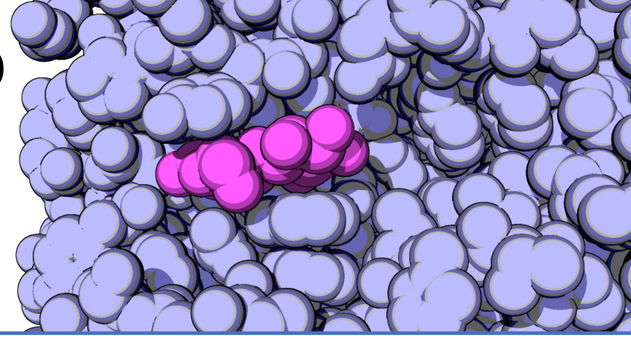
Molecular Dynamics Simulations Provide Insights into Stability of Hyperthermophilic Proteins

Logan Sheffield, Alexandra Martynova-Van Kley, Armen Nalian
Stephen F. Austin State University, Department of Biology



Introduction

Endoglucanases are enzymes that hydrolyze internal β-(1,4)-glycosidic bonds between the glucose monomers of cellulose. This hydrolysis plays a key role in the production of bioethanol, a renewable fuel source with lower greenhouse gas emissions than that of traditional fuels. Unfortunately, this process is currently limited due to the low thermostability of endoglucanases, which currently only allow for temperatures below 60°C (Ando et al., 2002).

The aim of this project is to use molecular dynamics simulations along with network analysis methods to study structurally similar endoglucanases of varying thermostability and search for identifiable differences in their respective stabilizing mechanisms. Ideally, these findings may in turn be applied to the design of enhanced endoglucanases for industrial applications.

Methods

- Molecules were solvated in water boxes then ionized to neutrality in VMD, followed by stepwise minimization.
- Simulations were performed using periodic boundary conditions, with 100ns production runs at 25°C, 50°C, 75°C, 100°C and 125°C for each of the three molecules.
- Root-mean-square fluctuation analysis (RMSF) was performed for each production run to visualize deviations of each residue throughout the simulations.
- Protein Energy Networks (PENs) were constructed for each simulation, in which a network model was built using residues as nodes and the averaged total nonbonded interactions as weighted edges as described by Vijayabaskar and Vishveshwara (2010).
- Hubs, the highly connected nodes in a network (degree >3), were identified and plotted as a function of energy.
- Clusters, connected components in a network, were identified from each PEN and plotted as a function of energy.
- Community subgraphs were constructed from cliques (k = 3) and a transition profile was plotted as function of energy.

RMSF Plots

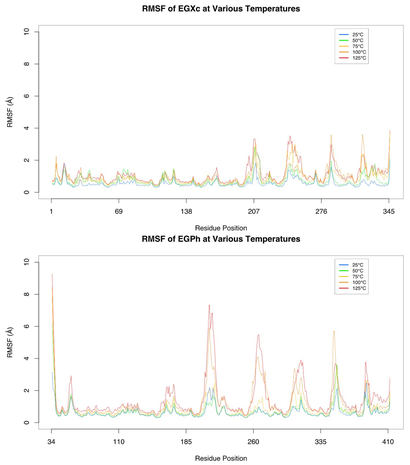


Figure 1: RMSF plots of endoglucanases from 25°C to 125°C.

These endoglucanases contain 8 looped regions, identified as the spiked regions in these plots. EGPh, a hyperthermophile (optimal temp ~97°C) has significantly greater fluctuations than the moderately thermophilic EGAc (optimal temp = 81°C) and the mesophilic EGXc (optimal temp = 45°C), which suggests a higher degree of flexibility within the looped regions.

PEN Hub Population

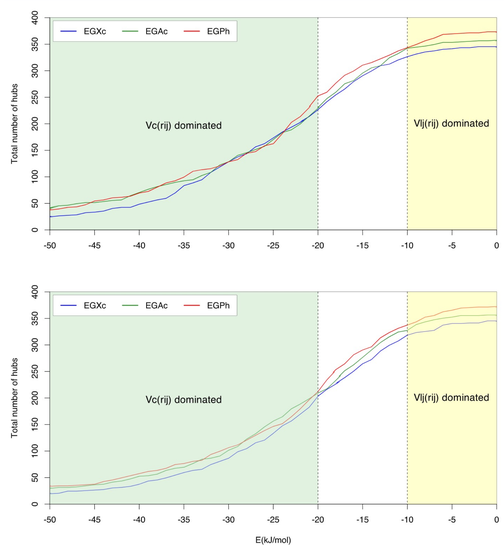


Figure 2: Hub population change in endoglucanases as a function of energy (E) at 25°C (top left), 75°C (top right) and 100°C (bottom left). Lennard-Jones interactions dominate the low-energy region (>-10kJ/mol) while Coulombic interactions dominate the higher energy region (<-20kJ/mol). The area between these regions is referred to as the transition region.

Hub population of EGPh is greater than its mesophilic and moderately thermophilic counterparts, both at the low energy and at the transition regions. This may suggest a more efficiently packed hydrophobic core.

PEN Cluster Population

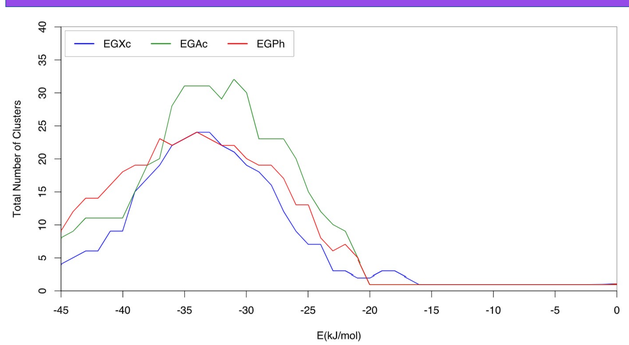


Figure 3: Cluster population change in endoglucanases with energy (E) at 25°C.

EGPh has more high-energy clusters than its counterparts and its peak is shifted towards the higher energy region, meaning its high-energy bonds are more segregated.

PEN Largest Community Transition Profile

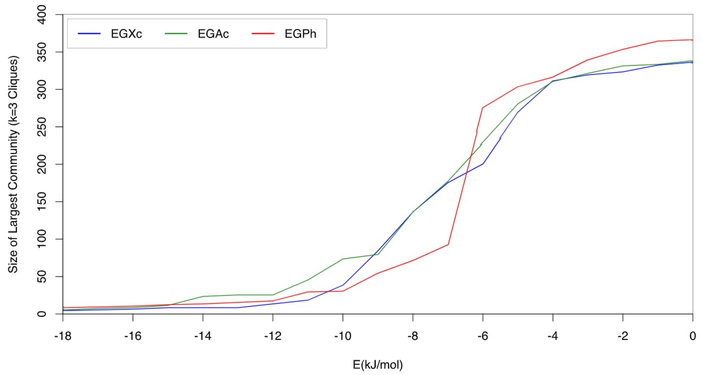


Figure 4: Largest community transition profile in endoglucanases.

EGPh displays larger community sizes in PENs at lower energy levels. Low energy communities help maintain a level of rigidity and increased global stability.

Results/Discussion

RMSF analysis (Figure 1) reveals large spikes throughout the looped regions of EGPh as temperature increases, which suggests a greater range of flexibility than its less thermostable counterparts. This is also supported by the observed cluster population changes in Figure 3, which revealed that EGPh's peak is shifted towards the high-energy region meaning a greater degree of segregation of high-energy bonds.

It should be further noted that EGPh does not have more overall high-energy bonds than EGAc, as shown in Figure 2. Rather, the largest community transition profile (Figure 4) reveals that EGPh has a larger community present throughout the lower energy region (>-6kJ/mol). These findings support the idea that EGPh's stability is derived from many low-energy hubs within the Lennard-Jones dominated region. This may suggest a better packed hydrophobic core, which would allow the enzyme to stay properly folded when exposed to high temperatures.

EGPh's clique population and largest community formation are at the low energy regime, which suggests this enzyme utilizes many low-energy communities to attain stability without too much rigidity. This, along with segregated electrostatic clusters, may allow adequate stability while allowing the flexibility to function properly at high temperatures.