Available online at www.sciencedirect.com
ScienceDirect
Journal of Molecular Graphics and Modelling 26 (2007) 124–134
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Comparative unfolding studies of psychrophilic and mesophilic uracil DNA glycosylase: MD simulations show reduced thermal stability of the cold-adapted enzyme

Magne Olufsen, Bjørn Olav Brandsdal, Arne Oskar Smalås *

The Norwegian Structural Biology Centre, Department of Chemistry, University of Tromsø, N-9037 Tromsø, Norway

Received 14 June 2006; received in revised form 17 October 2006; accepted 18 October 2006
Available online 24 October 2006

Abstract

Uracil DNA glycosylase (UDG) is a DNA repair enzyme involved in the base excision repair (BER) pathway, removing misincorporated uracil from the DNA strand. The native and mutant forms of Atlantic cod and human UDG have previously been characterized in terms of kinetic and thermodynamic properties as well as the determination of several crystal structures. This data shows that the cold-adapted enzyme is more catalytically efficient but at the same time less resistant to heat compared to its warm-active counterpart. In this study, the structure–function relationship is further explored by means of comparative molecular dynamics (MD) simulations at three different temperatures (375, 400 and 425 K) to gain a deeper insight into the structural features responsible for the reduced thermostability of the cold-active enzyme. The simulations show that there are distinct structural differences in the unfolding pathway between the two homologues, particularly evident in the N- and C-terminals. Distortion of the mesophilic enzyme is initiated simultaneously in the N- and C-terminal, while the C-terminal part plays a key role for the stability of the psychrophilic enzyme. The simulations also show that at certain temperatures the cold-adapted enzyme unfolds faster than the warm-active homologues in accordance with the lower thermal stability found experimentally.

Keywords: Molecular dynamics; Uracil DNA glycosylase; Protein unfolding; Cold adaptation; Psychrophilic

1. Introduction

Several cold-adapted enzymes from different organisms living in constantly cold environments have been identified and characterized in the last few decades. Such enzymes possess unique properties, attractive not only for the academic research but also to the biotechnological market. The emerging picture suggests that cold-adapted enzymes have higher catalytic efficiency, lower thermostability and improved flexibility in the structural parts directly involved in the catalytic cycle [1,2]. Indeed, most psychrophilic enzymes studied so far display reduced overall stability when compared to their mesophilic counterparts. The increased catalytic efficiency is usually accompanied by a decrease in thermal stability, which is believed to be a consequence of higher structural flexibility in the psychrophilic enzymes. When the temperature decreases from 37 to 0 °C the reaction rates are generally reduced by 30- to 80-fold [2], and in order to maintain a sufficient metabolic flux at low temperatures it has been suggested that psychrophilic enzymes have higher flexibility in the structural parts directly involved in the catalytic cycle [1]. However, enzymes are shown to use different structural adaptation strategies to achieve the above-mentioned features [3], and psychrophilic enzymes often possess decreased number of proline residues, increased number of glycine residues, low relative arginine content [arg/(arg + lys)] and lower numbers of ion pairs, aromatic interactions and/or hydrogen bonds compared to the mesophilic counterparts [3–5].

Cold-adapted uracil DNA glycosylase from cod (cUDG) has been shown to be up to 10 times more catalytic efficient (k_{cat}/K_m) in the temperature range from 15 to 37 °C compared to the warm-active human counterpart (hUDG) [6]. UDG is the first enzyme in the BER pathway, and catalyzes the hydrolysis of promutagenic uracil residues from single- or double-stranded DNA [7]. The crystal structure of the catalytic domain of UDG from several species are known: human (hUDG) [8], Atlantic

* Corresponding author. Tel.: +47 77644070; fax: +47 77644765.
E-mail address: Arne.Smalas@chem.uit.no (A.O. Smalås).

1093-3263/$ – see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2006.10.003

cod (cUDG) [9], virus 1 [10] and *Escherichia coli* [11]. The catalytic domain of hUDG and cUDG consists of 223 amino acids with a sequence identity of 75%, and the overall topology is a typical α/β protein [8]. The r.m.s.d. between the two crystal structures is 0.492 Å (all atoms were included). The higher catalytic efficiency of cUDG compared to hUDG has been postulated to arise from a more flexible DNA-recognition loop [12] and optimization of the electrostatic surface potential near the specificity pocket [9] in the cold-adapted enzyme.

cUDG has also been found to be much more pH and temperature labile than hUDG [13], indicating that cUDG is generally less stable compared to hUDG. Large opposing energies are involved when proteins go from the unfolded to the folded state, but the overall change in energy is small, typically ranging from 1 to 15 kcal/mol [14]. All contributions, both favorable and unfavorable, are therefore important when considering protein stability [14]. How proteins achieve their stability does not seem to follow any general rules [15], but the emerging picture is, for globular proteins, that the hydrophobic effect and burial of non-polar side chains stabilizes the native state [16,17]. Disulfide bonds, electrostatic interactions and hydrogen bonds are, however, also important for structural stability and can contribute favorably to protein stability [18–20]. The majority of reversible two-state folding proteins, including thermophilic, mesophilic and psychrophilic enzymes, show maximum stability around room temperature [17].

High temperature MD simulations have been shown to provide valuable insights into the unfolding process of several proteins: chymotrypsin inhibitor 2 (CI2) [21], barnase [22,23] and engrailed homeodomain (En-HD) [24]. Protein unfolding studies using MD simulations are carried out by increasing the temperature necessary to overcome the enthalpic forces stabilizing the 3D structure. Because of the differences in accessible time scales by experiments and computer studies, higher temperatures are needed when performing computer experiments relative to that observed in experimental unfolding studies. A MD study of CI2 has shown that the unfolding pathway is not widely affected by the simulation temperature, and thus raising the temperature only increases the reaction rate [25].

In this study, we have employed MD simulations to investigate the structure–function relationship further with focus on possible differences in the thermal unfolding pathway between a psychrophilic and a mesophilic UDG. We have especially analyzed the structures to see if there are important molecular contacts, which could explain the observed difference in thermal stability between these enzymes. Comparative high temperature (375, 400, 425 K) simulations for 6 and 9 ns have been carried out for the two enzymes, and the unfolding process has been carefully monitored both at an overall and a detailed structural level.

2. Methods

All MD simulations were carried out using the AMBER7 program package [26] with the recently developed force field parm99 [27]. The generalized Born methodology [28] was applied to describe the electrostatics from the solvent. Implicit solvent has been chosen in this study. The advantage of using implicit solvent models is the increase in computational efficiency. MD simulations with implicit solvent have successfully been used to investigate protein folding [29,30], and recent development in the implicit solvent methodology now provide a level of realism that challenge the results from the explicit solvent simulations [31]. The simulations were initiated with 100 cycles of energy minimization and the protein atoms constrained with a 5 kcal/mol $Å^2$ force constant. The temperature was slowly raised to the final temperature (375, 400 or 425 K) during a heating phase of approximately 20 ps. In this phase, the protein atoms were weakly restrained to their initial positions. SHAKE [32] was employed to constrain all bonds involving hydrogen atoms. A cutoff of 15 Å was used for the non-bonded interactions and every 1000 step the rotational and translational motion was removed. The time step was set to 2 fs, and the temperature was maintained by the Berendsen coupling algorithm [33]. The hydrophobic effect was taken into account by adding a surface energy term to the total energy [34].

Crystal structures were used as starting structures in the MD simulations for both hUDG [8] and cUDG [35] (Protein Data Bank entries 1AKZ and 1OKB, respectively). The crystal structures were recombinant enzymes with three mutations in the N-terminal end: P82M, V83E and G84F. UDG contains several histidines, and in the simulation the His 148 was charged, while the rest of the histidines were considered as neutral with the hydrogen atom placed on the Nε atom. These choices were based on data from NMR and continuum electrostatics calculations [36].

Hydrogen bonds and r.m.s.d. values of the Cα atoms were calculated with the Amber program package. In the hydrogen bond calculations a distance cutoff of 3.4 Å and an angle cutoff of 60° (the angle between: hydrogen-donor–acceptor) were applied. These calculations were also tested with no angle cutoff, but these calculations gave similar results, and the angle cutoff method was used in this study. The DSSP program [37] was used to find the secondary structure elements of each snapshot. The secondary structure elements were calculated for snapshots every 10th ps and for smoothening the plots, the number of residues in secondary elements was averaged every 10 snapshots. The surface area was calculated by the MSMS program [38], and a probe radius of 1.5 Å was applied to calculate the solvent accessible surface area (SASA).

3. Results and discussion

Organisms have adapted to extreme environments and survive at different temperatures by optimization of their proteins at some level and thereby maintain sufficient metabolic fluxes. Increased availability of sequences, kinetic, thermodynamic and structural information of proteins from psychrophilic, mesophilic and thermophilic enzymes have facilitated a deeper understanding of features specific for adaptation at the molecular level. Presently, much deeper insight has been

achieved on how proteins have adapted to cope with high temperatures without denaturation, as compared to the low temperature adaptation. Higher temperatures involve increased reaction rates and atomic mobility, while low temperature slows down the chemical reaction rates and reduced solvent viscosity. Hence, it may appear that adaptation to high and low temperature requires opposite changes, but several studies have shown that the picture is far more complicated than simply reversing the features responsible for high temperature adaptation when explaining how the challenges are met at low temperature.

Uracil DNA glycosylase is an excellent model system for the study of the mechanisms of enzymatic adaptation to low temperature. Its biological function is well-known, and several crystal structures of both cUDG and hUDG are available, including mutant and native structures, which have also been characterized in terms of stability and kinetics [6,8,9,13,35,39]. It is well-established that cold-adapted enzymes are less resistant to heat as compared to their warm-active homologues and this has also been shown to be the case for cUDG [13]. In light of the small differences revealed by structural analysis, computational simulation of the thermal unfolding process was a natural next step.

In order to gain further insights into the thermal stability of cUDG and hUDG eight comparative MD simulations at three different temperatures (375, 400 and 425 K) have been carried out. Molecular interactions and intramolecular contacts (like salt-bridges, hydrogen bonds and hydrophobic contacts) have been thoroughly monitored in order to search for specific features and structurally important regions that could explain the difference in thermal stability. Even higher temperatures were tested, but increasing the temperature beyond 425 K yielded complete loss of ordered structure within just a few ps. The resulting trajectories were too difficult to analyze as most of the sampling were on the fully unfolded polypeptide chains.

3.1. Structural fluctuation in the MD simulations

Computer aided thermal unfolding of proteins are usually performed at temperatures between 200 and 225 °C (473–498 K) [40], and a commonly used criterion for an unfolded state has been a Cα r.m.s.d. exceeding 10 Å [41]. According to this, the proteins studied here are all fully denaturated after approximately 5–6 ns. There are still some fluctuations in the r.m.s.d. values in the simulations at the two highest temperatures (Fig. 1A and B). However, the protein unfolding proceeds slowly in the 375 K simulations (Fig. 1A), and the structural ensembles show less fluctuations when compared to the 400 and 425 K simulations. The 375 K simulations were therefore extended to 9 ns to reach this threshold. cUDG and hUDG show similar unfolding rates for the first 5 ns of the simulations. However, the psychrophilic enzyme displays

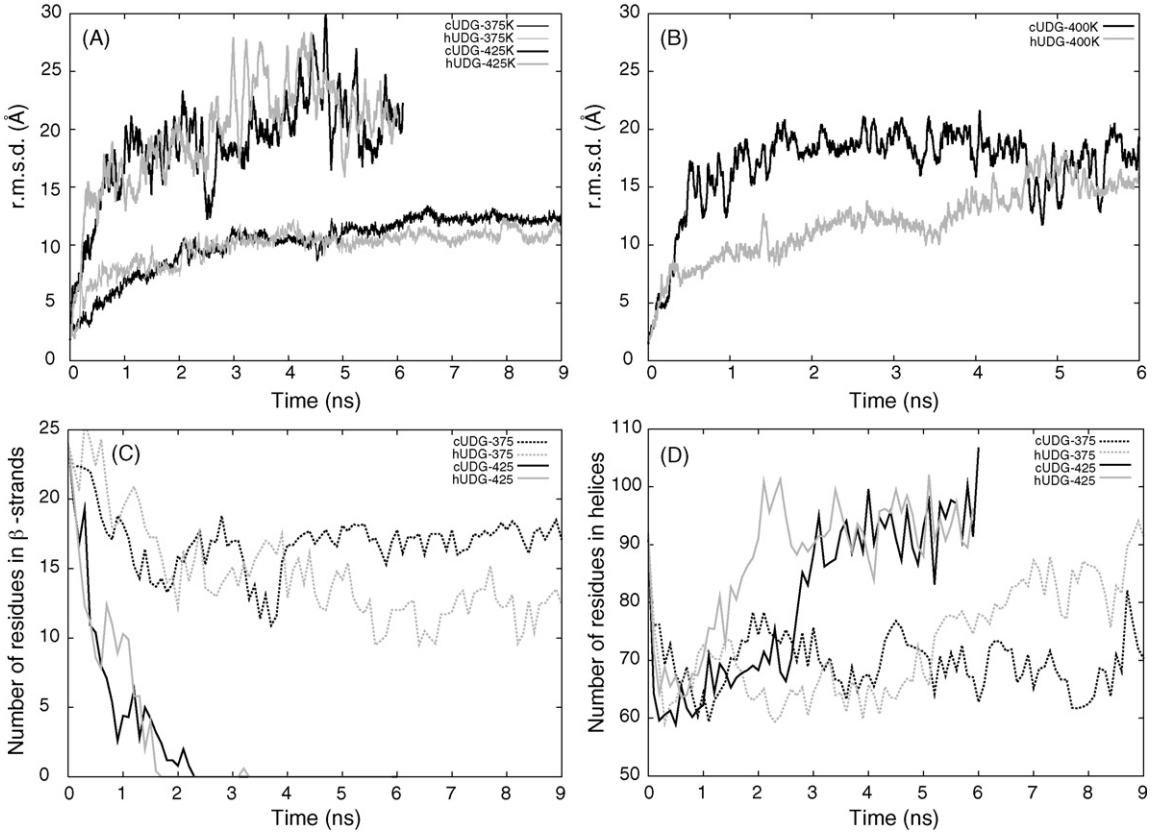

Fig. 1. Various plots from the simulations. (A) and (B) are average r.m.s.d. plots as a function of time during the simulation. The calculated r.m.s.d. values are based on all Cα atoms in the starting structure and from every snapshot during the simulation. (C) and (D) are number of residues in secondary structure elements during the simulation. The secondary structure elements are calculated by DSSP [37].

higher r.m.s.d. values in the last 4 ns (Fig. 1A), indicating greater structural changes in the cold-adapted enzyme. The unfolding rate is faster when the temperature is raised to 400 K, and the simulations show that the enzymes unfolded to a higher degree at this temperature when compared to the 375 K results. Interestingly, large differences in the unfolding rates between the psychrophilic and the mesophilic enzyme are observed in the 400 K simulations. The psychrophilic enzyme unfolds much more rapidly than the mesophilic counterpart, indicating reduced structural stability (Fig. 1B). Destruction of the native protein structures occurs very fast when the temperature is raised to 425 K, and unfolding of cUDG and hUDG proceeds at a similar rate. Moreover, the simulations show that the mesophilic enzyme requires higher temperatures to achieve similar degree of structural distortions as observed for the psychrophilic enzyme. This already indicates differences in the interactions stabilizing the 3D structure, and is also in agreement with the observed reduced stability of the cold-adapted enzyme.

4. Structural investigations

4.1. Secondary structure

The starting structure of both cUDG and hUDG contain 92 and 24 residues in helices (α-helices, π-helices and 3–10 helices) and β-strands, respectively. The amount of residues in β-strands decreases at approximately the same rate in the simulation at 375 and 400 K for both enzymes (Fig. 1C), while for the simulation at the highest temperature, all β-strands are lost at around 2 ns for both enzymes (Fig. 1C). Both enzymes also have similar extent of residues involved in helices in the 375 and 400 K simulations, while it actually increases for both enzymes the 425 K simulation as the structure get distorted (Fig. 1D) (residues in helices for the 400 K simulation is not shown). Visual inspection of the simulated trajectories shows that the unfolding process generates new helices, originating from areas that were initially loops and β-strands in their respective crystal structures. The catalytic domain of both cUDG and hUDG is distorted already in the beginning of the 425 K simulations, and after 600 ps both enzymes have r.m.s.d. values of more than 15 Å (Figs. 1A and 7). When the domain becomes distorted and the β-sheets are destroyed, solvent gains access to the core β-strands yielding rapid unfolding. On the contrary, helices are stable even after an opening of the catalytic domain. However, both NMR and MD studies have shown that the denatured state may comprise significant amount of secondary structure [42]. There are also indications that the unfolded protein can increase the helical content during the simulation and that helices can be formed in water in the absence of significant tertiary interactions [22].

4.2. Hydrogen bonding pattern

The starting structures of cUDG and hUDG contained 240 and 223 intramolecular hydrogen bonds (shorter than 3.4 Å), respectively, and many are lost early in the simulations. However, after approximately 1 ns the number of hydrogen bonds reaches a stabile plateau, which is steadily maintained throughout the simulations. The average number of hydrogen bonds for the last 5 ns of the cUDG simulations is 198, 179 and 165 for the 375, 400 and the 425 K simulation, respectively, while the corresponding numbers for the mesophilic enzyme are 191, 174 and 156. Thus, as the simulation temperature is increased, there is a concomitant decrease in number of intact hydrogen bonds. This is reasonable as the structures become more distorted as the simulation temperature is raised.

4.3. Accessible surface area

The solvent accessible surface area (SASA) is an important parameter for mapping unfolding. The SASA has been calculated for the crystal structures and of snapshots from the MD simulations. The initial SASA for cUDG and hUDG is 10259 and 10376 $Å^2$, respectively. SASA increases rapidly for the 375 K simulations by approximately 20%, and stays roughly at the same level throughout the simulations (Fig. 2A). In the 400 K simulation the SASA increase to roughly 40% at the end of the simulation (Fig. 2B). While at 425 K the enzyme rapidly unfolds, and already after only 100–200 ps the SASA

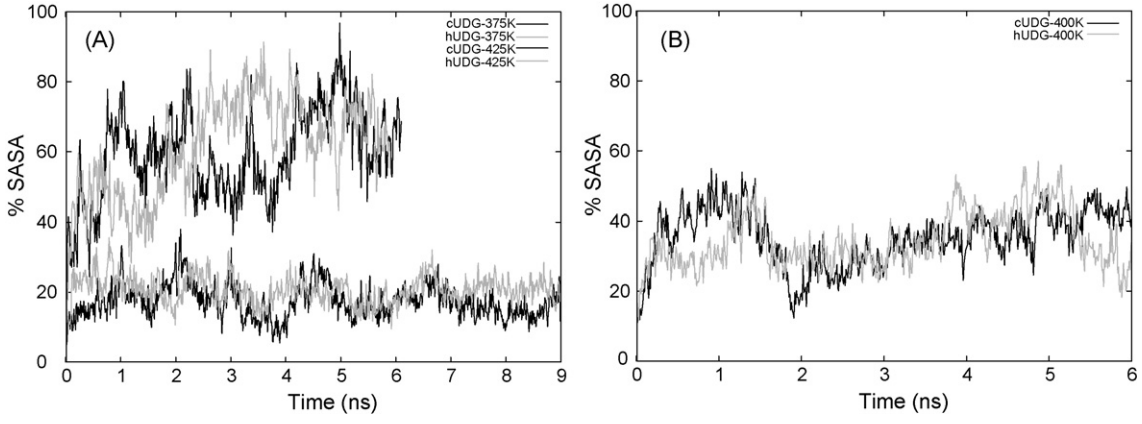

Fig. 2. Solvent-accessible surface area (SASA) as a function of simulation time. (A) shows percent increase in SASA from the 375 and the 425 K simulation. (B) shows percent increase in SASA from the 400 K simulation. The reference is the SASA from the crystal structures. The SASA are calculated by MSMS [38].

has increased by 40% (Fig. 2A). There are only minor differences in SASA among the enzymes as they unfold at different temperatures.

5. Stability of the N- and C-terminals

The N-terminal of cUDG unfolds at all three temperatures and as expected to a greater extent as the temperature is raised (Fig. 3A, C and E). While the N-terminal of the mesophilic variant of the enzyme also unfolds at all three temperatures (Fig. 3B, D and F), difference is observed when compared to cUDG. The N-terminal is more stable in the hUDG simulations as observed in the 375 and 400 K simulations. Structural investigations of the interactions involving the N-terminal show that there are several stabilizing contacts present. Numerous hydrophobic contacts between residues in the N-terminal helices (α1 and α2, α-helix 1 and 2, numbering according to Leiros et al. [9]) and residues in α6 and the loop between α6 and β1 are found. However, three hydrogen bonds (Ser88:Oγ-Asp133:Oδ2, Trp89:Nε1–Cys132:O and Trp89:Nε1–Thr129:O), which are conserved in the two enzymes, appear to be important for the N-terminal stability (Fig. 4). These hydrogen bonds connect the

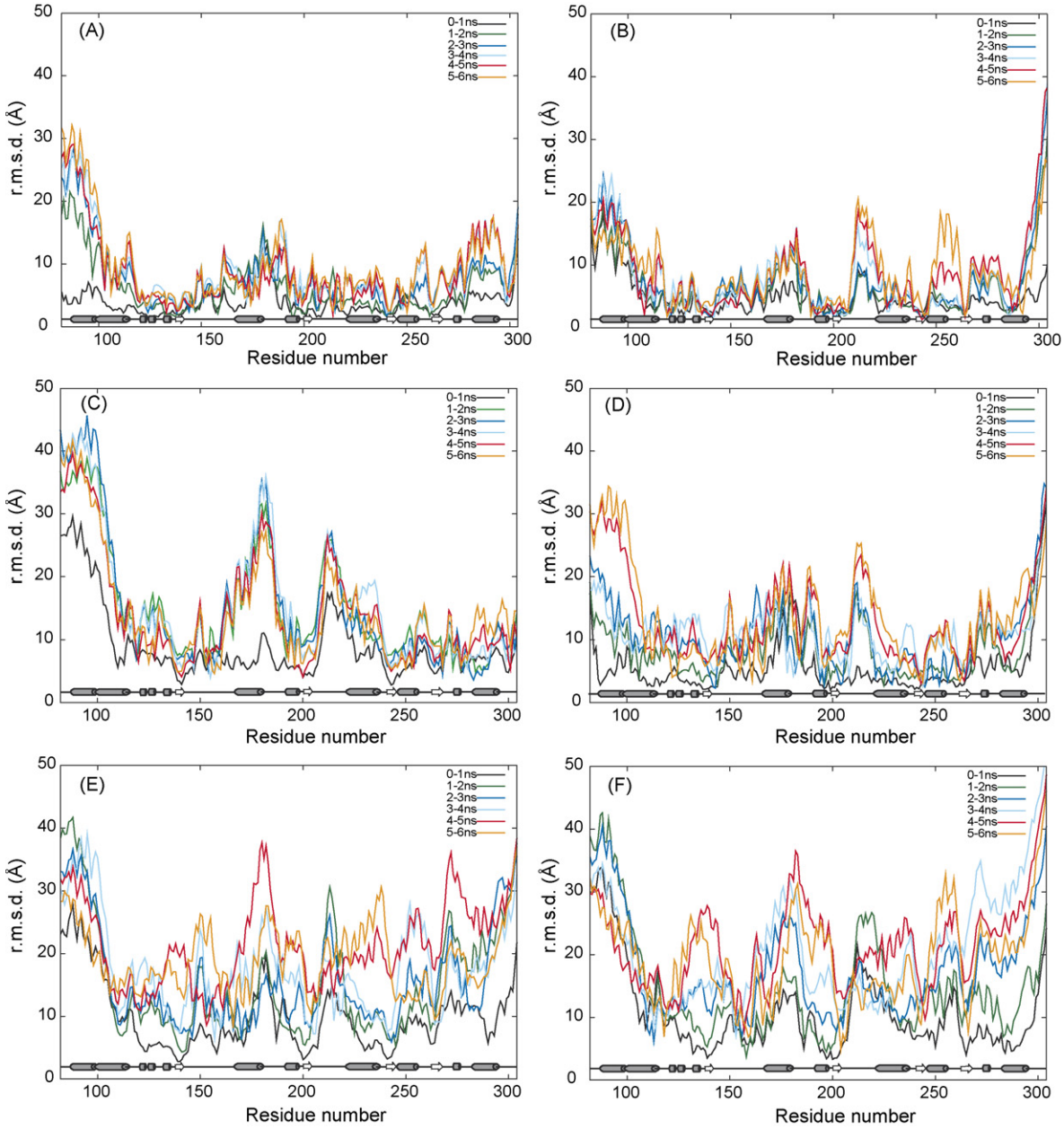

Fig. 3. Flexibility in the structure (r.m.s.d. per residue). The figure shows r.m.s.d. per residue. values in different part of the simulations, 0–1 ns is the r.m.s.d. values from the simulations from 0 to 1 ns, and so on. All the Cα atoms are used to calculate the r.m.s.d. values and the starting structure is used as a reference. The gray cylinder and the white arrow at the bottom of each plot is where helices and β-strands are located in the starting structure, respectively. (A), (C) and (E) is the r.m.s.d. values from the cUDG simulations at 375, 400, and 425 K, respectively. While (B), (D) and (F) is the r.m.s.d. values from the hUDG simulations at 375, 400 and 425 K, respectively.

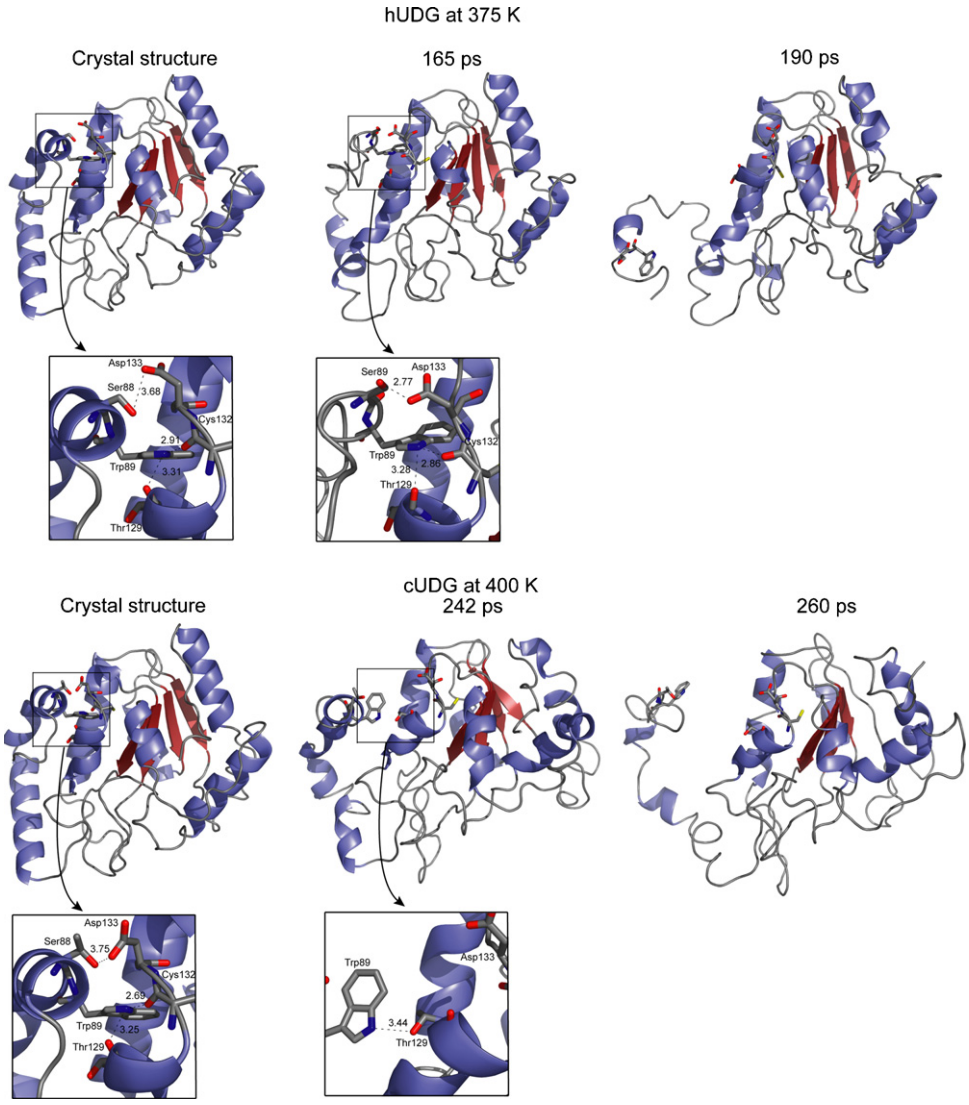

Fig. 4. Snapshots from the unfolding of hUDG at 375 K and cUDG at 400 K. The figure shows how the N-terminal unfolds from the rest of the structure. Some important residues are shown in ball-and-stick. The figure is generated in PyMOL [45].

N-terminal of helix α1 and the helix α4 and the loop between helix α4 and helix α5. In addition, the side chain of Trp89 packs into a hydrophobic area between helices α4 and α8.

Subtle differences in the hydrogen-bonding pattern observed during the simulations of cUDG and hUDG result in strikingly different behavior of their respective N-terminals. In the 375 K cUDG simulation the N-terminal starts to unfold at around 470 ps (Figs. 3A, 5A and 5B), as the previously mentioned hydrogen bond, Trp89:Nε1–Thr129:O, is lost. In contrast, the N-terminal starts to unfold at 177 ps in the 375 K simulation of hUDG, here the hydrogen bond between Ser88:Oγ and Asp133:Oδ2 is lost after 169 ps, and 8 ps later the Trp89:Nε1–Thr129:O hydrogen bond also disappear. Together, an immediate unfolding of the α1 and α2 helices is observed (Fig. 5A and B) in hUDG. Fig. 4 shows a detailed picture of the unfolding events occurring in the N-terminal of hUDG. The mesophilic enzyme starts to unfold earlier than the psychrophilic homologue at this temperature, but at around 2 ns the psychrophilic enzyme is more severely unfolded in the N-terminal.

When the temperature is increased to 400 K, the simulations show that the N-terminal of cUDG becomes highly flexible after only 70 ps. This is again a result of the loss of the Ser88:Oγ-Asp133:Oδ2 hydrogen bond. After the distortion of an additional hydrogen bond, Trp89:Nε1–Thr129:O, complete unfolding of the terminal is observed (Fig. 5D). Figs. 4 and 7 show the unfolding events observed for the N-terminal. Analysis of the simulation of hUDG at 400 K reveals that the hydrogen bond between the Ser88:Oγ and the Asp133:Oδ2 is lost after about 1025 ps (Fig. 5E), while the hydrogen bond between the Trp89:Nε1 and the Thr129:O disappear after 1075 ps (Fig. 5D). These events lead to a melting of the N-terminal in a similar manner as observed for cUDG.

In the 425 K simulation, molecular contacts are lost much faster than at lower temperatures, for example in the hUDG simulation important hydrogen bond in the N-terminal (Trp89:Nε1–Thr129:O) are lost after 110 ps and then the N-terminal starts to unfold (Fig. 5G). The corresponding in hydrogen bond in the psychrophilic enzyme is lost already after

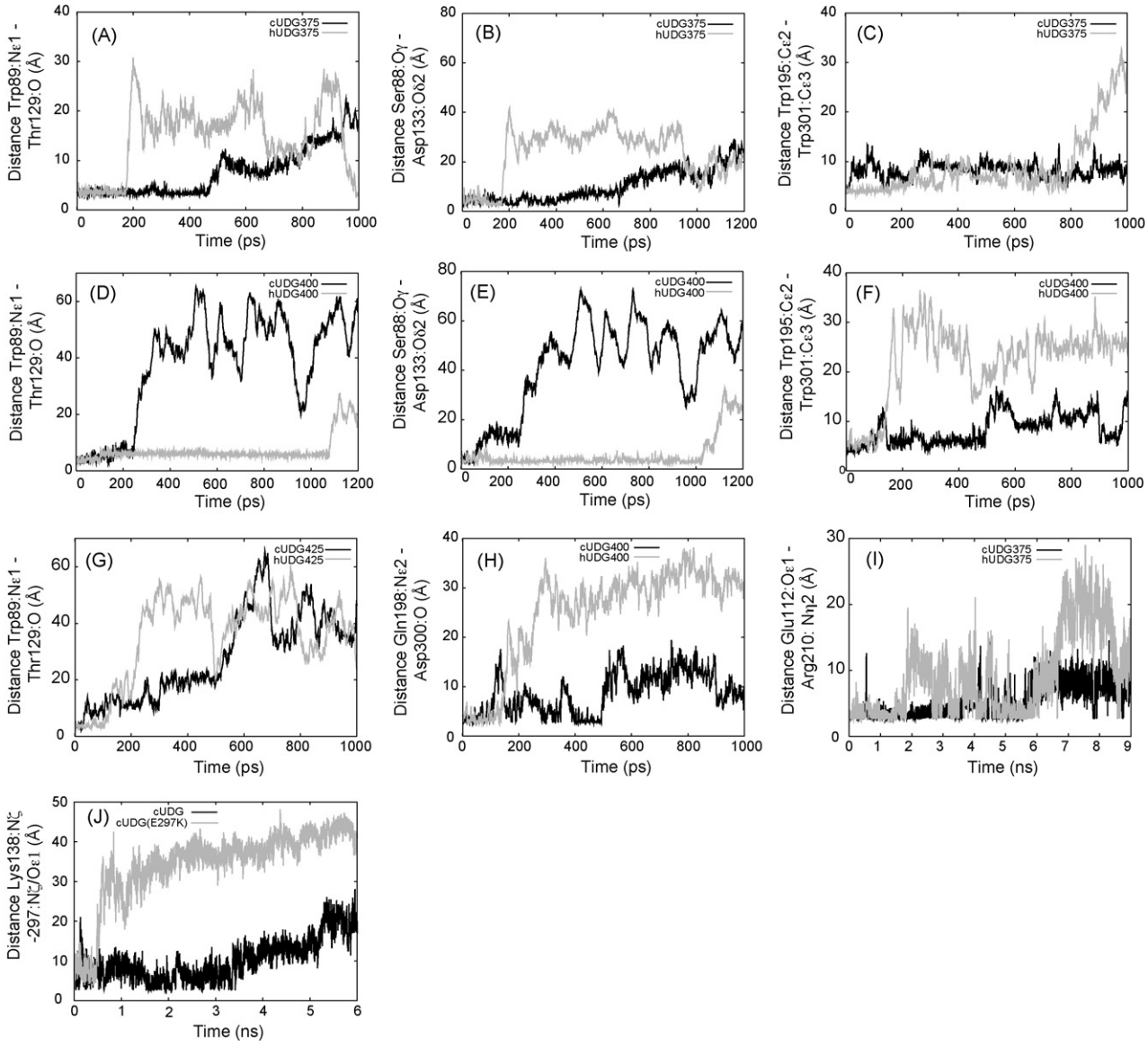

Fig. 5. Atom distance plots. Distances between selected atoms during simulation. (A), (D) and (G): distanced between the atoms Trp89:Nε1 and Thr129:O for the 375, 400 and 425 K simulation, respectively. (B) and (E): distance between the atoms Ser88:Oγ and Asp133:Oδ2 for the 375 and 400 K simulation, respectively. These above-mentioned atom pairs form hydrogen bonds close to the N-terminal. (C) and (F): distance between the atoms Trp195:Cε2 and Trp301:Cε3, for the 375 and 400 K simulation, respectively, which form a hydrophobic contact in the C-terminal. (H): distance between the atoms Gln198:Nε2 and Asp300:O in the 400 K simulation, these residues form a hydrogen bond in the C-terminal. (I): distance between atoms Glu112:Oε1 and Arg 210:Nη2. (J): distance between the atoms Lys138:Nζ and atom (Lys297:Nζ or Glu297:Oε1) in cUDG(E297K) and cUDG, respectively. Both simulations are done at 400 K simulation. There are different scales on the axes in the different plots.

25 ps (Fig. 5G). Many molecular contacts and interactions have been monitored in these simulations as well, and the unfolding of the two enzymes is very similar compared to the lower simulations temperatures. However, it is apparent that the kinetic energy of the system is so high that too many contacts are lost simultaneously making analysis of the simulations more difficult.

The different behavior of the N-terminal of cUDG compared to the hUDG is not obvious, but it is likely that subtle differences in the interactions contribute differently to the stability of the two N-terminals. In particular, the Trp89:Nε1–Thr129:O hydrogen bond seems to be important for N-terminal stability. Interestingly, the recently reported crystal structure of a thermophilic UDG from *Thermus thermophilus* HB8 (TthUDG) has a stabilizing [4Fe-4S] cluster close to the N-terminal [43]. TthUDG belongs to the family 4 UDG while hUDG and cUDG are from family 1. Even though the amino acid sequence homology is low between TthUDG and hUDG or cUDG, the topology and order of the secondary structure elements are very similar between the two UDG families [43]. High temperature adaptation thus seems to involve additional stabilization of the N-terminal part. Hence, increasing the stability of the N-terminal might make the mesophilic and the psychrophilic enzyme more thermostable. MD studies of Chymotrypsin inhibitor 2 have also shown that the N-terminal unfolds from the rest of the structure as one of the first events in

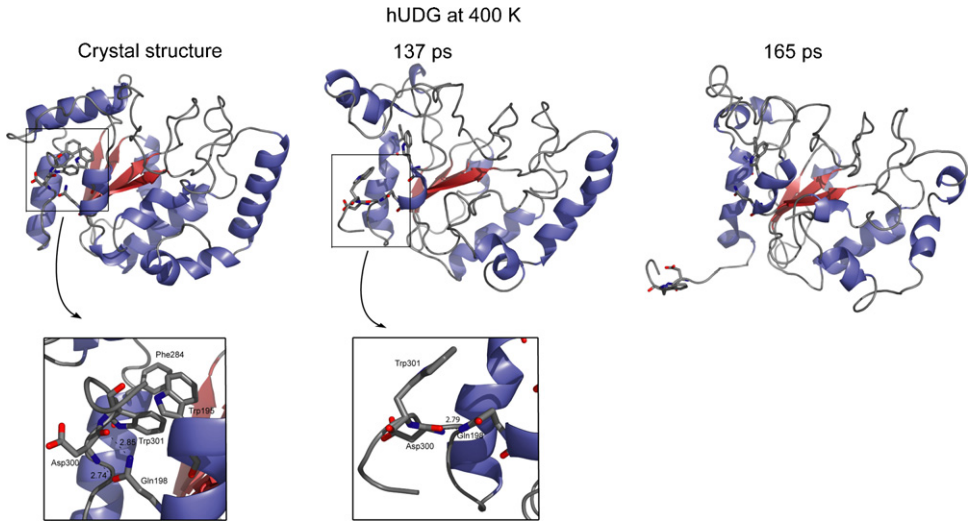

Fig. 6. Snapshots from the unfolding of hUDG at 400 K. The figure shows how the C-terminal unfolds from the rest of the structure. Some important residues are shown in ball-and-stick. The figure is generated in PyMOL [45].

the unfolding process [25,44], acting as an initiator of unfolding.

The very end of the C-terminal part (residue 293–304) is situated at the surface of the protein, and the side of the C-terminal facing away from the solvent is involved in a hydrophobic cluster with helices α7 and α11. In addition, the Trp301 side chain forms stacking interactions with the side chains of Phe284 and Trp195. There are also several hydrogen bonds that could be important for C-terminal stability, especially involving residue Gln198, forming hydrogen bond interaction with residue 300 (Asp in hUDG and Asn in cUDG) (Fig. 6). The fact that Trp301, Phe284, Trp195 and Gln198 are all conserved in the sequences studied by Leiros et al. [9], additionally indicates structural importance. The simulation shows that the C-terminal is far more stable in the cold-adapted enzyme when compared to its warm-active counterpart. The C-terminal of hUDG unfolds at all three simulations temperatures, while the C-terminal of cUDG only melts at 425 K.

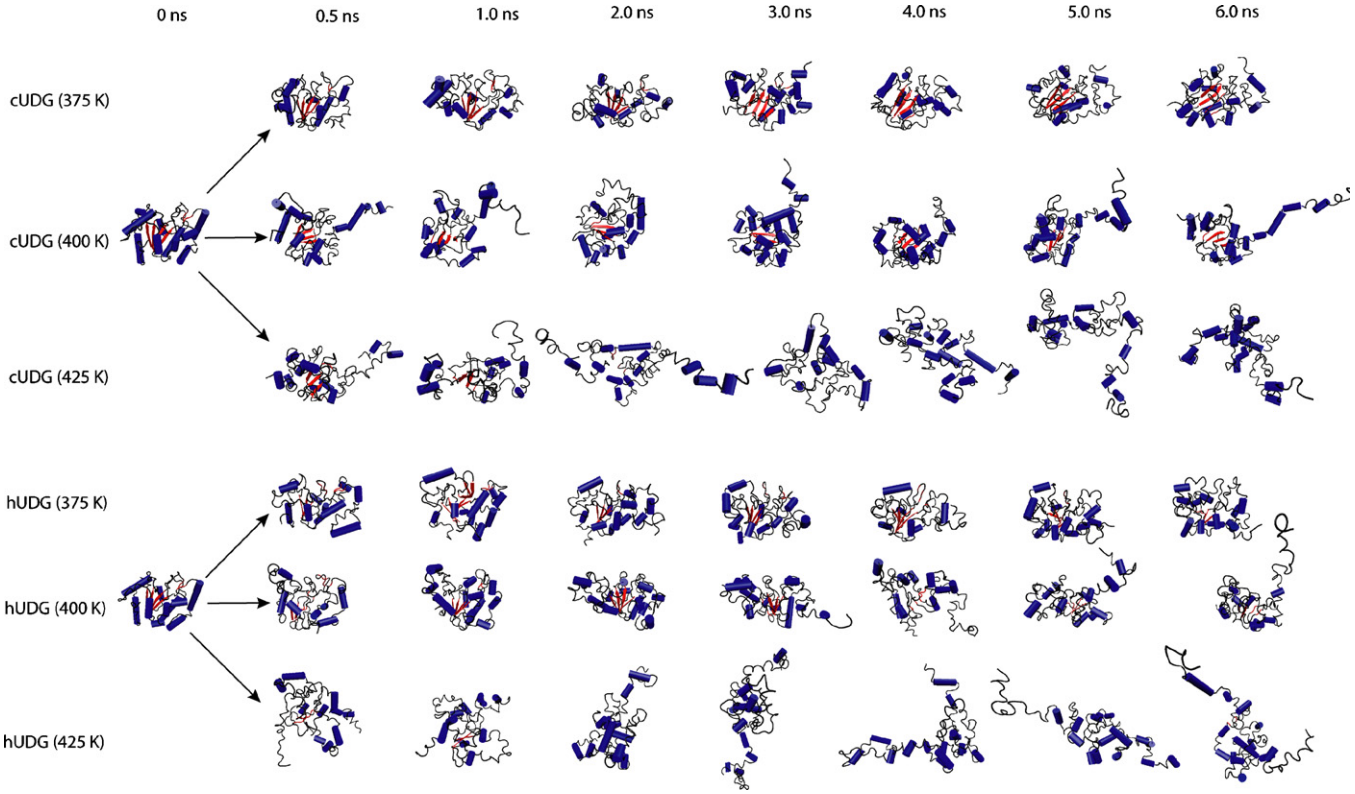

Fig. 7. Snapshots from the thermal unfolding simulations of cUDG and hUDG. The structures are made with the VMD program [46], helices, β-strands and loops are colored blue, red and gray, respectively.

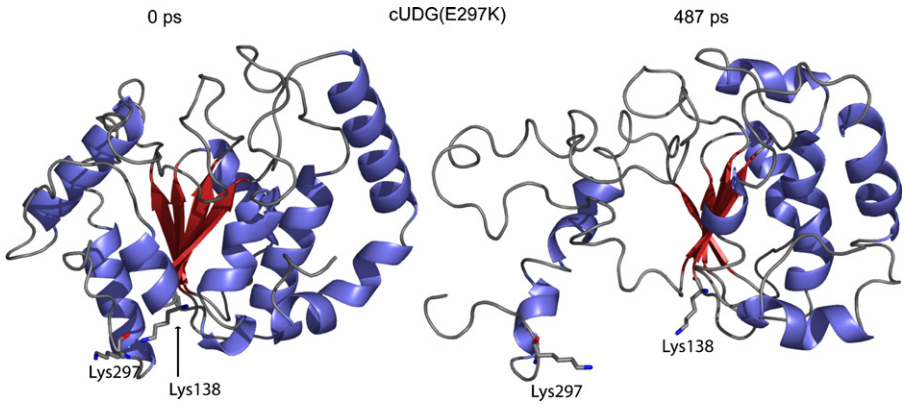

Fig. 8. Snapshots from the unfolding of cUDG(E297K) mutant at 400 K. The figure shows how the C-terminal unfolds from the rest of the structure. Some important residues are shown in ball-and-stick. The figure is generated in PyMOL [45].

Again, slightly different intramolecular interactions are responsible for the observed behavior in the C-terminal of cUDG and hUDG. The 375 K simulation of hUDG shows that around 450–500 ps, the hydrogen bonds between Gln198 and Asp300 disappear. Additional hydrophobic contacts in the C-terminal are also disrupted at around 800 ps, among others the Trp195–Trp301 side chain contact (Fig. 5C). These events seem to initiate melting of the C-terminal. The C-terminal of cUDG does not unfold at all at this temperature, but the last three residues have large fluctuations (Fig. 3A). Increasing the temperature to 400 K result in similar observations of the stability of the C-terminals. After only 140 ps several intermolecular contacts are lost, among others the hydrogen bond between Gln198 and residue 300 (Fig. 5H) and the hydrophobic contact between Trp195 and Trp301 (Fig. 5F). Together, these events lead to a rapid unfolding of the C-terminal as shown in Fig. 6. Again, the C-terminal is stable in the simulation of the psychrophilic enzyme, even though the interactions between residue Gln198 and Asn300 are not present after 500 ps. Favorable hydrophobic packing interactions between several residues in the C-terminal, among others the Trp195–Trp301 residues (Fig. 5F), is a possible explanation of this observation. The 425 K simulations show that both C-terminals unfold, but again with a higher stability in hUDG (Fig. 3E, F and 7).

Structural differences between the two enzymes may explain the differences observed in the stability of the C-terminal. First of all, residue 297 is a Lys in hUDG and a Glu in cUDG. In the crystal structure of cUDG this residue form a weak salt-bridge to Lys138 (5.4 Å between the Lys Nζ atom and the Glu Oε1 atom), while an unfavorable electrostatic interaction between two positively charged residues is observed in hUDG. To further investigate whether this explains the differences in unfolding of the C-terminal, two mutants of the enzymes were modeled from the crystal structures. The Lys297 residue in hUDG was mutated to a Glu and the Glu297 residue in cUDG was mutated to a Lys. Then 6 ns MD simulations of these two mutants were performed at 400 K. The C-terminal of the cUDG(E297K) mutant unfolds after ∼500 ps (Figs. 5J and 8), while the C-terminal of cUDG does not unfold at this temperature at all. This indicates that the weak salt bridge between Lys138 and Glu297 stabilizes the C-terminal in cUDG. The simulation of the hUDG (K297E) mutant shows that the enzyme still unfolds even when the mutations are introduced. Thus, other elements are also likely contributing to the observed difference in the unfolding of the C-terminal. From residue 293 to the C-terminal end, there are four residues in the sequence that make cUDG more hydrophobic than hUDG. The substitutions are K293L, K296T, D300N and E303A (mutation from hUDG to cUDG), all charged residues in hUDG, substituted to hydrophobic or uncharged in cUDG (for a full alignment see Leiros et al. [9]). It seems likely that the hydrophobic contacts are particularly important in the C-terminal, especially the hydrophobic/stacking contact between residue Trp195 and Trp301 (Fig. 5C and F). When this interaction is destroyed the C-terminal easily unfolds (Fig. 6).

6. Other differences in stability

Even if the largest differences in unfolding pathway among the two enzymes are in the terminals, there are also other parts of the structure possessing differences in unfolding pathways. In the 375 K simulation, the mesophilic enzyme unfolds to a high extent in the residues 207–226 region, whereas, the psychrophilic enzyme has a much lower r.m.s.d. per residue in this part of the structure (Fig. 3A and B). One possible reason for this difference is that the salt-bridge between residue Glu112 and Arg210 is not present for most of the simulation of hUDG whereas, it is intact for the first 5.5 ns of the simulation of cUDG (Fig. 5I).

In the 400 K simulation there are also other parts than the terminals of the structure that unfolds. Especially the segment from residue 205 to 220 unfolds rapidly for both enzymes, while in the 165–190 region the cUDG enzyme unfolds to a much higher extent than the mesophilic counterpart (Fig. 3C and D). There are several intermolecular interactions that may explain the difference in unfolding of this area. By analyzing atomic interactions trajectories it seems like residue 183, situated in a sharp turn, could explain much of the observed difference in this area. This residue is an Asp in hUDG and forms a salt bridge with Lys302 that is lost after 80 ps and does therefore probably not explain the observed differences. cUDG has a Gly at position 183 and one can therefore expect that the psychrophilic enzyme

becomes more flexible in this area of the structure. In addition, there is a salt-bridge between Asp180 and Lys/Arg282 (Lys in cUDG and Arg in hUDG), which is lost at around 85 ps for both enzymes. This interaction is however reformed again in the hUDG structure, and is actually intact for large parts of the simulation. Hence, these specific interactions can explain the differences observed in the 165–190 area between the two enzymes. In the 425 K simulation both enzymes unfolds rapidly. After a short period of simulation, both cUDG and hUDG are highly distorted (Fig. 3E, F and 7).

7. Concluding remarks

The high sequence identity and structural similarity between cUDG and hUDG allows for the construction of various hypotheses that can be tested experimentally. Such studies (covering cloning, protein production, purification, characterization, etc.) are demanding in terms of costs and labor. Integrating computational approaches with different levels of complexity with the biochemical and/or biophysical studies, therefore allows in many cases for a more rationalized experimental process. In this study we have used molecular dynamics simulations at high temperatures to investigate not only the specific features of the unfolding mechanism of two homologous DNA repair enzymes, but also to gain a deeper understanding of the structurally important regions modulating thermal stability of the two. It is important to note that the objective of the present study is not by any means to try to solve the problem of protein folding. Instead, these simulations are used as tools to increase our knowledge of the structure–function relationship in uracil DNA glycosylase.

The simulations presented here (at 375 and 400 K) show that the psychrophilic UDG has lower thermal stability compared to the mesophilic UDG, which is in accordance with experiments [13]. The mesophilic enzyme needs higher temperature to reach the same rate of unfolding compared to cUDG. However, once the thermal energy of the systems is sufficiently high (i.e. 425 K in this case), rapid unfolding is observed for both enzymes. Subsequent analysis of the resulting trajectories provides indication of the structural regions important for protein stability. It is particularly the terminals that have different behavior between the cold- and warm-adapted UDG enzymes. Important molecular contacts, probably decisive for the stability of both terminals have been pin-pointed.

Acknowledgements

This work has been supported by a grant from the Research Council of Norway (RCN) (NFR 154197/432), and the Norwegian Structural Biology Centre (NorStruct) is supported by the National program in Functional Genomics (FUGE) in the Research Council of Norway.

References

[1] G. Feller, Molecular adaptations to cold in psychrophilic enzymes, Cell. Mol. Life Sci. 60 (2003) 648–662.
[2] T. Lonhienne, C. Gerday, G. Feller, Psychrophilic enzymes: revisiting the thermodynamic parameters of activation may explain local flexibility, Biochim. Biophys. Acta 1543 (2000) 1–10.
[3] D. Georlette, V. Blaise, T. Collins, S. D'Amico, E. Gratia, A. Hoyoux, J.C. Marx, G. Sonan, G. Feller, C. Gerday, Some like it cold: biocatalysis at low temperatures, FEMS Microbiol. Rev. 28 (2004) 25–42.
[4] C. Gerday, M. Aittaleb, M. Bentahir, J.P. Chessa, P. Claverie, T. Collins, S. D'Amico, J. Dumont, G. Garsoux, D. Georlette, A. Hoyoux, T. Lonhienne, M.A. Meuwis, G. Feller, Cold-adapted enzymes: from fundamentals to biotechnology, Trends Biotechnol. 18 (2000) 103–107.
[5] A.O. Smalås, H.K.S. Leiros, V. Os, N.P. Willassen, Cold Adapted Enzymes, vol. 6, Elsevier Science B.V., Amsterdam, 2000, pp. 1–57.
[6] O. Lanes, I. Leiros, A.O. Smalås, N.P. Willassen, Identification, cloning, and expression of uracil-DNA glycosylase from Atlantic cod (gadus morhua): Characterization and homology modeling of the cold-active catalytic domain, Extremophiles 6 (2002) 73–86.
[7] T. Lindahl, B. Nyberg, Heat-induced deamination of cytosine residues in deoxyribonucleic-acid, Biochemistry 13 (1974) 3405–3410.
[8] C.D. Mol, A.S. Arvai, G. Slupphaug, B. Kavli, I. Alseth, H.E. Krokan, J.A. Tainer, Crystal-structure and mutational analysis of human uracil-DNA glycosylase—structural basis for specificity and catalysis, Cell 80 (1995) 869–878.
[9] I. Leiros, E. Moe, O. Lanes, A.O. Smalås, N.P. Willassen, The structure of uracil-DNA glycosylase from Atlantic cod (gadus morhua) reveals cold-adaptation features, Acta Crystallogr. D 59 (2003) 1357–1365.
[10] R. Savva, K. Mcauleyhecht, T. Brown, L. Pearl, The structural basis of specific base-excision repair by uracil-DNA glycosylase, Nature 373 (1995) 487–493.
[11] R. Ravishankar, M.B. Sagar, S. Roy, K. Purnapatre, P. Handa, U. Varshney, M. Vijayan, X-ray analysis of a complex of *Escherichia coli* uracil DNA glycosylase (ecudg) with a proteinaceous inhibitor. The structure elucidation of a prokaryotic UDG, Nucl. Acids Res. 26 (1998) 4880–4887.
[12] 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.
[13] O. Lanes, P.H. Guddal, D.R. Gjellesvik, N.P. Willassen, Purification and characterization of a cold-adapted uracil-DNA glycosylase from Atlantic cod (gadus morhua), Comp. Biochem. Physiol. 127 (2000) 399–410.
[14] A.R. Fersht, V. Daggett, Protein folding and unfolding at atomic resolution, Cell 108 (2002) 573–582.
[15] R. Scandurra, V. Consalvi, R. Chiaraluce, L. Politi, P.C. Engel, Protein stability in extremophilic archaea, Front. Biosci. 5 (2000) D787–D795.
[16] K.A. Dill, Dominant forces in protein folding, Biochemistry 29 (1990) 7133–7155.
[17] S. Kumar, C.J. Tsai, R. Nussinov, Maximal stabilities of reversible two-state proteins, Biochemistry 41 (2002) 5359–5374.
[18] C.N. Pace, Polar group burial contributes more to protein stability than non-polar group burial, Biochemistry 40 (2001) 310–313.
[19] C.N. Pace, R.W. Alston, K.L. Shaw, Charge-charge interactions influence the denatured state ensemble and contribute to protein stability, Protein Sci. 9 (2000) 1395–1398.
[20] S. Kumar, R. Nussinov, Salt bridge stability in monomeric proteins, J. Mol. Biol. 293 (1999) 1241–1255.
[21] A.J. Li, V. Daggett, Characterization of the transition-state of protein unfolding by use of molecular-dynamics—chymotrypsin inhibitor-2, Proc. Natl. Acad. Sci. U.S.A. 91 (1994) 10430–10434.
[22] A.J. Li, V. Daggett, Molecular dynamics simulation of the unfolding of barnase: characterization of the major intermediate, J. Mol. Biol. 275 (1998) 677–694.
[23] A. Caflisch, M. Karplus, Acid and thermal-denaturation of barnase investigated by molecular-dynamics simulations, J. Mol. Biol. 252 (1995) 672–708.
[24] U. Mayor, N.R. Guydosh, C.M. Johnson, J.G. Grossmann, S. Sato, G.S. Jas, S.M.V. Freund, D.O.V. Alonso, V. Daggett, A.R. Fersht, The complete folding pathway of a protein from nanoseconds to microseconds, Nature 421 (2003) 863–867.

[25] R. Day, B.J. Bennion, S. Ham, V. Daggett, Increasing temperature accelerates protein unfolding without changing the pathway of unfolding, J. Mol. Biol. 322 (2002) 189–203.
[26] D.A. Pearlman, D.A. Case, J.W. Caldwell, W.S. Ross, T.E. Cheatham, S. Debolt, D. Ferguson, G. Seibel, P. Kollman, Amber a package of computer-programs for applying molecular mechanics, normal-mode analysis, molecular-dynamics and free-energy calculations to simulate the structural and energetic properties of molecules, Comput. Phys. Commun. 91 (1995) 1–41.
[27] J.M. Wang, P. Cieplak, P.A. Kollman, How well does a restrained electrostatic potential (resp) model perform in calculating conformational energies of organic and biological molecules? J Comput. Chem. 21 (2000) 1049–1074.
[28] A. Onufriev, D. Bashford, D.A. Case, Modification of the generalized born model suitable for macromolecules, J. Phys. Chem. B. 104 (2000) 3712–3720.
[29] J. Karanicolas, C.L. Brooks, Integrating folding kinetics and protein function: Biphasic kinetics and dual binding specificity in a ww domain, Proc. Natl. Acad. Sci. U.S.A. 101 (2004) 3432–3437.
[30] Y.Z. Ohkubo, C.L. Brooks, Exploring flory's isolated-pair hypothesis: statistical mechanics of helix-coil transitions in polyalanine and the c-peptide from rnase a, Proc. Natl. Acad. Sci. U.S.A. 100 (2003) 13916–13921.
[31] M. Feig, C.L. Brooks, Recent advances in the development and application of implicit solvent models in biomolecule simulations, Curr. Opin. Struc. Biol. 14 (2004) 217–224.
[32] J.P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, Numerical-integration of cartesian equations of motion of a system with constraints—molecular-dynamics of *n*-alkanes, J. Comput. Phys. 23 (1977) 327–341.
[33] H.J.C. Berendsen, J.P.M. Postma, W.F. Vangunsteren, A. Dinola, J.R. Haak, Molecular-dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690.
[34] D. Sitkoff, K.A. Sharp, B. Honig, Accurate calculation of hydration free-energies using macroscopic solvent models, J. Phys. Chem. 98 (1994) 1978–1988.
[35] I. Leiros, O. Lanes, O. Sundheim, R. Helland, A.O. Smalås, N.P. Willassen, Crystallization and preliminary X-ray diffraction analysis of a cold-adapted uracil-DNA glycosylase from Atlantic cod (gadus morhua), Acta Crystallogr. D 57 (2001) 1706–1708.
[36] A.R. Dinner, G.M. Blackburn, M. Karplus, Uracil-DNA glycosylase acts by substrate autocatalysis, Nature 413 (2001) 752–755.
[37] W. Kabsch, C. Sander, Dictionary of protein secondary structure—pattern-recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (1983) 2577–2637.
[38] M.F. Sanner, A.J. Olson, J.C. Spehner, Reduced surface: an efficient way to compute molecular surfaces, Biopolymers 38 (1996) 305–320.
[39] E. Moe, I. Leiros, E.K. Riise, M. Olufsen, O. Lanes, A. Smalas, N.P. Willassen, Optimisation of the surface electrostatics as a strategy for cold adaptation of uracil-DNA *n*-glycosylase (UNG) from Atlantic cod (gadus morhua), J. Mol. Biol. 343 (2004) 1221–1230.
[40] V. Daggett, A. Fersht, The present view of the mechanism of protein folding, Nat. Rev. Mol. Cell Biol. 4 (2003) 497–502.
[41] V. Daggett, Molecular dynamics simulations of the protein unfolding/folding reaction, Acc. Chem. Res. 35 (2002) 422–429.
[42] C.J. Bond, K.B. Wong, J. Clarke, A.R. Fersht, V. Daggett, Characterization of residual structure in the thermally denatured state of barnase by simulation and experiment: description of the folding pathway, Proc. Natl. Acad. Sci. U.S.A. 94 (1997) 13409–13413.
[43] J. Hoseki, A. Okamoto, R. Masui, T. Shibata, Y. Inoue, S. Yokoyama, S. Kuramitsu, Crystal structure of a family 4 uracil-DNA glycosylase from *Thermus thermophilus* hb8, J. Mol. Biol. 333 (2003) 515–526.
[44] A.J. Li, V. Daggett, Identification and characterization of the unfolding transition state of chymotrypsin inhibitor 2 by molecular dynamics simulations, J. Mol. Biol. 257 (1996) 412–429.
[45] W.L. DeLano, The Pymol Molecular Graphics System, DeLano Scientific, San Carlos, CA, USA, 2002.
[46] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38.