Molecular dynamics studies on the thermostability of family 11 xylanases

Mikko Purmonen1, Jarkko Valjakka2,3, Kristiina Takkinen4, Tuomo Laitinen5 and Juha Rouvinen1,6

1Department of Chemistry, University of Joensuu, PO Box 111, 80101 Joensuu, Finland, 2Institute of Medical Technology, University of Tampere, 33014 Tampere, Finland, 3Tampere University Hospital, Biokatu 6-8, 33520 Tampere, Finland, 4VT Technical Research Centre of Finland, PO Box 1000, 02044 VTT, Finland and 5Department of Pharmaceutical Chemistry, University of Kuopio, PO Box 1627, 70211 Kuopio, Finland

6To whom correspondence should be addressed.
E-mail: juha.rouvinen@joensuu.fi

Twelve members of the family 11 xylanases, including both mesophilic and thermophilic proteins, were studied using molecular dynamics (MD). Simulations of xylanases were carried out in an explicit water environment at four different temperatures, 300, 400, 500 and 600 K. A difference in thermostolerance between mesophilic and thermophilic xylanases became clear: thermophilic xylanases endured heat in higher simulation temperatures better than mesophilic ones. The unfolding pathways seemed to be similar for all simulations regardless of the protein. The unfolding initiates at the N-terminal region or alternatively from the α-helix region and proceeds to the ‘finger region’. Unfolding of these regions led to denatured structures within the 4.5 ns simulation at 600 K. The results are in agreement with experimental mutant studies. The results show clearly that the stability of the protein is not evenly distributed over the whole structure. The MD analysis suggests regions in the protein structure which are more unstable and thus potential targets for mutation experiments to improve thermostability.

Keywords: denaturation/molecular dynamics/thermostability/unfolding pathway/xylanase family 11

Introduction

The use of many enzymes in industrial applications requires high thermostability. Thermostable enzymes have been created from mesophilic enzymes by using site-directed or random mutagenesis. One alternative is to use natively thermostable enzymes from corresponding organisms. Whereas mesophilic organisms live near room temperature, ‘heat-loving’ thermostophilic organisms live in considerably higher temperatures (optimal growth 50–80°C) and hyperthermophilic organisms in even more extreme temperatures (optimal growth >80°C) (Collins et al., 2005). Hyperthermophilic organisms are usually isolated from hot springs and other geothermally heated sources (Rees and Adams, 1995). The structures of natively thermostable proteins have shed light on these factors that contribute to the thermostability of proteins.

With molecular dynamics (MD), it is possible to study cases that are difficult to investigate experimentally, e.g. protein folding and unfolding. MD offers a detailed description of thermal motion at different temperatures and by following the conformational changes as a function of small time steps, it is possible to study, for example, unfolding or denaturation (irreversible unfolding) pathways of proteins. MD has been used a few times earlier to study protein unfolding. Chymotrypsin inhibitor 2 (CI2) was simulated at seven temperatures, ranging from 290 to 498 K and from 94 to 20 ns. It was proposed that an increase in temperature can accelerate protein unfolding without changing the unfolding pathway of proteins (Day et al., 2002). Para-nitrobenzyl esterase was simulated at 300 K until 500 ps using MD (Wintrode et al., 2003). The data suggested that adaptation for high temperature stability has resulted in a structure in which larger deviations from the native state are restricted and the conformational flexibility of thermophilic enzymes at room temperature is reduced, in comparison with the mesophilic ones. MD simulations in solution have also been performed for a rubredoxin from a hyperthermophilic archaean Pyrococcus furiosus (RdPf) and from a mesophilic organism Desulfovibrio vulgaris (RdDv) (Lazardis et al., 1997). Temperatures of 300, 375, 473 and 500 K were simulated and simulations extended from 0.2 to 1.0 ns. Although the unfolding behavior of RdPf and RdDv seemed to be very similar, simulation pathways differed slightly from each other. MD simulations have also been used as a tool to study the thermostability of haloalkane dehalogenase (DhIA), by identifying flexible regions from the structure (Pikkemaat et al., 2002). These results showed that it was possible to use MD simulations as a tool to identify flexible protein domains that can be used as a target for stability enhancement by introducing a disulfide bond. MD simulations at 300 K, 400 K, 600 K, and 800 K have also been used to study the thermostability of glucoamylase from Aspergillus awamori. An improvement in thermostability was achieved by introducing disulfide bond to lock one of the α-helices on the surface of the catalytic domain (Liu and Wang, 2003). In addition, some point mutations were tested for the same protein by using MD.

Xylanases have applications in food, feed and technical industries. For example, family 11 endo-β-xylanases are widely used in the pulp industry, e.g. in the biobleaching of craft pulp. Xylanases are also used in the baking industry to improve the handling of dough and the quality of the final baked product (Beg et al., 2001). The addition of xylanase to the rye-based diet of broiler chickens improved weight gain, feed intake and feed efficiency, and decreased water consumption (Mathlouthi et al., 2002).

We have previously determined several crystal structures of small-molecular weight family (11) endo-1,4-β-xylanases (EC 3.2.1.8). (Törrönen et al., 1994; Törrönen and Rouvinen, 1995; Hakulinen et al., 2003). Endo-1,4-xylanases break up xylan, a major hemicellulose polysaccharide in plant cell walls, into xylo-oligosaccharides with variable lengths. Xylans are heteropolysaccharides consisting of a β–xylopyranosyl backbone with variable side chains. Backbone
substituents depend on the source of the polymer, e.g., xylan received from softwood and hardwood differs significantly from each other. The most common substituents are α-linked 4-O-methyl glucuronic acid on C2, α-linked arabinose on C2 or C3, and acetyl esters on C2 or C3 of some xylose residues (Brett and Waldron, 1996). Xylanases are created from an α-helix and two β-sheets, A and B, forming a so-called β-sandwich structure (Fig. 1). The shape of xylanases resembles a right hand, where β-sheet A is composed of five antiparallel β-strands from A1 to A6 (sometimes the A1 strand is lacking among the protein family 11 xylanases and there is only a loop region, e.g. TRX II in Fig. 1) and β-sheet B contains β-strands B1–B9. The commonly used names for different parts of the xylanases are ‘fingers’, ‘palm’, ‘thumb’ and ‘cord’. Fingers are formed from A and B β-sheets, the palm is made up of an α-helix and a twisted β-sheet B, the thumb is a loop between strands B7 and B8, partly closing the cleft, and the cord is a long loop between B6 and B9 partly closing the side of the cleft (Törnönen et al., 1994). This family of xylanases consists of ~190–200 amino acids and has a diameter of 35 Å (12). Xylanases’ dimensions vary depending on their amount of amino acid residues, packing and the number and size of cavities. In a study by Hakulinen and co-workers, factors that affect the thermostability of family 11 endo-1,4-β-xylanases were represented by analyzing similarities and differences both at the amino acid sequence and three-dimensional structure level (Hakulinen et al., 2003). Several minor modifications appeared to be responsible for increased thermostability, such as a higher Thr/Ser ratio, an increased number of charged residues and a higher number of residues in the β-strands and stabilization of the α-helix region.

In this study, MD have been used to study the unfolding of the family 11 xylanases. The simulations were carried out at different temperatures, 300, 400, 500 and 600 K. High-temperature simulations eventually led to the breakdown of the tertiary structure of protein. An analysis of the results will give an estimation of the initial steps of denaturation in this protein family. We conducted MD simulations for 12 members of family 11 xylanases of which three-dimensional structures are available in the Protein Data Bank (Berman et al., 1994). The studied xylanases were from Trichoderma reesei (I and II, TRX I and TRX II) (Törnönen et al., 1994; Törnönen and Rouvinen, 1995), Bacillus circulans (BCX) (Wakarchuk et al., 1997), Aspergillus kawachiii (AKX) (Fushinobu et al., 1998), Aspergillus niger (ANX) (Krengel and Dijkstra, 1986), Bacillus agaradaraeens (BAX) (Sabini et al., 1999), Trichoderma harzianum (THX) (Campbell et al., 1993), Chaetomium thermophilum (CTX) (Hakulinen et al., 2003), Dictyoglomus thermophilum (DTX) (McCarthy et al., 2000), Thermomyces lanuginosus (TLX) (Gruber et al., 1998), Paecilomyces Varioti Bainier (PVX) (Kumar et al., 2000) and Nonomuraea flexuosa (NFX) (Hakulinen et al., 2003). A list of the studied xylanases, their PDB-codes and optimum growth temperatures are shown in Table I. A sequence comparison of the studied family 11 xylanases is shown in Fig. 2. These enzymes can be divided into mesophilic (TRX II, BCX, THX, TRX I, AKX, ANX and BAX) and thermophilic (NFX, CTX, DTX, TLX and PVX) enzymes. Whereas BAX, DTX, BCX and NFX were from bacterial sources, the other studied xylanases were from a fungal source (Hakulinen et al., 2003).

Methods

The MD simulations were carried out using the Amber 7 program (Assisted Model Building with Energy Refinement) (Pearlman et al., 1995; Case et al., 2005) installed on our Dell Precision computers operating under Linux with Inter(R) Xeon(TM) 2.40 GHz processors. Atom interactions were described with Amber ff99 (Wang et al., 2000). First, before starting MD, all bad contacts need to be removed, exclusively minimizing water molecules and the whole system, containing protein and water, without SHAKE (Ryckaert et al., 1977). In addition to minimization (0.3 ns), the system was also heated and brought into equilibrium before MD simulation (15 ps). Simulations were performed using a canonical NVT ensemble. A constant temperature for the simulated system was implemented with a weak-coupling algorithm, with which the system was coupled to an external heat bath that was fixed at the desired temperature using a default time constant (τT = 1.0 ps) for the heat bath coupling for the system (Berendsen et al., 1984). The solvent around the proteins was described with explicit water molecules consisting of TIP3P water molecules (Jorgensen et al., 1983).

![Fig. 1. (A) Topology diagram and (B) a schematic 3D structure of the family 11 xylanases. Family 11 xylanases are constructed with two β-sheets and an α-helix forming a so-called β-sandwich structure. The different parts of the enzyme are named after the right hand (Törnönen et al., 1994), thus containing regions named fingers, thumb, palm, cord and α-helix.](https://academic.oup.com/peds/article-abstract/20/11/551/1563006/Molecular-dynamics-studies-on-the-thermostability/552)
The resultant structure was a protein surrounded with water molecules. The range of the truncated octahedron solvent buffer was 10 Å from the surface of the protein, so the total amount of the water molecules varied from 4278 to 5494 molecules, depending on the area and shape of the molecule surface. Sodium and chloride ions were used as counter-ions in order to uncharge the molecules. The number of counter-ions varied from 0 to 14 among the xylanases. A cut-off distance of 8 Å was used to carryout long-range electrostatics with Particle Mesh Ewald (PME) and for van der Waals forces. Simulations were conducted in periodic boundary conditions (PBC) using SHAKE, in which bonds involving hydrogen atoms were constrained. Simulations were carried out for 4.5 ns including 3 000 000 cycles. The magnitude of a time step was 1.5 fs. Every 500th step was saved to the trajectory, resulting in a set of 6000 coordinates for the analysis of the unfolding pathway. The results were analyzed using an Amber module, ptraj. In addition, programs MOIL-View (Simmerling et al., 1995), VMD and Microsoft Excel were used to analyze trajectories. Each simulation required ~1 month of CPU time, which corresponds a 4.5 ns computational model of unfolding. Other programs used in this study were ClustalW 1.8.1, BOXSHADE 3.2.1, PyMOL 0.97 (DeLano Scientific) and HBPLUS, below described.

Density of the system was calculated to be ~1.04 g/cm³ at all temperatures (300, 400, 500 and 600 K). The HBPLUS program uses the same minimum angles and maximum distances as Baker and Hubbard (Baker and Hubbard, 1984). Minimum angles between D-H-A, H-A-AA and D-A-AA are 90.0°, maximum angles for amino-aromatic interactions D-A-AX and H-A-AX are 20.0° and maximum distances for D-A, H-A, and S-S are 3.9, 2.5 and 3.0 Å, respectively.

Table I. Studied family 11 xylanases and their properties

<table>
<thead>
<tr>
<th>Organism</th>
<th>Symbol</th>
<th>PDB code</th>
<th>Number of amino acid</th>
<th>Optimum growth (°C)</th>
<th>pH optimum</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Thermophilic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Dictyoglomus thermophilum</em></td>
<td>DTX</td>
<td>1F5J</td>
<td>199</td>
<td>75</td>
<td>6.5</td>
<td>Sapag et al. (2002)</td>
</tr>
<tr>
<td><em>Nonomuraea flexuosa</em></td>
<td>NFX</td>
<td>1M4W</td>
<td>197</td>
<td>85</td>
<td>6.0-7.0</td>
<td>Xiong et al. (2004)</td>
</tr>
<tr>
<td><em>Thermomyces lanuginous</em></td>
<td>TLX</td>
<td>1YNA</td>
<td>194</td>
<td>70-80</td>
<td>6.5</td>
<td>Xiong et al. (2004)</td>
</tr>
<tr>
<td><em>Paeclomyces variot</em></td>
<td>PVX</td>
<td>1PVX</td>
<td>194</td>
<td>50</td>
<td>4.0</td>
<td>Beg et al. (2001)</td>
</tr>
<tr>
<td><em>Chaetomium thermophilum</em></td>
<td>CTX</td>
<td>1H1A</td>
<td>191</td>
<td>70-75</td>
<td>5.5-7.0</td>
<td>Sapag et al. (2002)</td>
</tr>
<tr>
<td><strong>Mesophilic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Bacillus agaradhaerens</em></td>
<td>BAX</td>
<td>1QH7</td>
<td>207</td>
<td>40-45</td>
<td>4.0-4.5</td>
<td>Beg et al. (2001)</td>
</tr>
<tr>
<td><em>Trichoderma reesei II</em></td>
<td>TRX II</td>
<td>1ENX</td>
<td>190</td>
<td>45</td>
<td>4.5-5.5</td>
<td>Kumar and Nussinov (2001)</td>
</tr>
<tr>
<td><em>Trichoderma harzianum</em></td>
<td>THX</td>
<td>1XND</td>
<td>190</td>
<td>50</td>
<td>5.0</td>
<td>Beg et al. (2001)</td>
</tr>
<tr>
<td><em>Bacillus circulanus</em></td>
<td>BCX</td>
<td>1XNB</td>
<td>185</td>
<td>45</td>
<td>6.0-7.0</td>
<td>Srinivasan and Rerle (1999)</td>
</tr>
<tr>
<td><em>Aspergillus kawachi</em></td>
<td>AKX</td>
<td>1BK1</td>
<td>182</td>
<td>50</td>
<td>2.0</td>
<td>Christov et al. (1999)</td>
</tr>
<tr>
<td><em>Aspergillus niger</em></td>
<td>ANX</td>
<td>1UKR</td>
<td>181</td>
<td>60</td>
<td>3.0</td>
<td>Sapag et al. (2002)</td>
</tr>
<tr>
<td><em>Trichoderma reesei I</em></td>
<td>TRX I</td>
<td>1XYN</td>
<td>178</td>
<td>45</td>
<td>3.0</td>
<td>Beg et al. (2001)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>50</td>
<td>4.0</td>
<td>De Vries and Visser (2001)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>24 h at 50°C</td>
<td>3.5-4.0</td>
<td>Lappalainen et al. (2003)</td>
</tr>
</tbody>
</table>

Thermophilic and mesophilic xylanases are divided into separate categories. Enzymes can be found from the Protein Data Bank using the PDB codes mentioned above. Optimum growth temperatures have been found from the literature according to references.

Results

The simulations at different temperatures

The denaturation of protein typically occurs on a micro-second time scale (Duan and Kollman, 1998). Unfortunately, it was not possible to perform MD simulations within the typical time period in which proteins normally denaturate. Therefore, shorter simulation times are needed, and it is necessary to increase the temperature to detect the denaturation of protein. Although the drawback is that the system is more artificial, it is probably more realistic at the beginning of the simulation, so it may reveal details of the initial unfolding process which might be the most interesting when one considers the use of mutagenesis in improving thermostability.

In total, 48 simulations were made for 12 xylanases at 300, 400, 500 and 600 K. The differences between native structures and simulated structures were investigated with CA RMS deviations. RMS deviation is a numerical measure of the difference between two structures. An example of the effect of higher temperatures is shown in Fig. 3, where higher temperatures inflict a more rapid denaturation. As can be seen in Fig. 3, the RMS deviation values of mesophilic TRX II are considerably higher than the corresponding RMS deviation values of DTX at 600 K. In addition, a notable observation was found also for other thermophilic xylanases.

The effect of increased temperature to the protein structure

A more comprehensive picture was obtained from the simulation when the RMS deviation values were presented as a function of the amino acid residue. Figure 4 shows the...
This presentation clearly highlights some areas in which differences from the native structure are large. These areas are the ‘thumb’, α-helix, ‘cord’ and finger/N-terminal areas. One also see that simulations at 300, 400 and 500 K are quite similar, but the 600 K simulation shows markedly larger changes.

An even more complete view can be obtained if the RMSD values are presented as a function of time and residue number simultaneously. This is shown in Fig. 5, where RMSD plots of TRX II simulations are represented at four different temperatures, 300, 400, 500 and 600 K. On the x-axis there is an amino acid sequence, on the y-axis time and on the z-axis RMS deviation values. If RMS deviations fluctuate substantially in certain regions, it can indicate those areas where there are reversible conformational changes. If RMS deviations rise but do not descend back to the initial level, this can indicate an irreversible change in the structure and lead to denaturation.

At 300 K, as expected, either mesophilic or thermophilic xylanases did not show any signs of denaturation at the end point of the simulation (4.5 ns); the structures were stable and remained close to the crystal structure during the whole simulation.
unfolding yet. At 500 K, local unfolding occurred mainly in the mesophilic but to some extent also in the thermophilic xylanases and at 600 K, we observed larger unfolding which led to comprehensive denaturation, especially in mesophilic xylanases. Consequently, on a 4 ns time scale only the use of 600 K causes denaturation of the protein structure and offers a sketch of the denaturation pathway.

The influence of the initial structure

The initial three-dimensional structures of the studied protein family 11 xylanases were obtained from the Protein Data Bank. However, protein molecules may exist natively in different conformations. For example, protein crystal structures often contain more than one molecule in the asymmetric unit which may have different conformations. We studied the influence of the initial structures by performing simulations for TRX II and DTX with the A and B molecules of the asymmetric unit at 600 K. Although the structures are similar (RMSD 0.1 Å), there were differences in the simulations. In TRX II, molecule B unfolded slightly faster and ended the structure in which the average RMSD was ~2 Å higher. On the other hand, the unfolding pathways were similar regardless of the choice of the molecule. We also modified the structures of TRX II with the program Moleman (Kleywegt et al., 2001), which enables the random shifting of CA atoms. The use of an average 0.1 Å change in the structure did not significantly change the unfolding behavior. An average 0.2 Å change, however, caused serious structural problems in the simulations indicating that the changes in the structure were too large. The influence of different minimization times (0.3, 0.6 and 0.9 ns) was also studied for TRX II and DTX at 300 K. Again, RMSD values between different simulations varied between 1 and 2 Å.

Analysis of hydrogen bonds

Intramolecular hydrogen bonding is known to improve the thermostability of proteins and to increase fractional polar surfaces (Vogt et al., 1997). Intramolecular hydrogen bonds were calculated from pdb-structures with the HBPLUS program with default parameters at different time spots (McDonald and Thornton, 1994). Generally, thermophilic xylanases of family 11 have more hydrogen bonds and slightly more hydrogen bonds per residue in their initial structures than mesophilic ones. With mesophile proteins the number of hydrogen bonds vary from 157 to 214 corresponding to 0.88–1.03 hydrogen bonds per residue, whereas among thermophiles it is from 184 to 205, with a correspondence of 0.96–1.03. The number of intramolecular hydrogen bonds decreased during simulations at 600 K as the RMS deviations increased. At the beginning of the simulations, there is a rapid 14–26% decrease in the number of hydrogen bonds. Later, the decrease in the number of hydrogen bonds is smaller. The differences in the number of hydrogen bonds between mesophilic and thermophilic xylanases during simulations are small and within error limits of the simulations.

Discussion

The influence of the initial structure

The influence of the initial structure on the MD simulation results has not been extensively studied. However, some previous papers have shown that the differences in initial conformation can lead to quite different results in simulations. In the case of acetyl-CoA dehydrogenase, the key residue was in bad conformation, which make it difficult to investigate the reaction mechanism theoretically. In the second case, it was concluded that different initial structures of xylose isomerase from S. olivochromogenes affect the simulation results of the two Mg²⁺ cofactors, substrate and...
enzymatic residues during hydride transfer (Garcia-Viloca et al., 2004). We tested the effect of similar initial structures using A and B molecules of TRX II and DTX alike and used structures distorted by the program called Moleman. The significance of the initial structure for the simulation is slight. The differences in the total energy are due to the different numbers of water molecules in the simulation. In any case, care should be taken when defining the initial coordinates for simulations. In the Protein Data Bank, there are several structures in which part of the protein structure, e.g. some loops, can be disordered or completely missing in the model. Therefore, it is much more challenging to perform simulations for such models.

**Time scale and the effect of the simulation temperature**

At present, nanosecond simulations are routine but they are still too short to observe events in protein motion that occur on the microsecond to second time scale (Daggett, 2000). Because of a lack of sufficient computer power, we were forced to use higher simulation temperatures to accelerate the protein unfolding. Previous studies have suggested that denaturation, using higher temperatures, happens without changing the unfolding pathway (Daggett et al., 2002). Daura and co-workers have also shown that protein folding and unfolding are independent of the temperature and the most common intermediates are the same at different temperatures (Daura et al., 1999). The use of high temperatures allows real-time MD unfolding simulations to be tested in some cases. For example, the engrailed homeodomain protein has very-high refolding and folding rate constants and the half-life of unfolding has been extrapolated to be \( \sim 7.5 \text{ ns} \) at 100°C, which enables real-time MD simulations to be conducted (Mayor et al., 2000). For testing, we have also made simulations in the longer time scale (9 ns) at 400 K for TRX II (mesophilic) and DTX (thermophilic). The RMSD plot shows that DTX protein stays in a folded state but RMSD values continue to increase for TRX II indicating that in the future with better computing power, it might be possible to use lower (more realistic) temperature and longer simulation time to study thermostability.

MD simulations of the studied family 11 xylanases were performed at four different temperatures, 300, 400, 500 and 600 K. By increasing the simulation temperature, it allows us the possibility to observe many interesting events, e.g.

---

**Fig. 5.** RMSD plots of TRX II at different temperatures including (A) 300, (B) 400, (C) 500 and (D) 600 K. The x-axis represents amino acid sequence number, the y-axis time (trajectory number). Information used to construct this plot was saved every 500 time-steps, so 6000 trajectories correspond to 4.5 ns, and the z-axis represents the amount of movement (Å). White, grey, yellow, red and black represent the amount of movement from 0 to 3, 3 to 6, 6 to 9, 9 to 12 and over 12 Å, respectively. Values have been calculated with the use of CA atoms.

---

M.Purmonen et al.,

556

Downloaded from https://academic.oup.com/peds/article-abstract/20/11/551/1563006/Molecular-dynamics-studies-on-the-thermostability
by guest
on 16 September 2017
protein unfolding on a much shorter time scale. The use of a significantly high temperature (600 K) is well founded because the denaturation of some structures can be obtained in 4.5 ns. The 300 K simulation can be used to represent native thermal fluctuation. The 400 K simulations do not greatly differ from the native thermal fluctuation during simulation time. Instead, the differences for 400 and 500 K simulations are easier to notice simply because at 500 K simulations local denaturations happen. The highest simulation temperature (600 K) represents highly vigorous circumstances in which global denaturation happens within 4.5 ns. When comparing simulation results calculated at different temperatures, we can see that structural changes occur in the same regions regardless of the temperature. The only difference is that the structural changes have proceeded the furthest at 600 K. Therefore, the denaturation pathway of family 11 xylanases seems to be highly similar regardless of the used simulation temperature and different initial structures of xylanases.

Structural changes at high temperatures in different regions

Structural changes during the simulations were screened by plotting RMSD values as a function of amino acid residues and time, calculating intramolecular hydrogen bonds, and visualization, by using the VMD program (Humphrey et al., 1996). The unfolding of mesophilic TRX I and thermophilic NFX xylanases are shown in Fig. 6 as snapshots. In the following, we describe in detail general structural changes in different regions of xylanases.

The N-terminal end is a region that is constructed from loops of variable length and A1 (not in all structures), A2, A3, B1, B2 and B3 β-strands. Irreversible structural change of the N-terminal can lead to a comprehensive denaturation. Denaturation can happen when β-strands are gradually lost, starting from the N-terminus (ANX and TRX I) and the structural changes can proceed to the whole finger area (CTX). In some cases, even though the RMSD values of the finger area are large, an increase in the movement of the N-terminal does not lead to protein denaturation (NFX, simulation at 500 K). As we consider the importance of the N-terminal and finger of xylanase family 11, one should note that thermophilic xylanases have generally longer N-terminal regions than mesophilic ones. Mesophilic xylanases often have a long, structurally irregular loop in the N-terminal, whereas thermophilic xylanases commonly have a β-strand in the same position. This may explain larger RMSD values for mesophilic xylanases in the N-terminal. There is also experimental evidence that the N-terminal affects the thermostability of protein family 11 xylanases. TRX II mutants DS1 and DS5 have been stabilized with a C110–C154 disulfide bridge and other site-directed mutations, whereas the highly thermostable DB1 mutant was a combination of the DS5 mutant improved with the N-terminal C2–C28 disulfide bridge (Turunen et al., 2001; Jänis et al., 2004; Xiong et al., 2004). The importance of the N-terminal has also been shown in the study of Sun et al., where the thermostability and catalytic activity of A. niger xylanase A (AnxA) was improved by an N-terminus substitution with the corresponding region of Thermomonospora fusca xylanase A (TrxA) (Sun et al., 2005). Fenel et al. have also stabilized the structure of TRX II by engineering a disulfide bridge into the N-terminal (Fenel et al., 2004). A combination of weakly stabilizing mutations with a disulfide bridge in the α-helix structure of TRX II have been studied and produced experimentally; several combinations of mutations substantially increased thermostability in comparison with the wild type TRX II (Turunen et al., 2001).

The finger region is a stable region that is made up of several β-strands and β-turns, including A4, A5 and A6 strands, which are connected with loops to B4, B5 and B6 strands. The finger region forms a central part of the hydrophobic core of the protein. Consequently, the loss of its structure is important for the denaturation of protein family 11 xylanases. If the β-sheet structure is lost, the whole protein structure is threatened by comprehensive collapse (MD simulations of CTX, BAX, BCX, ANX and TRX I at 600 K). If the structure in the finger region stayed native-like, denaturation of the protein structure did not occur (4.5 ns MD simulations of DTX, NFX and TLX at 600 K). The denaturation of the finger region can proceed gradually from the N-terminal end (MD simulations of ANX and TRX I at 600 K) or the whole finger region can lose the structure at once (CTX).

An α-helix is situated at the back of the xylanases behind the palm. Although this helix structure is very stable in the simulations, its C-terminal end is able to fluctuate. Experimental mutations have proven the importance of this region to the thermostability of protein family 11 xylanases (Turunen et al., 2001; Jänis et al., 2004; Xiong et al., 2004), and xylanases in general (Wakarchuk et al., 1997; Georis et al., 2000). However, the unfolding does not in most of the cases begin from here, according to the simulations. The most well known modification in this region is the C110–C154 disulfide bridge which significantly stabilizes the structure. TLX and PVX have a native C110–C154 disulfide bridge which may have a significant role in addition with a stable N-terminal to in preventing the unfolding.
The cord is a flexible loop connecting B6 and B9 strands. The cord is placed on the outer surface of the protein and is connected to the movements of the thumb, fingers, and α-helix. In the case of the CTX cord area, the 'cord loop', the B9 strand, the N-terminal of B8 and the loop connecting B9 and B8 areas all seem to especially affect denaturation.

The thumb is a short loop, which is the most fluctuating region of xylanase family 11, placed between B7 and B8 strands. The movement of the loop affects the movement of B8 and B7 strands. Mobility is probably essential for the function of the enzyme (Muilu et al., 1998). The fluctuation of the thumb region seems to be widespread in family 11 xylanases. However, the fluctuation of the thumb does not affect the rest of the structure, which remains stable. This is seen in the 600 K simulations of thermostable xylanases, as e.g. in DTX.

The palm is the part of the hydrophobic core and the most stable region in the protein family 11 xylanases. It consists of a β-sheet B containing β-strands B4–B9 coiled into a form of a palm. The region consisting of the longest β-strands B6 and B7 is especially stable. In several MD simulations, the palm region was the last region to maintain the native structure, e.g. CTX, BCX and TRX II. In xylanase family 11, catalytic amino acids are situated in the palm cleft (Törnönen et al., 1994).

Structural factors affecting thermostability

Unfortunately, similar thermostability parameters are not available for all the studied xylanases. The only experimental parameter that can be used to evaluate relative thermostolerance is the optimum growth temperature of an organism, which is available for all cases except for BAX. Correlation between the average RMSD values and temperatures is presented in Fig. 7. The RMSD values represent average values during the whole 600 K simulation. There is a correlation between the RMSD values and growth temperature. However, it must be kept in mind that because RMSD values between different test simulations varied between 1 and 2 Å, this comparison should be considered only as an indicative one. Most of the thermophilic enzymes have lower RMS deviation values than the mesophilic ones, which correspond to the experimental optimum temperature data. PVX, DTX, NFX and TLX are the most thermostable structures according to the MD simulations. AKX and molecule A from TRX II with the immediate presence of CTX and BCX follow the most thermostable structures, leaving THX, ANX and TRX I to be the least thermtolerant structures. BAX is not included in the dissection because of a lack of experimental data, but it would be placed on the thermophilic side of the segment of the line rather than on mesophilic side.

Crystal analysis, sequence comparison and mutagenesis research shows that mesophilic and thermophilic xylanases of family 11 are very similar (Fig. 2). Comparison of the amino acid sequence was made with programs BOXSHADE 3.2.1 written by K. Hofmann and M. Baron, and ClustalW 1.8.1. (Higgins et al., 1994). As a consequence of the similarity between mesophilic and thermophilic xylanases, it is likely that the increase in thermostability derives from several small modifications. Although factors that affect thermostability have been studied for a long time, none of the studies have been able to offer a prioritized list of all factors.

The number of hydrogen bonds is known factor that may affect the thermostability of proteins. It was found that thermophilic xylanases typically contain more hydrogen bonds natively, but the differences in the amounts of hydrogen bonds during simulation is difficult to estimate because of the error in simulation. The simulations pointed out the importance of the N-terminal for the initiation of unfolding. All the thermophilic enzymes (NFX, CTX, DTX, TLX and PVX) have the B1 β-strand containing five residues, whereas among the mesophiles, only BAX, TRX II and THX have the B1 strand containing five, five and four amino acids, respectively. In addition, thermophilic enzymes have charged residues at the N-terminal. NFX has acidic Asp12 and CTX basic His11 in the B1 strand. TLX and PVX contain both acidic (Glu7) and basic (His10) amino acids at the B1 strand, whereas the only thermophilic enzyme that does not have acidic/basic residues in the B1 strand is DTX. BAX is the unique enzyme from the mesophilic xylanases, containing one basic (His11) amino acid. Another interesting observation is that TLX and PVX have a significant conserved sequence including B1 and B2 strands; only Gly1 of PVX is replaced with PCA1 compared with the TLX structure. Both TLX and PVX show outstanding thermostolerance and maintain their common xylanase shape during the simulations. In addition, the native C110–C154 disulfide bridge stabilizes both structures.

In this study, we used MD simulations for the first time to study the unfolding and denaturation of a large enzyme family. Because a large number of proteins, five thermophilic and seven mesophilic enzymes were studied, we can try to sketch a general view of the structural factors which contribute to the unfolding of family 11 xylanases. The most interesting finding is the importance of the N-terminal and α-helix for the initiation of the unfolding process. This supports previously published experimental data which discussed the stabilizing effect of a disulfide bridge mutated in the N-terminal of family 11 xylanases (Turunen et al., 2001; Jänis et al., 2004; Xiong et al., 2004). Simulations also revealed that the palm area is the most resistant against unfolding. Interestingly, this is also the area in which catalytically important residues are situated.

The MD simulations used in this study may offer an alternative strategy for the study of factors that affect the thermostability of proteins. MD simulations can offer a...
References


Funding

Finnish National Technology Agency (TEKES).

Received December 15, 2006; revised September 20, 2007; accepted September 25, 2007

Edited by Jane Clarke