th and \(j\)th eigenvectors for the systems of \(a\) and \(b\), respectively. 10 means the first 10 eigenvectors. The RMSIP value was computed from the equilibrated portion of the simulation trajectory. For diagonal elements, the RMSIP value was computed from two equal halves of the simulation trajectory. The ED analysis was performed with the tool available in the GROMACS software package.
Results and discussion

Structural stability of the tmRBP and ecRBP

To investigate the structural changes in tmRBP and ecRBP, four parallel simulations were performed for the pair at temperatures 300, 380, 450, and 500 K in explicit water. The root mean square deviations (RMSD) of the Cα atoms with respect to the experimental structure were obtained for the two proteins (Figure 2 and Table 1). As can be seen, the structures from the 300 K simulations show small deviations from the crystal structures exhibiting average RMSD values of .28 and .25 nm, respectively, for ecRBP and tmRBP (Table 1). In the trajectory run at 380 K, the RMSD increases up to an equilibrium value of .54 and .37 nm for ecRBP and tmRBP, respectively. Thus, for these two proteins, changing the temperature up to 380 K exhibits different influences on the stability of their structure in the timescale of the simulations, validating the present experiments’ results. At 450 and 500 K simulations, the RMSD values of both ecRBP and tmRBP increase gradually during the 100 ns simulations and exhibit a maximum deviation of 2.1 nm (Figure 2), indicating that the structures of the two proteins undergo a constant evolution at high temperatures. However, tmRBP displays lower RMSD values than that of ecRBP, implying the unfolding transitions of the tmRBP occurred much later than the ecRBP. This is in agreement with the observed reduced stability of the ecRBP (Cuneo et al., 2008). Overall, the structures of tmRBP and ecRBP are in the folded state and are closer to the crystal structures at 300 and 380 K in the timescale of the simulations, while the structures undergo major structural changes leading to unfolded states at 450 and 500 K. These changes can also be seen from the time evolution of the secondary structure of ecRBP and tmRBP at four temperatures (Figure 5 and Supplementary Figure S1), which will be discussed in a later section. In addition, the solvent accessible surface area (SASA) for tmRBP and ecRBP was calculated to explore the structural changes of these two proteins (Table 1). As can be seen from Table 1, the hydrophobic SASA from the 300 K simulations for ecRBP and tmRBP is 68.19 and 70.52 nm², respectively. The hydrophobic SASA increases up to the values of 72.50 and 72.79 nm², respectively, for ecRBP and tmRBP at 380 K, showing slight structural changes for ecRBP. At 450 and 500 K simulations, the hydrophobic SASA of ecRBP and tmRBP increases up to maximum average values of 84.91 and 87.43 nm², respectively. In a word, tmRBP has a higher hydrophobic SASA compared to ecRBP at four temperatures simulations, which contributes to the high stability of the tmRBP. Such high stability of tmRBP is consistent with our previous analysis that the unfolding transitions of the tmRBP occurred much later than the ecRBP at high temperature simulations.

Furthermore, to characterize the intramolecular interactions in these two homologous proteins, the contact maps were drawn by calculating the averages of the inter-residue distances at four different temperatures. Since the proteins undergo major structural changes at 450 and 500 K evidenced from RMSD calculations, the calculations were based on the trajectories obtained from the last 40 ns of the simulations. In general, there are gradual changes in the interactions of inter-residue with increasing temperature (Figure 3). As shown in Figure 3(a) and (b), the proximity of residues in β1 and α9, α1 and α9, β3 and α9, β4 and α9, α8 and β10, α8 and β11, β9 and β10, β8 and β9, and α4 and β9 is quite similar at both 300 and 380 K. Additionally, other tertiary contacts β1 and β2, β1 and β3, β2 and β4, β4 and β5, β1 and α2, β2 and α3, β3 and α3, β4 and α3, α4 and α5, α4 and β8, α4 and β9, α5 and β6, α5 and β7, α6 and β6, β6 and β7, α6 and β7, a7 and β7, a8 and β8, and α8 and β9 are well characterized at 300 and 380 K consistent with experimental data. However, a quick look at Figure 3(c) and (d) reveals that the dimensional structures of these two proteins from 450 to 500 K simulations are extremely different from the starting structure, indicating denaturation of ecRBP and tmRBP. As the protein undergoes unfolding, most of the interactions of
inter-residue listed above are destabilized, and the inter-residue distances increase to more than 1.5 nm. However, the tertiary contacts α4 and α5, β6 and α5, and β10 and α9 for tmRBP are quite stable at 500 K in the time-scale of our simulations. Overall, there is no obvious change in the inter-residue distances at 300 and 380 K simulations. This is in good agreement with our previous analysis that these two proteins retain their initial structure at 380 K within the length of the simulations in this study. On the basis of the above analysis, we can deduce that low-temperature simulations (300 and 380 K) do not have enough energy to overcome the barrier around the starting conformation in our limited time simulations. However, high-temperature unfolding simulations do show the rapid denaturation of the native structure of tmRBP and ecRBP.

**Dynamics features of the tmRBP and ecRBP**

To explore the flexibility of the structural elements in the two homologous proteins, the dynamics of the proteins at different temperatures were studied based on the root mean square fluctuations (RMSF) of the protein Cα with respect to the average coordinates (Figure 4). As can be seen from Figure 4, the overall fluctuations obtained at 300 K are similar in tmRBP and ecRBP, and the RMSF values increase marginally for ecRBP at 380 K. The difference in the flexibility profiles of ecRBP at 300 and 380 K is in good agreement with our previous analysis that ecRBP exhibits marginally large structural changes within the simulation performed in this study, while the flexibility patterns of the residues of tmRBP and ecRBP are qualitatively similar in both 300 and 380 K simulations. As expected, the loop and terminal regions are more flexible compared to the regular secondary structure regions of the proteins, since they do not have a large network of tertiary interactions, while the C-terminals of these two proteins exhibit a significant sensitivity to thermal stress and the N-terminals were considerably stable. However, the RMSF values calculated at 450 and 500 K are significantly large indicating the sampling of larger conformational space at these two temperatures. This is owing to the loss of secondary structure at these two high temperatures (450 and 500 K). In a word, the RMSF values of ecRBP exhibit greater fluctuations than tmRBP at the same temperature simulations. These results can also be seen from the ED analysis of MD simulations which will be discussed in a later section.

**Secondary structure evolution of the tmRBP and ecRBP**

A much better understanding of the time evolution of the secondary structure of tmRBP and ecRBP can give further information about their structural flexibilities. As a consequence, the difference of dynamics consequences in tmRBP and ecRBP can be addressed. Thus, the nature of the structural transitions during the simulations was illustrated by investigating the evolution of the secondary structure as a function of time (Figure 5 and Supplementary Figure S1) and the average number of residues for various secondary structural regions is reported (Table 2). As can be seen from Table 2, tmRBP exhibits a lower content of β-sheets, α-helices, and a great number of residues in coil-like conformation than ecRBP at 300 K. At 380 K, tmRBP exhibits a higher content of α-helices, and a few number of residues in coil-like conformation and β-sheets than ecRBP. This may be due to the fact that several coils have transformed into α-helices during simulation. On the contrary, these two proteins present a lower content of β-sheets and α-helices at high temperatures simulations (450 and 500 K), indicating major structural changes with respect to the experimental structure. These changes also can be seen from Supplementary Figure S1, the secondary structure of tmRBP and ecRBP is quite conserved at 300 K simulation. At 380 K simulation, the sheet 11, helices 2, 3, and 8 in ecRBP begin unfolding after 5 ns, while helix 8 in tmRBP begins unfolding after 60 ns. These lead to the conclusion that these two proteins, changing temperature up to 380 K, show different influences on the stability of its structure during the 100 ns MD simulations, which is consistent with the results from experiments. As can be seen from Figure 5, both tmRBP and ecRBP lose their original secondary
Figure 3. Contact maps indicating the distances between the center of masses of each residue and the other residues of ecRBP (A) and tmRBP (B) at four temperatures.
structure elements considerably at 450 and 500 K in our simulations. Compared to tmRBP, the unfolding of ecRBP is much faster and already starts after 2 ns at 500 K. After 5 ns simulations, the secondary and, thus, the tertiary structure are almost completely lost in ecRBP, whereas these in tmRBP are still partially structured. This can be seen from the unfolding of β1. Moreover, the orders in which they lose their secondary elements are dissimilar to each other. As shown in Figure 5 at 450 K simulations, the helices 3 and 8 in ecRBP undergo unfolding first; while in tmRBP, the helix 8 undergoes unfolding first and the helix 3 is completely gone after 30 ns. Helices 4, 5, and 9 in ecRBP begin unfolding after 9 ns, whereas in tmRBP, helix 9 begins unfolding after 50 ns and helix 4 seems to be stable within the time scale of the simulation retaining its structure to a larger extent. Remarkably, strands 10 and 11 are completely gone after 20 ns in tmRBP; while in ecRBP, they disappear partially after 10 ns, but reform later. Most importantly, the helices 6 and 7 in tmRBP undergo unfolding after 22 ns simulations, while in ecRBP, they disappear partially after 70 ns. In general, 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 which indicate a lower thermal stability of ecRBP than tmRBP.

Figure 4. Root mean square fluctuations of residues of ecRBP (a) and tmRBP (b) at four temperatures. The β-strands and the α-helix are represented by block arrows and wavy arrows, respectively.

Figure 5. Time evolution of the secondary structure elements of ecRBP (a,b) and tmRBP (c,d) at 450 and 500 K. The β-strands and the α-helix are represented by β and α, respectively.
Hydrogen bonding is one of the dominant interactions to maintain an intact secondary structure, which plays an important role in protein folding and stabilization. Thus, the average number of intramolecular hydrogen bonds (H-bonds) is calculated at different temperature simulations for tmRBP and ecRBP (Table 1). As can be seen, the hydrogen bonds decrease from: (a) about 205 at 300 K to about 137 at 380 K, about 88 at 450 K, and about 66 at 500 K in the case of ecRBP; (b) about 161 at 300 K to about 152 at 380 K, about 102 at 450 K, and about 66 at 500 K in the case of tmRBP. In general, there is a concomitant decrease in the number of intact hydrogen bonds with an increase in the temperature for ecRBP than tmRBP, and tmRBP has more hydrogen bonds than ecRBP at the same temperature simulations. On the basis of the above analysis, this is reasonable as the structure of ecRBP becomes more distorted compared to tmRBP as the simulation temperature increases.

Salt bridges in hyperthermophilic and thermophilic proteins were considered as the primary influential factors on their high thermal stability compared to their mesophilic counterparts (Folch, Rooman, & Dehouck, 2008). To better understand the effects of salt bridges on the thermal stability of ecRBP and tmRBP in detail, the probability distributions of all possible N(Lys/Arg)-O(Asp/Glu) interatomic distances were computed at four different temperatures (Figure 6). In general, similar probability distribution curves were obtained for these distances in ecRBP and tmRBP at the same temperature. However, the probabilities corresponding to the shorter distances (less than .6 nm) in ecRBP (Figure 6(a)) are lower than that in tmRBP (Figure 6(b)). This suggests that salt bridges contribute to the thermal stability of tmRBP at the higher temperature. In addition, a significant increase in the probabilities corresponding to the longer distances (more than .6 nm) in ecRBP and tmRBP was observed at high temperature, which indicates that several salt bridges are disrupted and several new salt bridges are formed during denaturation of ecRBP and tmRBP in our simulations. In a word, salt bridges play an important role on the thermal stability of tmRBP compared to ecRBP.

| Table 2. Average secondary structure contents in ecRBP and tmRBP obtained at four different temperatures simulations. |
|-------------------------------|------------------|------------------|------------------|------------------|
|                               | 300 K            | 380 K            | 450 K            | 500 K            |
| IURP                          |                  |                  |                  |                  |
| coil                          | 44.98 [16.59%]   | 52.39 [19.33%]   | 87.08 [32.14%]   | 126.81 [46.79%]  |
| β-sheet                       | 60.58 [22.35%]   | 60.25 [22.23%]   | 62.29 [22.97%]   | 22.17 [8.18%]    |
| β-bridge                      | .62 [.22%]       | 3.81 [.141%]     | 7.58 [2.80%]     | 16.01 [5.91%]    |
| Bend                          | 29.91 [11.03%]   | 32.35 [11.94%]   | 53.96 [19.91%]   | 76.77 [28.33%]   |
| Turn                          | 15.89 [5.87%]    | 23.20 [8.56%]    | 24.13 [8.90%]    | 26.25 [9.69%]    |
| α-Helix                       | 118.82 [43.85%]  | 97.05 [35.81%]   | 33.31 [12.29%]   | 1.20 [.44%]      |
| 30-Helix                      | .097 [.036%]     | .96 [.35%]       | 1.32 [.49%]      | .77 [.28%]       |
| 2FN9                          |                  |                  |                  |                  |
| coil                          | 53.15 [18.26%]   | 54.49 [18.73%]   | 88.97 [30.57%]   | 133.88 [46.01%]  |
| β-Sheet                       | 57.39 [19.72%]   | 62.02 [21.31%]   | 55.87 [19.20%]   | 25.76 [8.85%]    |
| β-Bridge                      | .77 [.26%]       | 1.22 [.42%]      | 7.57 [2.60%]     | 16.01 [5.50%]    |
| Bend                          | 31.47 [10.81%]   | 31.56 [10.84%]   | 55.22 [18.98%]   | 83.18 [25.58%]   |
| Turn                          | 22.62 [7.78%]    | 23.57 [8.10%]    | 26.66 [9.16%]    | 28.80 [9.90%]    |
| α-Helix                       | 124.94 [42.93%]  | 116.77 [40.13%]  | 50.94 [17.51%]   | 1.86 [0.64%]     |
| 30-Helix                      | .57 [.20%]       | .87 [.24%]       | 2.88 [.99%]      | .75 [.26%]       |

Figure 6. Probability distributions corresponding to the interatomic distance of N(Lys/Arg)-O(Asp/Glu) of ecRBP (a) and tmRBP (b).
Table 3. Average interaction energies (kcal/mol) of the main secondary structure elements in ecRBP at four different temperatures simulations.

| 1URP | 300 K       | 380 K       | 450 K       | 500 K       |
|------|-------------|-------------|-------------|-------------|
| β1-β2| -579.4 ± 9.8| -570.9 ± 5.1| -558.5 ± 4.1| -537.3 ± 2.6|
| β1-β3| -442.8 ± 3.2| -437.9 ± 10.0| -415.1 ± 5.7| -369.6 ± 7.3|
| β1-α1| -743.1 ± 12.0| -733.8 ± 9.6| -426.5 ± 4.6| -364.9 ± 3.2|
| β2-α2| -1249.2 ± 10.2| -1274.3 ± 6.9| -1187.5 ± 12.7| -1022.8 ± 15.2|
| β3-β4| -415.5 ± 8.7| -407.8 ± 7.6| -355.9 ± 7.3| -251.2 ± 3.3|
| β3-α2| -821.1 ± 9.1| -826.2 ± 8.2| -833.3 ± 8.4| -728.6 ± 9.5|
| β3-α3| -735.8 ± 5.7| -664.9 ± 4.3| -934.3 ± 9.5| -534.5 ± 4.5|
| β4-α4| -760.8 ± 9.9| -700.4 ± 8.1| -950.0 ± 12.0| -595.6 ± 6.7|
| β4-β5| -325.8 ± 4.9| -314.9 ± 1.4| -390.1 ± 7.7| -227.9 ± 6.6|
| α4-α5| -1797.7 ± 11.0| -1759.5 ± 13.0| -1166.3 ± 9.8| -1352.6 ± 7.8|
| β6-α5| -1004.8 ± 7.4| -951.9 ± 5.3| -895.3 ± 9.2| -860.7 ± 8.1|
| β6-β7| -662.6 ± 1.5| -651.4 ± 2.5| -1162.2 ± 11.8| -461.3 ± 5.5|
| β6-β8| -460.4 ± 5.6| -456.4 ± 6.9| -580.8 ± 4.0| -343.4 ± 2.1|
| β6-α6| -986.7 ± 8.9| -959.3 ± 9.1| -1082.3 ± 4.5| -753.1 ± 5.0|
| β7-α5| -988.1 ± 4.6| -979.5 ± 9.4| -881.2 ± 7.6| -843.3 ± 6.5|
| β8-α4| -970.3 ± 4.2| -935.6 ± 7.7| -889.7 ± 8.1| -735.7 ± 6.1|
| β8-α7| -826.4 ± 11.7| -857.2 ± 5.3| -734.8 ± 5.4| -667.1 ± 3.7|
| β9-α4| -951.8 ± 5.1| -992.4 ± 9.7| -876.2 ± 6.9| -701.5 ± 3.8|
| β9-α7| -793.4 ± 4.3| -971.7 ± 6.5| -576.3 ± 5.4| -566.8 ± 4.8|
| β8-α9| -282.9 ± 9.9| -285.3 ± 3.6| -263.5 ± 3.2| -198.8 ± 9.9|
| β9-β10| -157.8 ± 7.5| -220.9 ± 8.3| -182.1 ± 5.6| -157.97 ± 4.1|
| β10-β11| -246.9 ± 8.7| -230.4 ± 5.4| -222.9 ± 7.6| -156.9 ± 4.9|
| β10-α9| -932.4 ± 3.8| -940.3 ± 8.8| 864.3 ± 6.2| -703.0 ± 9.6|
| β10-α8| -446.6 ± 5.9| -450.2 ± 9.6| -401.7 ± 4.7| -389.8 ± 2.8|

Table 4. Average interaction energies (kcal/mol) of the main secondary structure elements in tmRBP at four different temperatures simulations.

| 2FN9 | 300 K       | 380 K       | 450 K       | 500 K       |
|------|-------------|-------------|-------------|-------------|
| β1-β2| -595.8 ± 5.1| -589.0 ± 4.6| -572.2 ± 6.5| -586.1 ± 7.2|
| β1-β3| -454.3 ± 1.7| -450.1 ± 4.2| -425.3 ± 3.4| -400.1 ± 9.5|
| β1-α1| -1340.0 ± 13.5| -1329.3 ± 8.2| -1219.4 ± 11.3| -978.1 ± 5.6|
| β2-α2| -1333.9 ± 10.2| -1310.8 ± 13.6| -1274.9 ± 13.5| -1164.8 ± 10.8|
| β3-β4| -3535.3 ± 5.5| -3424.2 ± 9.8| -462.6 ± 5.8| -457.5 ± 6.4|
| β3-α2| -839.3 ± 6.8| -1033.5 ± 4.6| -1158.7 ± 9.8| -1318.7 ± 12.9|
| β3-α3| -667.4 ± 5.5| -662.5 ± 5.1| -868.0 ± 8.7| -868.9 ± 9.7|
| β4-α3| -658.7 ± 3.1| -650.3 ± 6.5| -867.1 ± 7.3| -863.2 ± 5.0|
| β4-β5| -390.6 ± 8.2| -360.2 ± 2.5| -513.7 ± 7.2| -481.6 ± 7.4|
| a4-α5| -2399.9 ± 12.7| -2369.3 ± 14.8| -2801.9 ± 11.1| -2181.2 ± 15.8|
| β6-α5| -1330.0 ± 13.4| -1328.8 ± 12.6| -1634.7 ± 15.4| -1553.3 ± 14.1|
| β6-β7| -717.2 ± 6.1| -749.9 ± 8.7| -1009.1 ± 9.5| -962.6 ± 7.9|
| β6-β8| -491.0 ± 3.5| -486.1 ± 6.4| -632.6 ± 7.4| -651.3 ± 9.2|
| β6-α6| -1257.7 ± 7.1| -1234.4 ± 9.1| -1633.5 ± 14.3| -1261.9 ± 16.5|
| β7-α5| -1207.5 ± 8.9| -1090.4 ± 10.1| -968.7 ± 9.5| -874.1 ± 6.9|
| β8-α4| -1442.9 ± 10.7| -1562.3 ± 13.6| -1347.2 ± 14.7| -1833.8 ± 14.7|
| β8-α7| -975.9 ± 8.8| -1018.4 ± 7.9| -983.1 ± 9.5| -1400.5 ± 11.4|
| β9-α4| -1442.1 ± 8.0| -1479.8 ± 12.7| -1327.3 ± 12.3| -1794.5 ± 10.3|
| β9-α7| -9240.9 ± 9.1| -978.3 ± 8.4| -890.6 ± 6.5| -1316.6 ± 7.5|
| β9-β9| -304.8 ± 7.7| -397.4 ± 5.3| -332.5 ± 4.3| -421.2 ± 3.4|
| β9-β10| -350.1 ± 8.2| -387.6 ± 5.9| -369.2 ± 5.2| -432.8 ± 7.5|
| β10-β11| -431.4 ± 1.1| -465.3 ± 5.4| -427.5 ± 4.8| -519.4 ± 6.2|
| β10-α9| -1507.2 ± 15.6| -1610.5 ± 14.9| -1520.6 ± 7.3| -1883.2 ± 12.1|
| β10-α8| -890.7 ± 6.5| -915.2 ± 7.6| -824.9 ± 8.2| -652.7 ± 8.4|
Intramolecular interactions of the tmRBP and ecRBP

Intermolecular interactions between various secondary structural elements constitute in tmRBP and ecRBP play an important role in their overall stability. In order to understand the unfolding pathways of tmRBP and ecRBP and to identify the energetic determinants responsible for such a hierarchical unfolding, the interaction energies between the major secondary elements in tmRBP and ecRBP were calculated (Tables 3 and 4). As can be seen from Tables 3 and 4, most of the intramolecular interaction energies are less favorable in ecRBP compared to tmRBP. Especially, the favorable interaction energies of β1 with α1, α4 with α5, β6 with α6, β7 with α5, β7 with α5, β8 with α4, β9 with α4, and β10 with α4 reduce marginally in ecRBP compared to tmRBP. The decrease in the interaction of β11 with the other parts of ecRBP with an increase in the temperature explains the observation that this is the first to undergo unfolding. Overall, an increase in the favorable interaction energies in tmRBP with respect to ecRBP is suggested to be responsible for the high stability of tmRBP. Given the conservation of structures of these two proteins, it is tempting to conclude that tmRBP is more stable than ecRBP based on these interaction energies. The low favorable interaction energies in ecRBP compared to tmRBP may also be responsible for the faster unfolding of the former during MD simulations (Figure 5 and Supplementary Figure S2).

Table 5. Root mean square inner products (RMSIP) between the first 10 eigenvectors for MD simulation trajectories at different temperatures.

| E300 | E380 | E450 | E500 | T300 | T380 | T450 | T500 |
|------|------|------|------|------|------|------|------|
| E300 | .887 | .683 | .442 | .241 | .234 | .278 | .301 | .238 |
| E380 | .627 | .392 | .22 | .203 | .348 | .281 | .232 | .289 |
| E450 | .406 | .181 | .135 | .238 | .244 | .238 | .244 | .238 |
| E500 | .354 | .164 | .256 | .181 | .183 | .289 | .239 | .259 |
| T300 | .795 | .521 | .442 | .422 | .232 | .238 | .238 | .238 |
| T380 | .521 | .392 | .235 | .238 | .244 | .238 | .238 | .238 |
| T450 | .329 | .235 | .238 | .238 | .238 | .238 | .238 | .238 |
| T500 | .325 | .238 | .238 | .238 | .238 | .238 | .238 | .238 |

Note: E stands for ecRBP and T stands for tmRBP.

**Intramolecular interactions of the tmRBP and ecRBP**

Intramolecular interactions between various secondary structural elements constituted in tmRBP and ecRBP play an important role in their overall stability. In order to understand the unfolding pathways of tmRBP and ecRBP and to identify the energetic determinants responsible for such a hierarchical unfolding, the interaction energies between the major secondary elements in tmRBP and ecRBP were calculated (Tables 3 and 4). As can be seen from Tables 3 and 4, most of the intramolecular interaction energies are less favorable in ecRBP compared to tmRBP. Especially, the favorable interaction energies of β1 with α1, α4 with α5, β6 with α6, β7 with α5, β7 with α5, β8 with α4, β9 with α4, and β10 with α4 reduce marginally in ecRBP compared to tmRBP. The decrease in the interaction of β11 with the other parts of ecRBP with an increase in the temperature explains the observation that this is the first to undergo unfolding. Overall, an increase in the favorable interaction energies in tmRBP with respect to ecRBP is suggested to be responsible for the high stability of tmRBP. Given the conservation of structures of these two proteins, it is tempting to conclude that tmRBP is more stable than ecRBP based on these interaction energies. The low favorable interaction energies in ecRBP compared to tmRBP may also be responsible for the faster unfolding of the former during MD simulations (Figure 5 and Supplementary Figure S2).

**Essential dynamic analysis of the tmRBP and ecRBP**

To further explore the dynamic properties of tmRBP and ecRBP in our simulations, the ED analysis on the Cα atoms was performed. The overall flexibility of tmRBP and ecRBP was explored by calculating the trace of the diagonalized covariance matrix of the atomic positional fluctuations. We have gained the following values for tmRBP: .63 nm² at 300 K, 1.13 nm² at 380 K, 5.76 nm² at 450 K, and 15.79 nm² at 500 K, whereas for ecRBP, we have .76 nm² at 300 K, 4.19 nm² at 380 K, 7.26 nm² at 450 K, and 20.18 nm² at 500 K. These results lead to the conclusion that these two proteins show different influences on the stability of its structure with an increase in the temperature, while ecRBP is more flexible compared to tmRBP in the same temperature simulations.
It is consistent with our previous analysis that the rate of unfolding transitions of the tmRBP occurred much later than the ecRBP at high temperature simulations.

To evaluate quantitatively the essential subspaces in tmRBP and ecRBP, we also computed the RMSIP between the first 10 eigenvectors of each simulation (Table 5). As can be seen from Table 5, the diagonal elements indicate the overlap in the individual trajectories. At different temperature simulations, the RMSIP values are high, indicating the highest degree of overlap within the individual trajectories. In addition, the RMSIP values between the same proteins are high at lower temperatures (300 and 380 K), while they are significantly low at higher temperatures (450 and 500 K), implying that the highest degree of overlap between the same proteins occurs within the individual trajectories at lower temperatures (300 and 380 K). However, the RMSIP values between the two different proteins at four temperatures are significantly low, indicating that the essential subspaces of the two different proteins are overlapping insignificantly at four temperatures. These results lead to the conclusion that each protein has similar direction of motion, while these two proteins display different directions of motion at different simulated temperatures. To further investigate thoroughly the dynamical differences between tmRBP and ecRBP, the first eigenvector was analyzed at four temperatures simulations (Figure 7 and Supplementary Figure S2). The arrows indicated the dynamical behaviors of the three-dimensional correlated motion of the involved regions, while the width of the ribbon shows the backbone motion. As can be seen from Figure 7 and Supplementary Figure S2, the direction of motion of tmRBP is different from that of ecRBP at the four temperatures simulations, while the conformational space sample of tmRBP and ecRBP exhibits an increasing trend as the temperatures rise. Compared to tmRBP, the amplitude of the fluctuations is higher in ecRBP, indicating that the flexibility of ecRBP is higher than the corresponding flexibility of tmRBP at the same temperature. These results are in good agreement with the results of RMSIP and further indicate that internal motions of the two proteins are significantly different at same temperature.

**Hierarchical unfolding pathway of tmRBP and ecRBP**

The structural changes of tmRBP and ecRBP are explored through characterization of the projected trajectories onto the first two PC spaces (Figure 8). On these projections at 300 and 380 K, clusters of stable states are well defined in tmRBP and ecRBP. We can also see that tmRBP and ecRBP have different distributions in the PC spaces.
space at these two temperatures. Thus, the essential sub-
space and the structural fluctuations in these two proteins
can be extremely different with increasing temperature.
The distributions of ecRBP in the PC space larger than
that of tmRBP and tmRBP tend to show drastically
reduced coverage than ecRBP, indicating ecRBP under-
goes remarking structural changes than tmRBP. Such
large structural fluctuations at higher temperature may be
associated with an initial unfolding process that changes
the essential coordinate definition. The earlier experimen-
tal studies (Cuneo et al., 2008) have also shown that the
ecRBP has a lower stability than tmRBP. We want to
further investigate the differences in the stability of
tmRBP and ecRBP using thermal unfolding studies by
carrying out MD simulations at elevated temperature.
Thereby, several representative snapshots at different
times along the unfolding processes of the studied pro-
teins are displayed (Figure 9 and Supplementary
Figure S3). At 450 K in the MD simulation (Figure 9),
the unfolding processes of ecRBP began with the melt-
ing of helices 3 and 8 around 0.8 ns. From the snapshot
at 50 ns, it was found that the distorted helix 3 was
refolded. The distortion of helices 4, 5, and 9 in ecRBP
happened at around 9 ns. The unfolding of the helix 1
happened at around 16 ns. Helix 7 was retained until
70 ns in our simulation. Compared to ecRBP, the unfold-
ing of tmRBP evolved much more slowly which was
also indicated by the change of RMSD. Helix 8 was
firstly melted in tmRBP as shown in Figure 9 and then a
distortion of helix 3 was observed around 3.2 ns. Differ-
ent to ecRBP, the unfolding of β10 and β11 in tmRBP
happened at around 14.5 ns. Helix 4 was retained until
the end of the simulation in tmRBP. At 500 K, both the
two proteins become highly coiled early in the simula-
tion, but only a few rearranged sheets remain in their
structure. The unfolding pathway analysis leads us to
conclude that the rate of unfolding is faster in ecRBP
compared to tmRBP and the sequence of unfolding
events is different in these two proteins.

Conclusions
Focused on the hyperthermophilic tmRBP and mesophi-
lic ecRBP, we performed MD simulations at four differ-
ent temperatures (300, 380, 450, and 500 K) to
investigate the thermal stability and unfolding pathways
of these two homologs.

Detailed analysis of these MD trajectories in terms of
secondary structure content, solvent accessibility,
intramolecular hydrogen bonds, salt bridges, and inter-
molecular interactions indicates the differences in the
structural stability between tmRBP and ecRBP. The
unfolding studies of tmRBP and ecRBP suggest that the
rate of unfolding is faster in ecRBP compared to tmRBP
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 high
flexibility of ecRBP than tmRBP at the same temperature.
The ED analysis also indicates that tmRBP and ecRBP
exhibit different principal directions of the motions.

MD simulations and ED analysis provide the above
structural and dynamic properties, which are inaccessible
in the static crystal structure. This study can provide an
argument for further experimental and theoretical investi-
gations into the stability of RBP.

Supplementary material
The supplementary material for this paper is avail-
able online at http://dx.doi.org/10.1080/07391102.2015.
1084480.

Funding
This work is supported by the National Natural Science
Foundation of Heilongjiang [grant number B201420], [grant
number QC201012]; Program for new century excellent talents
in Heilongjiang Provincial University [grant number 1252-
NECT-022].

References
Amadei, A., Linssen, A., & Berendsen, H. J. (1993). Essential
dynamics of proteins. *Proteins: Structure, Function, and
Genetics, 17*, 412–425.
Amadei, A., Ceruso, M. A., & Di Nola, A. (1999). On the con-
vergence of the conformational coordinates basis set
obtained by the essential dynamics analysis of proteins’
molecular dynamics simulations. *Proteins: Structure, Func-
tion, and Genetics, 36*, 419–424.
van der Spoel, D., van Maaren, P. J., & Berendsen, H. J. C. (1998). A systematic study of water models for molecular simulation: Derivation of water models optimized for use with a reaction field. The Journal of Chemical Physics, 108, 10220–10230.

van der Spoel, D., Lindahl, E., Hess, B., Van Buuren, A., Apol, E., & Meulenhoff, P. (2008). Gromacs user manual version 3.3. www.gromacs.org.

Tang, L., & Liu, H. (2007). A comparative molecular dynamics study of thermophilic and mesophilic ribonuclease HI enzymes. Journal of Biomolecular Structure and Dynamics, 24, 379–392.

Thompson, J., Higgins, D., & Gibson, T. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research, 22, 4673–4680.

Turner, P., Mamo, G., & Karlsson, E. N. (2007). Potential and utilization of thermophiles and thermostable enzymes in biorefining. Microb Cell Fact, 6, 1–23.

Vieille, C., & Zeikus, G. J. (2001). Hyperthermophilic enzymes: Sources, uses, and molecular mechanisms for thermostability. Microbiology and Molecular Biology Reviews, 65, 1–43.

Vogt, G., Woell, S., & Argos, P. (1997). Protein thermal stability, hydrogen bonds, and ion pairs. Journal of Molecular Biology, 269, 631–643.