Non-equilibrium Dynamics of Nanoscale Soft Matter Deformation

Austin D. Fergusson

Thesis submitted to the faculty of Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
In
Engineering Mechanics

Douglas P. Holmes
Anne E. Staples
Sunghwan Jung

July 24, 2014
Blacksburg, VA

Keywords: lipid bilayer, elastin, resilin, inverse temperature transition, umbrella sampling, potential of mean force

Non-equilibrium Dynamics of Nanoscale Soft Matter Deformation

Austin D. Fergusson

ABSTRACT

Life is soft. From the fluid-like structure of lipid bilayers to the flexible folding of proteins, the realm of nanoscale soft matter is a complex and vibrant area of research. The lure of personalized medicine, advanced sensing technology, and understanding life at a fundamental level pushes research forward. This work considers to areas: (1) lipid bilayer dynamics in the presence of substrate defects and (2) the inverse temperature transition of elastic proteins. Molecular dynamics simulations as well as umbrella sampling were employed. The behavior of the bilayers discussed in the work provides evidence that small defects on confining surfaces can promote nucleation of lipid tethers. Results the second part of this work indicate elastin-like peptides experiencing inverse temperature transitions may be capable of performing amounts of work similar to RNA polymerase; additionally, resilin's inverse temperature transition may be closely linked to the molecule's ability to efficiently transmit energy through the similar coil-β secondary structure transition seen in both cases. These insights into the inverse transition temperature are relevant for the design of bio-inspired sensors and energy storage devices.

Contents

Chapter 1| Supported Lipid Bilayer Compression Dynamics 1
1.1 Supported lipid bilayers (SLBs) and their significance in biology 1
1.2 Methods 2
1.3 Results and Discussion 5
Chapter 2| Protein Heat Engine: An exploitation of the inverse temperature transition of elastic proteins 9
2.1 Inverse Transition Temperature 9
2.2 Methods 9
2.2.1 The PACE Force Field 10
2.2.2 Protein Simulation Procedures 12
2.2.4 Umbrella Sampling 13
2.3 Results and Discussion 16
2.3.1 Inverse Temperature Transition 16
2.3.2 Umbrella Sampling 20
2.4 Conclusions 23
References 24

List of Figures

Figure 1. The simulated system contains a DPPC bilayer consisting of polar head groups (black) and hydrophobic tails (grey), a supporting substrate (red), and water layers above the membrane and below the substrate. 2
Figure 2. Ratio of the maximum and average lipid head group heights versus compressive area strain plots for the defect free case. 5
Figure 3. Views of the SLB along the simulation box diagonal. 6
Figure 4. Mean curvature (nm-1) versus compressive area strain plot of the SLB lower leaflet. 7
Figure 5. Schematic showing the work performed by an elastin peptide during the inverse temperature transition. 13
Figure 6. Schematic showing how frames were selected for umbrella sampling. All images are of the resilin exon III peptide. 14
Figure 7. Radius of gyration, Rg, values for the extended and folded states of the elastin peptide at various temperatures. 16
Figure 8. Radius of gyration vs. time plot for the elastin peptide at different temperatures. 16
Figure 9. Radius of gyration vs. time plot for the resilin exon III simulations. 19
Figure 10. (a) Free energy curve for the elastin peptide as calculated from umbrella sampling at 60°C. 22

List of Tables

Table 1. Steady state radius of gyration values for the $(GVGVP)_6$ peptide 18

Chapter 1| Supported Lipid Bilayer Compression Dynamics

1.1 Supported lipid bilayers (SLBs) and their significance in biology

Due to their ability to provide electrical insulation, incorporate receptor proteins, and suppress nonspecific ligand binding, supported lipid bilayers (SLBs) are used in biosensors.[1] Their mechanical properties have also been extensively investigated.[2, 3] Transmembrane phenomena, such as the ion channel function,[4] have been explored by suspending SLBs across nanopores.[5-7] The suspended SLB regions minimize substrate effects on system behavior[8] and provide an ideal environment for these investigations. Applications of SLBs commonly assume that the supporting substrate is continuous and defect-free.[9] However, this assumption is unrealistic *in vivo*. The lipid bilayer that forms the cell membrane is supported by an F-actin mesh that contains large gaps between its fibers, suggesting that a supporting scaffold model would be more appropriate.[10] While experiments characterized the influence of the cytoskeleton on static membrane properties[11] and prior computational works[12-14] by Xing *et al.* and Lin *et al.* investigated the changes in static membrane properties related to confinement by a supporting substrate, the effects of nanoscale defects within the substrate on membrane dynamics are still poorly understood.

We investigate these dynamics through molecular dynamics (MD) simulations of the phospholipid dipalmitoylphosphatidylcholine (DPPC). Two types of defects are simulated: (1) a system with a hole in the center of the substrate (hereafter referred to as the center-defect substrate) that is shown in Fig. 1b, and (2) a system with the substrate edges removed (referred to as the edge-defect substrate), shown in Fig. 1c. We also examine the effect of a third defect, namely, substrate protrusions, on SLB dynamics. These three scenarios are investigated by compressing the systems at a constant strain rate, which allows us to observe how the SLB accommodates the induced stress, both in the presence and absence of substrate defects.

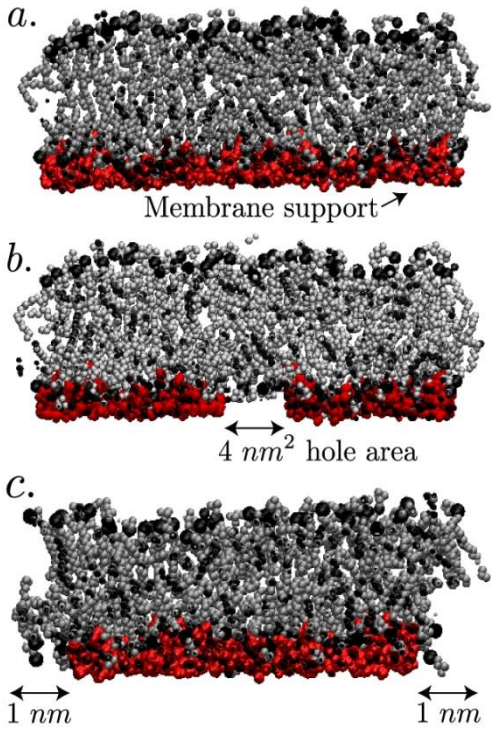

Figure 1. The simulated system contains a DPPC bilayer consisting of polar head groups (black) and hydrophobic tails (grey), a supporting substrate (red), and water layers above the membrane and below the substrate. The water layers present within the system are omitted from these images for clarity. The dimensions of this system are: 12.584 nm × 12.633 nm × 10.474 nm. (a) Side view of a slice of the supported lipid bilayer and the substrate. (b) A cross-sectional view of the 4 nm^2 hole present in the center of the substrate during the center defect simulations. (c) Cross-sectional view of the substrate showing the 1 nm of substrate that was removed around the edges for the edge defect simulation.

1.2 Methods

The simulations are conducted with GROMACS[15] to perform: (i) energy minimization, (ii) a canonical NVT (constant volume and constant temperature) ensemble simulation over 5 ns with the temperature controlled at 323 K by rescaling the molecular velocities every 1 picosecond (ps), and (iii) an NPT (constant pressure and temperature) ensemble based simulation over 13 ns using a Nose-Hoover thermostat[16, 17] with a coupling time of 1 ps, and a Parrinello-Rahman

barostat[18, 19] with a coupling time of 5 ps.[16-20] All simulations employ a time step of 0.002 ps. We use a simulation temperature and pressure of 323 K and 1 bar, respectively. The temperature is above the gel transition temperature of DPPC, and is commonly used in the literature.[21-26] Visualization of the structures is performed using VMD.[27]

The initial structure of DPPC is acquired from a previous study[28]: 128 lipid molecules arranged into two monolayers – referred to as leaflets – with 64 lipids each, as shown in Figure 1. We use a modified GROMACS 53a6 force field including Berger lipid parameters for our simulations.[28-30] The systems are simulated in bulk water (modeled using the SPC potential) with periodic boundary conditions.[31] The SETTLE algorithm is used to hold the geometry of each solvent water molecule fixed.[32] The geometry of each water molecule in the substrate is held fixed using the SHAKE algorithm.[33] Coulombic and van der Waals interactions have a cut-off distance of 1.2 nm.[31, 34] Long-range charge interactions are computed using a Particle Mesh Ewald (PME) summation scheme.[34] After equilibration, the area per lipid head group is 0.62 nm^2, which is in good agreement with previous work that reports this value as 0.629 ± 0.013 nm^2.[35] We construct a larger system by replicating the equilibrated structure along the x- and y-directions. The larger structure is desirable for compression simulations to ensure that the x- and y- dimensions of the lipid bilayer are several times larger than the substrate defects. This structure is subjected to another NPT equilibration that is simulated for 2 nanoseconds using the same thermostat and barostat as before. Once the density and temperature data converge, the system is ready for data production runs.

The supporting substrate consists of water molecules within ~2 nm from the bottom of the lipid bilayer. They are restrained from translational motion by applying an energy penalty to each atom. This penalty must be overcome to produce movement. We refer to the supporting substrate

as the “water slab”. The water slab restricts the ability of the bilayer to adjust during compression. Although the choice of support differs from typical experiments with SLBs, the simulations are still qualitatively comparable due to the polar nature of the support. To create the desired nanoscale defects in the substrate, the translational constraints on specific water molecules are removed, thereby allowing the bilayer to deform through these regions. The center-defect substrate has a 4 nm^2 hole in the center while the edge-defect substrate has 1 nm of the substrate removed along the outer edges.

Once this structure equilibrates, we conduct compression simulations. The pressure control during all compression simulations is coupled along the x- and y-directions. The corresponding control in the z-direction is decoupled to allow system deformations that are normal to the plane of the bilayer to differ from those in the tangential directions. The system is compressed along the plane of the membrane (i.e. in the x- and y-directions) at four deformation rates: 8×10^{-5}, 1×10^{-4}, 2×10^{-4}, and 3×10^{-4} nm/ps. As a control, we also simulate a defect-free substrate at compression rates of 8×10^{-5} and 1×10^{-4} nm/ps. Simulations are run until the bilayer reaches an area strain of ~30%. The simulations times for the different compression rates are (fastest – slowest): 10 ns, 15 ns, 30 ns, and 35 ns. Analysis becomes difficult thereafter due to the amount of bilayer deformation. We make the assumption that no lipid molecules switch leaflets because the flip-flop rate for DPPC is on the timescale of hours[36] which is orders of magnitude longer than the simulations performed in this work. Additionally, after reviewing the simulation trajectories, no lipid flip-flop was observed.

The mean curvature of both leaflets is calculated by applying a finite difference approximation of the Laplacian to the surface that is interpolated from the positions of the lipid head groups. We consider the compressive area strain given by

$$\varepsilon_A = \frac{A_0 - A_{new}}{A_0} \tag{1}$$

where A_0 represents the lateral area of the membrane in its uncompressed state and A_{new} is the same at a specific time step. Since this work only examines membrane compression, $\grave{o}_A$ is always positive.

1.3 Results and Discussion

In each of the defect-free simulations, the SLB remains planar until it experiences a critical area strain when the lower leaflet buckles. Buckling, which occurs at the corners of the simulation box, is considered to arise when a lipid is forced out-of-plane. † Compression causes the substrate molecules to pile up at the corners of the box. This protrusion promotes deformation. The compression dynamics of the SLB are biased toward delamination at the promoter site. We postulate that membrane delamination at these sites represents the initial stages of lipid tether formation.

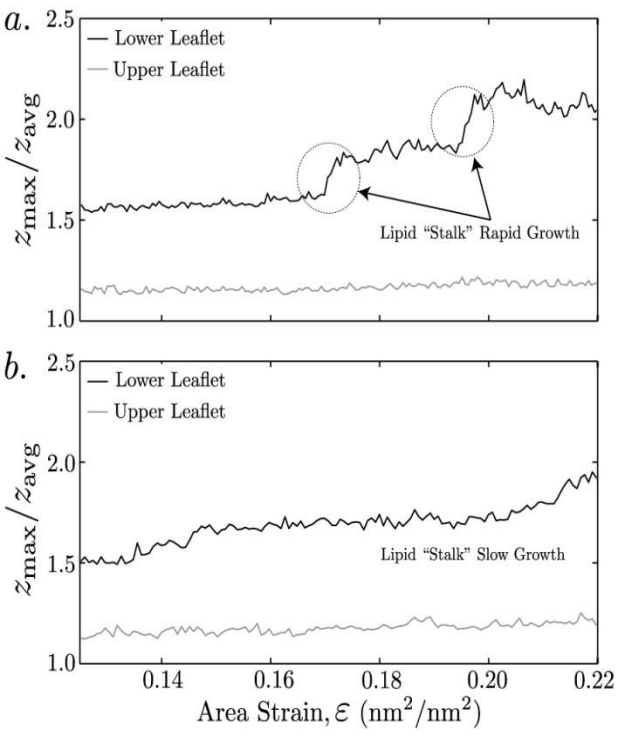

Figure 2. Ratio of the maximum and average lipid head group heights versus compressive area strain plots for the defect free case. (a) The 8 × 10-5 nm/ps compression rate plot. The graph displays distinct jumps representing the elongation of the lipid stalk in the lower leaflet. (b) The 1×10^{-4} nm/ps compression rate plot, which shows a smoother, continuous growth process of the lipid stalk in the lower leaflet. The initial elongation of the "stalk" begins with a lower compressive area strain at the faster compression rate.

Further compression drives additional lipid head groups in the lower leaflet upwards into the interior of the bilayer, leading to the growth of a lipid "stalk" in the buckling region. We monitor the ratio of the z-coordinate of the highest lipid head group to the mean z-coordinate of the head groups in a specific leaflet for different compression rates and present these results in Fig. 2. For a relatively slow compression rate (8×10^{-5} nm/ps), stalk growth follows a step-like process while for the higher compression rate (1×10^{-4} nm/ps) it is more continuous. The relationship between mean curvature and area strain is the same at both compression rates in the absence of substrate defects.

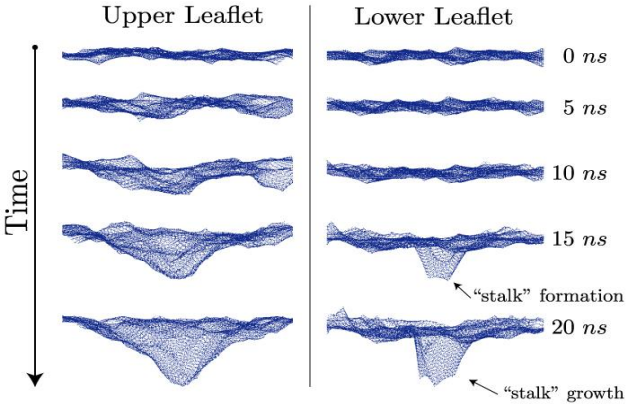

Figure 3. Views of the SLB along the simulation box diagonal. The upper (left) and lower (right) leaflet surfaces are shown every 5 nanoseconds. The images shown represent a section of each leaflet surface that spans the simulation box in order to observe the deformation near the center-defect in greater detail.

A relationship between the volume enclosed by a lipid protrusion and the compressive area strain of the bilayer has been recently proposed.[37], Employing this relationship, the lipid stalks considered here will likely continue to elongate into tubes or tethers since the volume enclosed by the protrusions is minimal[37] while the area strain is high. The negligible enclosed volume is an artifact that arises from the choice to use a substrate constructed from water molecules that are restrained from translation.

Figure 3 shows the case where a 4 nm^2 hole in the supporting substrate is responsible for the primary mode of deformation. Snapshots of the upper and lower leaflets from the slowest compression rate simulation are displayed. When the lipid bilayer lies on a substrate that contains the defect, the primary response to compression is extrusion of the lower leaflet through

the center defect. Simultaneously, the upper leaflet begins to curve toward the center of the defect. The extrusion process produces a more rapid increase in curvature with increasing area strain than is displayed for the defect-free system. The mode of deformation transitions from one that is center-defect controlled to being substrate protrusion controlled with increasing compression rate. After this initial extrusion, buckling occurs at the corners of the simulation box in a manner similar to that for a defect-free substrate. However, the compression rate has a significant influence on the deformation mode that the SLB assumes.

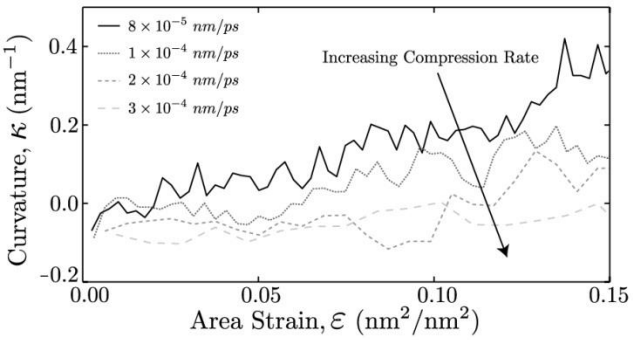

Figure 4. Mean curvature (nm-1) versus compressive area strain plot of the SLB lower leaflet. The data presented here have been time averaged every 200 ps for the simulations containing a 4 nm2 hole in the center of the supporting substrate.

At the highest compression rate, the substrate protrusions control the deformation mode in a manner that is indistinguishable from the behavior of a bilayer compressed on a substrate without a center defect. Two possible explanations for this behaviour are: (1) the compression rate is faster than the rate the lipid membrane can adjust to the strain, or (2) The substrate deformation at the corners of the simulation box is large enough to bias the system response toward those defects instead of the hole in the support. The mean curvature of the lower leaflet is plotted against membrane area strain for different compression rates in Fig. 4.

Next, we consider the edge-defect scenario. The topology of the edge-defect substrate is analogous to supporting the bilayer with evenly spaced pillars. The pillars sit on a surface that undergoes equal bending about both horizontal axes. As bending increases, the separation distance between the tops of the supporting pillars decreases. Eventually these pillar tops join one another, recreating a continuous substrate. The membrane compression causes lipid

molecules to extrude through portions of the open space surrounding the substrate. The bilayer deforms preferentially through the spaces at the corners of the substrate.

During compressive biaxial loading, nanoscale substrate defects produce nontrivial changes in SLB deformation compared to a defect-free substrate. Such nanoscale defects could be present even in carefully fabricated substrates. This suggests that a defect-free substrate model may be invalid even for microscale SLBs. Finally, small defects within a supporting substrate are capable of inducing lipid tether nucleation.

Chapter 2| Protein Heat Engine: An exploitation of the inverse temperature transition of elastic proteins

2.1 Inverse Transition Temperature

The family of elastic proteins exhibits a conformational transition above certain temperatures[38-45]. This behavior, which occurs above a protein's inverse transition temperature (ITT), represents the direct conversion of heat into mechanical work. When harnessed properly, the inverse temperature transition could produce useful mechanical work[46-53]. This inverse temperature transition is so named due to the increased order exhibited by the proteins following the conformational change, which would seem to violate the second law of thermodynamics[39, 54]. The second law is not in fact broken during the transformation because the decrease in the protein's entropy is counterbalanced by a greater increase in entropy of the highly ordered water that was previously associated with the protein's structure[39, 54-56]; therefore, the net change in entropy of the system is positive. The event that triggers this change is named hydrophobic collapse[54]. During hydrophobic collapse, the hydrophobic amino acid residues that were previously exposed to solvent molecules associate to reduce their solvent accessible surface area. This work will focus on the transitions of elastin and resilin whose ITT's are 20°C - 40°C[39, 56-58] and 50°C[59], respectively.

2.2 Methods

Highly parallelized molecular dynamics (MD) simulations were run to examine the conformational transition pathways exhibited by both elastin $(GVGVP)_6$ and resilin exon III $(GYSGGRPGGQDLG)_5$ during their respective inverse temperature transitions. The initial protein structures were generated using AmberTools[60] by stringing together the desired amino

acid chains. C- (COO^-) and N-terminal (NH_3^+) groups were added to the structure during simulation preprocessing in GROMACS 3.3.4[61]. Due to the long relaxation times usually exhibited by proteins, a model combining both efficiency and accuracy was needed to study this problem. Coarse grained MD simulations are orders of magnitude faster than their all-atom counterparts, but efficiency comes at a price. The reduction in resolution would likely miss important information regarding the transition, or potentially be incapable of modeling the transition at all. Alternate modeling options were needed. The solution to these modeling issues was found in the form of the PACE[62, 63] force field recently developed by Wu *et al.*

2.2.1 The PACE Force Field

The PACE force field constitutes a hybrid coarse-grained/united-atom model. The bulk of the particles in many biological simulations are water molecules. This water uses valuable computational resources but provides little additional information as an all-atom model. Thinking along these lines, the developers of PACE utilize coarse-grained water in their model while keeping proteins represented by a united-atom formulation. The proposed method allows for greater detail of protein dynamics while reducing the necessary computational time. This model represents a compromise between efficiency and accuracy. The potential function for this force field is:

$$
\begin{aligned}
V = & \sum_b \frac{1}{2} K_{bond,b}(r_b - r_{0,b})^2 + \sum_a \frac{1}{2} K_{angle,a}(\theta_a - \theta_{0,a})^2 + \sum_d K_{dih,d}[1+\cos(n_d\zeta_d - \zeta_{0.d})] \\
& + \sum_{1-4\,pair} 4\varepsilon_{14,ij}\left(\frac{\sigma_{14,ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{14,ij}^{6}}{r_{ij}^{6}}\right) + \sum_i \frac{1}{2} K_{imp,i}(\xi_i - \xi_{0,i})^2 + \sum_d \sum_{n_d=1}^{N_d} K_{dih_{n_d},d}[1+\cos(n_d\zeta_d - \zeta_{0_{n_d},d})] \\
& + \sum_{short} 4\varepsilon_{short,ij}\left(\frac{\sigma_{short,ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{short,ij}^{6}}{r_{ij}^{6}}\right) + \sum_{i\neq j} 4\varepsilon_{CGW-CGW,ij}\left(\frac{\sigma_{CGW-CGW,ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{CGW-CGW,ij}^{6}}{r_{ij}^{6}}\right) \\
& + \sum_{i\neq j} 4\varepsilon_{CGW-UA,ij}\left(\frac{\sigma_{CGW-UA,ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{CGW-UA,ij}^{6}}{r_{ij}^{6}}\right) + \sum_{i\neq j} 4\varepsilon_{vdW,ij}\left(\frac{\sigma_{vdW,ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{vdW,ij}^{6}}{r_{ij}^{6}}\right) \\
& + \sum_{|i-j|>2} 4\varepsilon_{attr,O-N}\left(\frac{\sigma_{attr,O-N}^{12}}{r_{O-N,ij}^{12}} - \frac{\sigma_{attr,O-N}^{6}}{r_{O-N,ij}^{6}}\right) + \sum_{|i-j|>2}\left(\frac{C_{rep,O-C_\alpha}}{r_{O-C_\alpha,ij}^{12}} + \frac{C_{rep,O-C}}{r_{O-C,ij}^{12}} + \frac{C_{rep,C-N}}{r_{C-N,ij}^{12}}\right) + \sum_{i\neq j} \frac{q_i q_j}{4\pi\varepsilon_0\varepsilon_r r_{ij}}
\end{aligned}
\tag{2}
$$

Where $\varepsilon_{a\text{-}b,ij}$ and $\sigma_{a\text{-}b,ij}$ are LJ parameters, the $K_{xxx,x}$ terms are spring constants, and the C terms are parameters relating to the charged interactions between particles[64-66]. PACE has produced reasonable results for the secondary structure of the AK17[65], Fs[65], and GB1p[65] peptides when compared to experimental data. Comparison of protein folding kinetics with the literature can be misleading because the PACE model increases the speed of protein folding by up to 5-10 times the rates[65] seen in experiments. This acceleration is beneficial for examining previously unobservable slow protein kinetics.

PACE was implemented in GROMACS 3.3.4[61]. This older version of GROMACS was used to construct the topology files (molecular level definition of the system) needed to perform the protein simulations. All simulations were performed on GROMACS 4.5.x[15] or later. There is one exception to the previous statement; all umbrella sampling simulations were run using GROMACS 3.3.4 due to the large changes in how the software package accepts umbrella sampling parameters between the 3.x and 4.x versions.

2.2.2 Protein Simulation Procedures

Every simulation requires the definition of a control volume which will be referred to as a "simulation box" for the remainder of this work. The system of interest, in this case a peptide, was centered inside the box and solvent molecules were added until there was no room left. The contents of the control volume will most likely be at artificially high energy states due to close proximity of the newly added molecules. To prevent system instability, the two step process of energy minimization followed by equilibration was performed twice. The peptides were restrained during the first stage to allow the solvent molecules to adjust around the peptide structure without displacing the peptide. No restraints were applied during the second stage. Energy minimization was performed using the steepest descent algorithm[15]. The algorithm follows the steepest descending energy gradient by an amount, δ, at each step. If the new energy is lower than the previous value, the step is successful. If the new energy is larger than the previous value, the step is considered unsuccessful and the algorithm returns to the previous value and repeats this process using a smaller step of ~0.2δ. Each equilibration step was performed using the NPT ensemble, i.e. the number of particles, the temperature, and the pressure were all held constant. The temperature and pressure of the system were maintained at 150 K and 1 atm, respectively, using the Berendsen temperature and pressure coupling algorithms[67]. Additional NPT equilibration was performed at each temperature for 5 ns prior to beginning the data collection simulations. Elastin (resilin) was simulated for 1.5 μs (500 ns) with a time step of 4 fs. Hydrogen atoms were represented by virtual sites to facilitate the use of the larger time step.

2.2.4 Umbrella Sampling

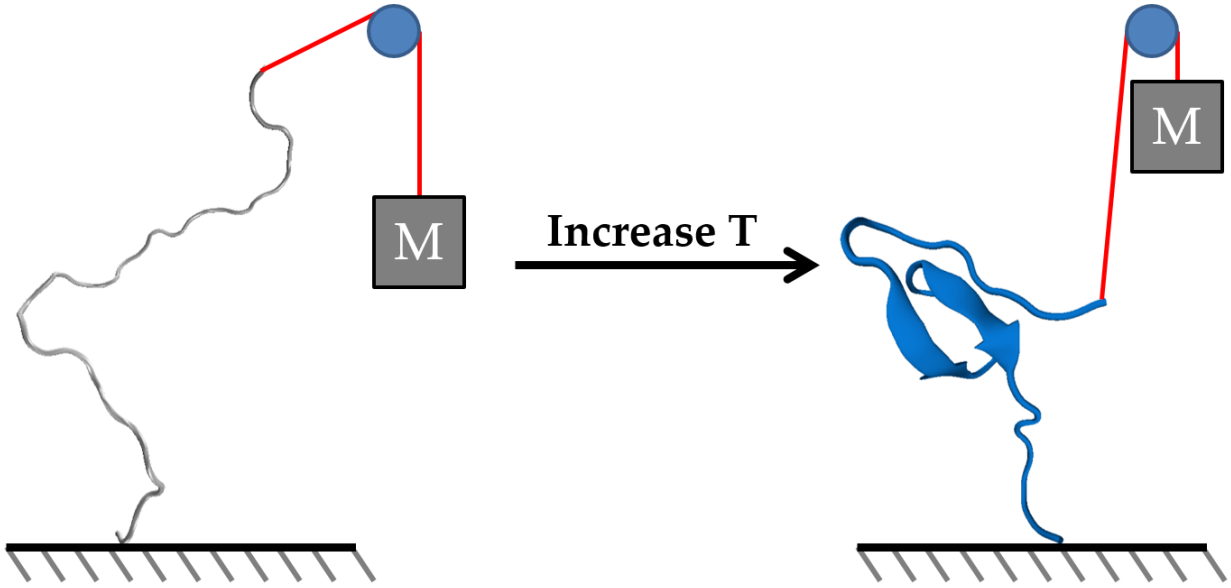

Figure 5. Schematic showing the work performed by an elastin peptide during the inverse temperature transition.

The energy near a subset of the transition trajectory is required to adequately calculate the average force along that path and by extension the approximate work performed across the inverse temperature transition. There are several techniques available to calculate free energies between two states. The free energy difference, ΔF, between these states is related to the work, W, performed to move from one to the other by the non-equilibrium work relation developed by Jarzynski[68].

$$\left\langle \exp(-\beta W) \right\rangle = \exp(-\beta \Delta F) \tag{3}$$

Therefore, the average work performed during a non-equilibrium process to move a system from state A to state B is greater than or equal to the free energy difference between the states. Concisely, this relation links equilibrium properties to non-equilibrium processes. In the case of a set of finite-time simulations, the calculated value of ΔF serves as the upper limit of the work

the peptide could perform during the inverse temperature transition[68]. In this work, I apply the umbrella sampling method to derive the free energy surface.

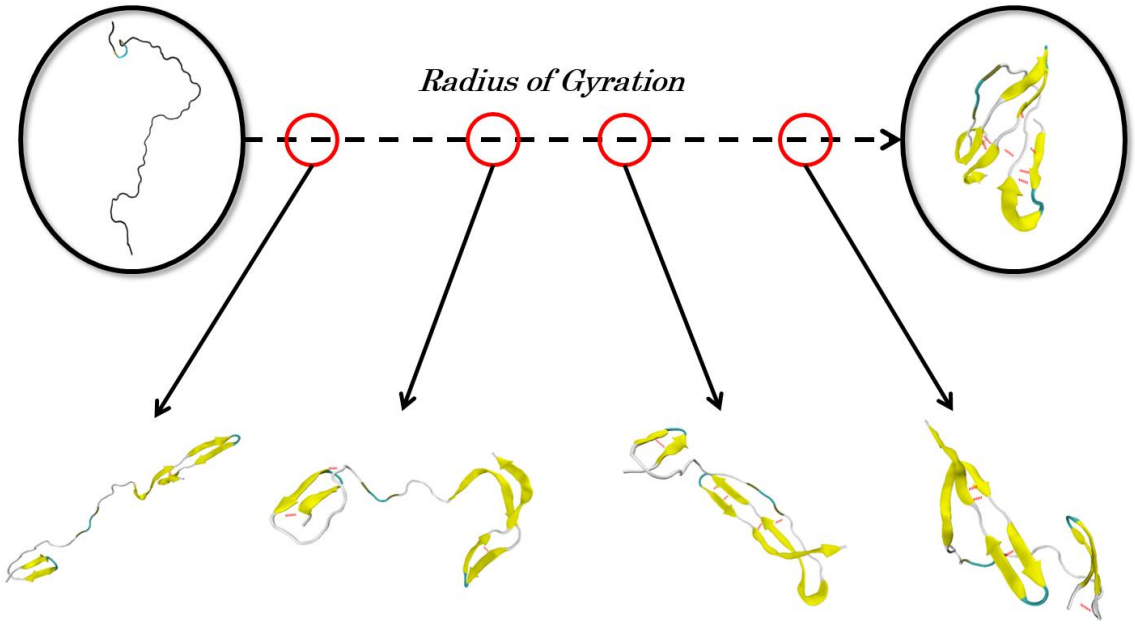

Figure 6. Schematic showing how frames were selected for umbrella sampling. All images are of the resilin exon III peptide.

Umbrella sampling was performed to quantify the maximum possible work a model peptide could perform when moving between the extended and contracted states. Specifically this method calculates the potential of mean force (PMF), or free energy, of the peptide along the conformal change pathway between an initial state A and the final state B (figure 6). In order to adequately sample the energy surface between the two configurations, additional simulations were run while holding the reaction coordinate, ξ, constant. Sampling windows were chosen from the original simulation trajectory such that the values of the reaction coordinate change by a constant value between neighboring windows. The radius of gyration was chosen as the reaction coordinate for resilin and elastin because both of their inverse temperature transitions move the peptides from an elongated, disordered state to a compact, ordered configuration. During an

umbrella sampling simulation, the model protein's radius of gyration is restrained to its initial value for that sampling window with a harmonic (spring) potential with the spring constant k = 1000 kJ nm^{-2} mol^{-1}. Through this restraint, the model protein is only allowed to sample its configuration space near the initial reaction coordinate value. Sufficient sampling is achieved when the individual energy histograms display nearly Gaussian distributions and overlap with neighboring histograms in regions of high sampling. Larger values of k produce faster convergence of the energy histograms but reduce sampling of the reaction coordinate space; therefore, a larger number of sampling windows would be required to accurately calculate the free energy curve. Insufficient overlap of the histograms leads to artificial peaks or oscillations in the free energy landscape.

Each umbrella sampling window was simulated for a total of 51 ns and 21 ns for elastin and resilin, respectively. The first nanosecond of all sampling simulations was a short NPT equilibration simulation and was excluded from the analysis. The separation between neighboring sampling windows was approximately 0.025 nm. The PLUMED[69] plug-in was used with GROMACS to perform umbrella sampling using the protein's radius of gyration as a non-standard reaction coordinate. Weighted Histogram Analysis (WHAM) was performed on the umbrella simulation data sets using the native GROMACS utility "g_wham[70]". The g_wham utility calculates the autocorrelation times for each sampling window and performs bootstrap error analysis to assess the statistical uncertainty of the free energy profile.

2.3 Results and Discussion

2.3.1 Inverse Temperature Transition

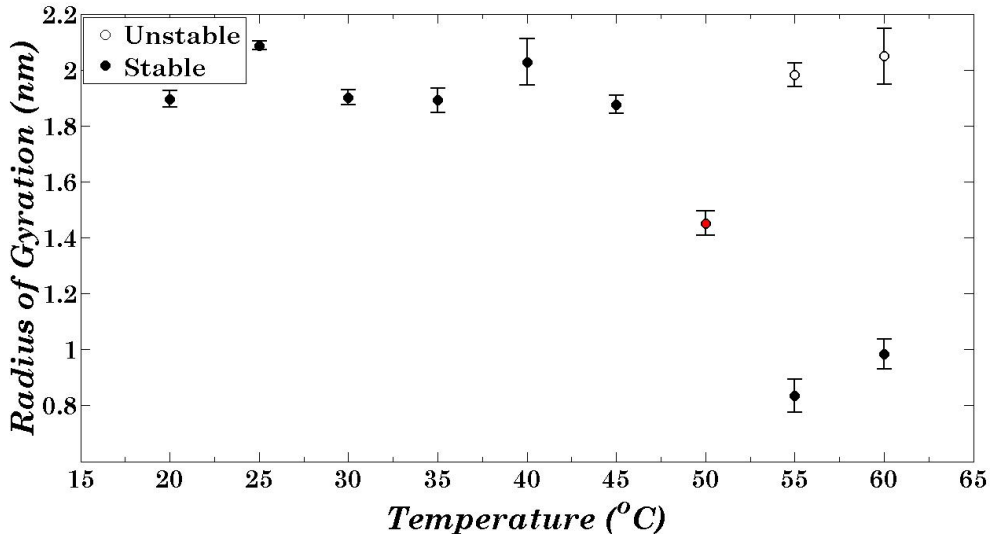

Figure 7. Radius of gyration, Rg, values for the extended and folded states of the elastin peptide at various temperatures. Filled circles mark the stable states while empty circles mark unsteady states. The lone circle filled with red (50°C) represents the first temperature where the inverse temperature transition occurs. Since the peptide was still in the process of transitioning at 1.5 μs, the 50°C point was calculated from the final 400 ns of the simulation trajectory, which is representative of the final configuration the peptide had paused at.

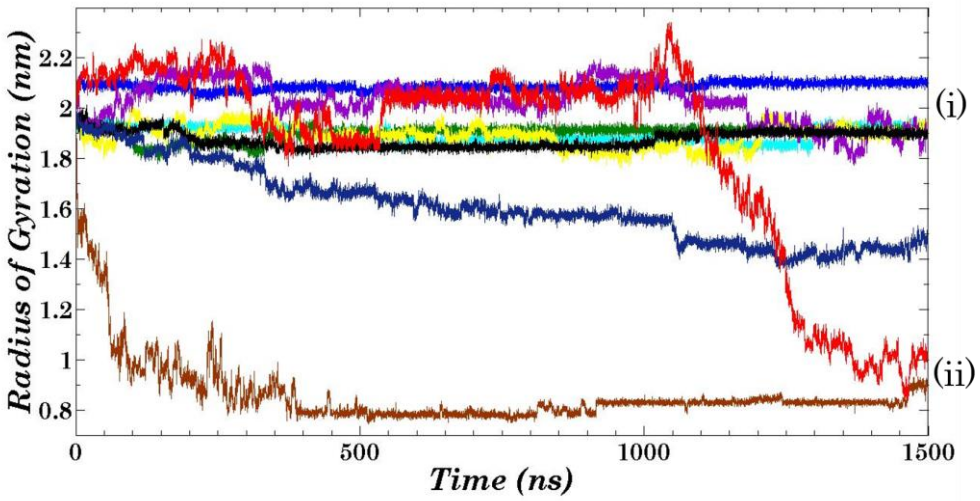

Figure 8. Radius of gyration vs. time plot for the elastin peptide at different temperatures. There are two main configurations: (i) elongated with R_g ~ 2nm and (ii) folded with R_g ~0.95nm. The peptides from the three highest temperature simulations experienced inverse temperature transitions. (light blue) 20°C, (blue) 25°C, (green) 30°C, (yellow) 35°C, (purple) 40°C, (black) 45°C, (dark blue) 50°C, (brown) 55°C, (red) 60°C

There are many different values for the ITT of elastin listed in the literature. These values range from 20°C - 40°C[39, 56-58]. The range of reported values suggests a difference between the measured temperature of a solution (*i.e.* average thermal energy) and the actual temperature of the peptides or proteins of interest. In this work, elastin was simulated at 9 different temperatures on the range 293 K – 333 K (20°C - 60°C) in 5 K increments (figure 8). A transition from the extended (disordered) state to the folded (ordered) state was observed for all $T \geq 323$ K (50°C); therefore, the ITT of elastin is bracketed between 318 K (45°C) and 323 K (50°C). According to these results, the molecular temperature at which elastin transitions is higher than previously reported values. I propose both the previously reported values and those presented here are compatible since the temperature measured in prior experimental works is representative of the average thermal energy of all of the molecules in the experimental systems. This allows for instances where individual components of those systems deviate from the average. Following this logic, it is not difficult to imagine a situation where molecular collisions impart enough energy to a peptide to raise its molecular "temperature" above its transition temperature while the mean temperature of the system remains below the ITT. Therefore, the previously documented experimental ITT values can instead be thought of as the temperatures where the probability of any one peptide transitioning is high enough that observation is possible.

The observed transition of the elastin peptide yields a final configuration containing β-structure (turns/strands) and a reduction in random coil structure. In two out of the three cases where the elastin peptide transitioned, it completed the change of configuration in ~150-200 ns, but the transition in the T = 50°C case occurred over a timespan >1500ns. The order-of-magnitude difference in the observed transition rates is curious, but comparison of these values with the

Table 1. Steady state radius of gyration values for the $(GVGVP)_6$ peptide

T (°C)	R_g (nm)	T (°C)	R_g (nm)
20	1.8952 ± 0.0294	40	2.0276 ± 0.0834
25	2.0868 ± 0.0161	45	1.8756 ± 0.0327
30	1.9018 ± 0.0274	55	0.8344 ± 0.0594
35	1.8911 ± 0.0428	60	0.9844 ± 0.0536

literature is difficult due to the mixed nature of the PACE model. The transition rate difference may be a transient effect due to 50°C being near the peptide's transition temperature.

Figure 7 displays the observed states, both steady and unsteady, of the elastin peptide for the full temperature range (20°C - 60°C). Since the peptide at 50°C had not finished transitioning by 1.5 μs, the average R_g value was taken from the final 400 ns of the simulation. The final 400 ns show a pause of the transition at a value of R_g of ~1.4 nm. As I discuss later in this work, a radius of gyration value of 1.4 nm places the elastin peptide within one of the shallower energy basins in its free energy surface; therefore, it is not unreasonable to see the transition slow or pause at this point. The radius of gyration values for each point in figure 7 are provided in table 1. The values across the transition appear reasonable when compared with MD simulations of the elastin-like $(GPGSQ)_5$ peptide of *A. aurantia* silk[40]. The change in the silk peptide's radius of gyration across its transition is ~52% of the pre-transition value. The elastin peptide's radius of gyration change is ~48-49% of its pre-transition value.

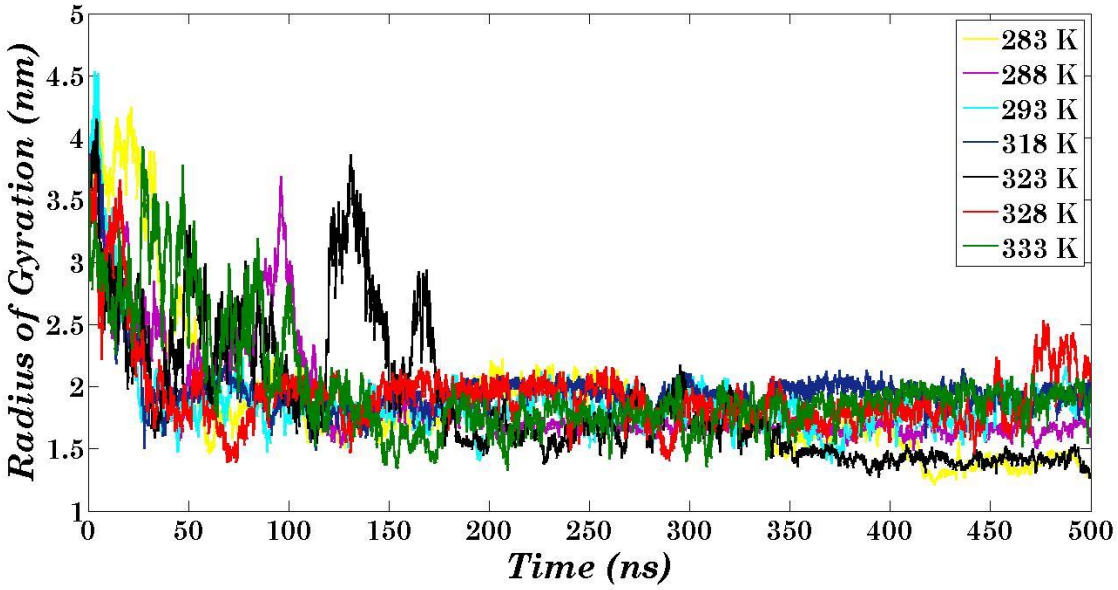

Figure 9. Radius of gyration vs. time plot for the resilin exon III simulations. The strange peaks in the 288 K (purple) and 323 K (black) simulation data are the result of the peptide rebounding back to an elongated structure.

During the transition, the resilin peptide repeatedly folded back on itself to create a repeating β-strand pattern. The peptide forms multiple β-turns and β-strands by the end of the inverse temperature transition, which is similar to the structure observed by Qin *et al.*[41] Their experiments show resilin having an ITT of 50°C, which contrasts with the behavior observed in this work. Every resilin simulation (10°C - 60°C) displayed a coil-β transition. The resilin exon III peptide discussed in this work has a much lower ITT than the whole resilin protein. As discussed by Urry[54, 55], the increase in a peptide's hydrophobicity serves to reduce the ITT, so the whole resilin protein is proportionally less hydrophobic than the exon III peptide studied here. Interestingly, as can be seen in figure 9, the 50°C and 15°C simulations both show a period where the peptide collapses ($R_g \approx 1.6$ nm) followed by the rapid increase in radius of gyrations ($R_g \approx 3.2$ nm). During this time, the ends of the peptide fold over the middle and move close together before quickly rebounding.

2.3.2 Umbrella Sampling

All of the elastin peptide's umbrella sampling windows were taken from the 60°C simulation. The potential of mean force (PMF) profile can be seen in figure 10. The profile displays two flat regions near the largest reaction coordinate values. Both of the plateaus correspond with values for the radius of gyration that the elastin peptide maintains at temperatures below the ITT; additionally, the flat regions provide a plausible explanation for why the 60°C simulation required ~1 μs of simulation time before it began to transition while the peptide from the 55°C simulation transitioned within the first 100 ns. An elastin peptide could remain in either of these metastable states while the temperature is above the ITT because the neighboring energy landscape does not provide any configurations that are more energetically favorable. The inverse temperature transition should occur once thermal vibrations push the peptide to a configuration having a radius of gyration of ~1.8 nm. There are several small basins present in the PMF profile where the peptide could pause in its transition. The simulation data supports the existence of these basins in the free energy profile. Multiple simulations display more gradual change for radius of gyration values near 1.6 nm and 1.44 nm; however, the thermal fluctuations are strong enough to prevent the peptide transition from getting stuck in these regions. The lowest temperature transition data (50°C) shows the elastin peptide spending prolonged periods (>300 ns) in both energy basins. In fact, the peptide had yet to move out of the 1.44 nm radius of gyration energy basin as of 1.5 μs. The free energy difference between the elongated state and the final folded configuration is ~39.8 kJ mol^{-1}. As a comparison, the work performed by RNA polymerase during one stroke is ~25 kJ mol^{-1}.[71] If even 50% of the free energy change from the transition can be harnessed, a single "stroke" of the elastin peptide transition produces ~20 kJ mol^{-1} of work, which is reasonably close to the RNA polymerase value. Contraction is the

primary useful motion of the elastin peptide transition. It is not unreasonable to imagine a modified version of the peptide[54] used as part of a sensor when coupled with a mechano-sensitive channel.

Resilin exon III umbrella sampling was performed on windows from the 50°C simulation. The resulting PMF profile (figure 10) shows a steep gradient leading from the largest radius of gyration values to ~2.5 nm where a plateau region is located. One large energy basin exists in this profile with several smaller basins within it. The change in free energy due to folding is ~21.5 kJ mol^{-1}. Experimental evidence suggests resilin is primarily used as the protein equivalent to the capacitor[41, 72]. In insects, prior to jumping, energy is transmitted down the resilin molecule and stored in a resilin-chitinous cuticle composite structure before being rapidly released with little energy loss due to resilin's high elasticity[73]. I mention this because the elastic portions of the resilin molecule (exon I and exon III) display transitions to β-structure during this energy input/release process[41], which is reminiscent of the conformation change observed during the inverse temperature transition. While the energy input is high enough through the action of temperature or stress, the elastic portions of resilin acquire and maintain the β-turn/strand structure; but, if the input is removed, the protein regains its random coil configuration and transmits energy through the molecule. Thus, the β-structure is a key component for the successful transmission and storage of energy in naturally occurring resilin.

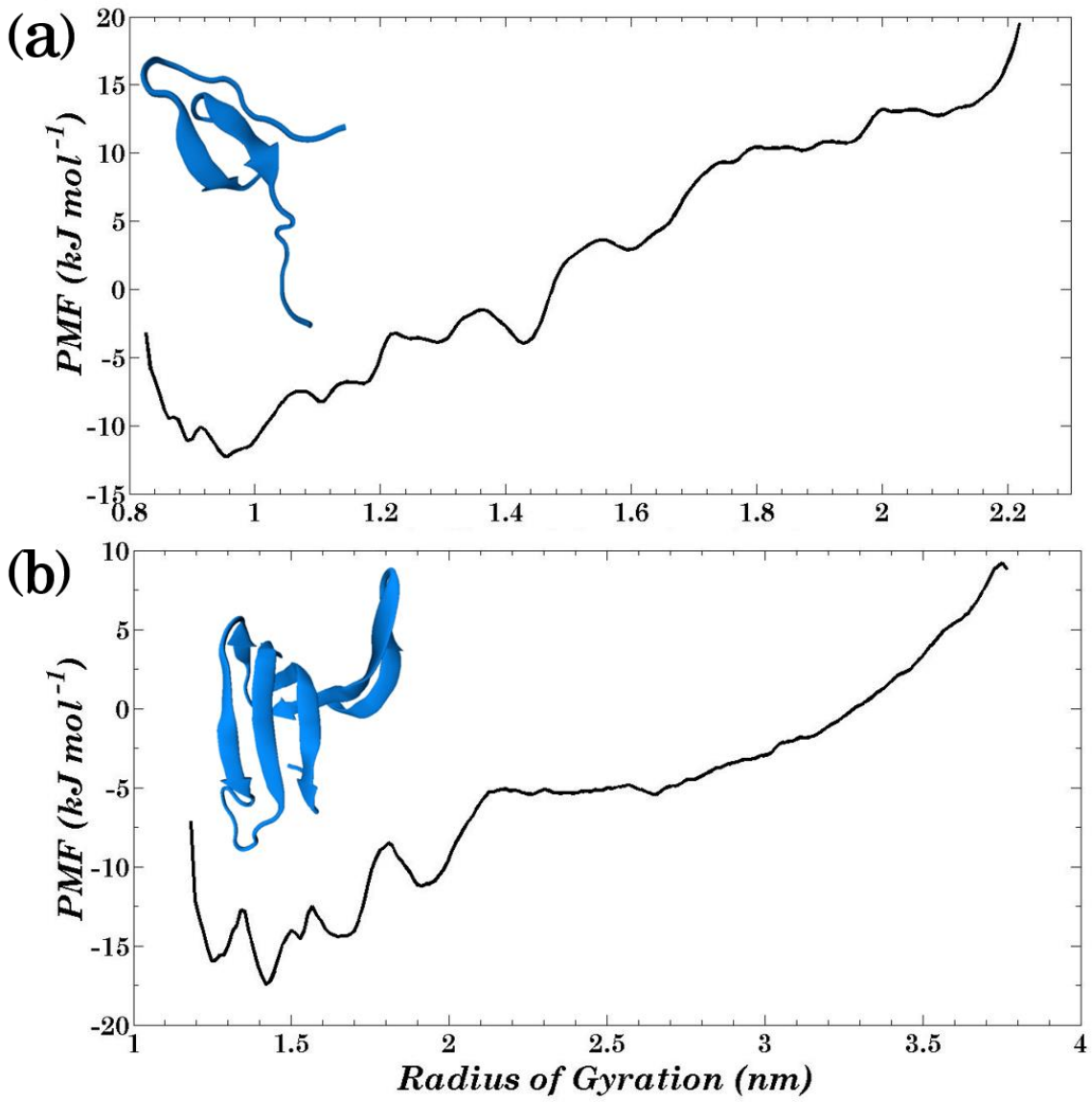

Figure 10. (a) Free energy curve for the elastin peptide as calculated from umbrella sampling at 60°C. The peptide configuration found in the lowest energy basin is shown in blue. The initially random coil structure transformed into β-structure (strands and turns) following the inverse temperature transition. The free energy difference between the initial and final configurations is ~39.8 kJ mol^{-1}. The statistical uncertainty of the energy minimum is 4.43 kJ mol^{-1}. (b) Free energy curve for the resilin exon III peptide as calculated by umbrella sampling at 50°C. The final configuration of the peptide is shown in blue. Following transition, the peptide takes on a configuration primarily consisting of β-strands and β-turns. The free energy difference between the elongated state and the folded configuration is ~21.5 kJ mol^{-1}. The statistical uncertainty at the energy minimum is 4.73 kJ mol^{-1}.

2.4 Conclusions

This work has addressed the inverse temperature transitions of an elastin peptide and a resilin exon III peptide. Both peptides acquire β secondary structure following transition, which is consistent with previous MD studies and experimental works[39, 41, 58, 72-74]. The free energy differences between the elongated and folded states of each peptide were computed with umbrella sampling along the radius of gyration reaction coordinate. The elastin peptide's inverse temperature transition produces a change in free energy, which, if utilized properly, could produce similar values of work as naturally occurring RNA polymerase[71]. Therefore, I propose that elastin or more generally elastin-like peptides can be integrated into nanoscale sensors when combined with mechano-sensitive channels to detect a variety of stimuli and convert that information into mechanical work. The results of this work when examined in context with prior experimental studies of resilin suggest that the coil-β secondary structure transition is a key component of resilin's use in nature as an efficient energy transfer and storage device. To address this connection, simulations of the full resilin molecule under biologically relevant stresses should be performed to observe the energy transmission and storage pathways. The benefits of greater computational efficiency coupled with faster protein folding kinetics make the PACE model an ideal candidate for future modeling work on this system. Additional simulation work should be performed to better understand the behavior and utility of the elastic peptides presented in this work when coupled to simple mechanical systems.

References

1. Z. Wu, B. Wang, Z. Cheng, X. Yang, S. Dong and E. Wang, *Biosensors and Bioelectronics*, 2001, **16**, 47-52.
2. M. Staykova, D. P. Holmes, C. Read and H. A. Stone, *Proc Natl Acad Sci*, 2011, **108**, 9084-9088.
3. E. Sackmann, *Science*, 1996, **271**, 43-48.
4. E. Reimhult and K. Kumar, *Trends Biotechnol.*, 2008, **26**, 82-89.
5. S. Kresak, T. Hianik and R. L. C. Naumann, *Soft Matter*, 2009, 4021-4032.
6. B. A. Nellis, J. H. S. Jr. and S. H. Risbud, *Acta Biomaterialia*, 2011, 380-386.
7. A. Simon, A. Girard-Egrot, F. Sauter, C. Pudda, N. P. D'Hahan and L. Blum, *Journal of Colloid and Interface Science*, 2007, 337-343.
8. E. Reimhult and K. Kumar, *Cell Press*, 2008.
9. S. Varma, M. Teng and H. L. Scott, *Langmuir*, 2012, **28**, 2842-2848.
10. K. H. S. Richard J. Barfoot, Benjamin R. G. Johnson, John Colyer, Robert E. Miles, Lars J. D. Jeuken, Richard J. Bushby, and Stephen D. Evans, *Langmuir*, 2008, **24**, 6827-6836.
11. J. X. Feng Yi, A. Michelle Smith, Atul N. Parikh, David A. LaVan, *Soft Matter*, 2009, **5**, 5037-5041.
12. C. Xing and R. Faller, *J. Phys. Chem. B*, 2008, **112**, 7086-7094.
13. C. Xing, O. H. S. Ollila, I. Vattulainen and R. Faller, *Soft Matter*, 2009, **5**, 3258-3261.
14. X. Lin, C. Wang, M. Wang, K. Fang and N. Gu, *J. Phys. Chem. C*, 2012, **116**, 17960-17968.
15. B. Hess, *J. Chem. Theory Comput.*, 2008, **4**, 435-447.
16. S. Nose, *Molecular Physics*, 1983, **100**, 191-198.
17. W. G. Hoover, *Phys. Rev. A*, 1985, **31**, 1695-1697.
18. M. Parrinello and A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182-7190.
19. M. L. Klein and S. Nose, *Molecular Physics*, 1983, **50**, 1055-1076.
20. G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**.
21. J. F. Nagle, *Biophys. Journal*, 1993, **64**, 1476-1481.
22. H. Heller, M. Schaefer and K. Schulten, *J. Phys. Chem.*, 1993, **97**, 8343-8360.
23. S. Leekumjorn and A. K. Sum, *Biochimica et Biophysica Acta*, 2007, 354-365.
24. J. F. Nagle and M. C. Wiener, *Biochimica et Biophysica Acta*, 1988, **942**, 1-10.
25. J. F. Nagle and S. Tristram-Nagle, *Biochimica et Biophysica Acta*, 2000, **1469**, 159-195.
26. M. Kodama, M. Kuwabara and S. Seki, *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1982, **689**, 567-570.
27. W. Humphrey, A. Dalke and K. Schulten, *Journal of molecular graphics*, 1996, **14**, 33-38.
28. D. P. Tieleman and H. J. C. Berendsen, *J. Chem. Phys.*, 1996, **105**, 4871-4880.
29. D. P. Tieleman, S. J. Marrink and H. J. C. Berendsen, *Biochimica et Biophysica Acta*, 1997, **1331**, 235-270.
30. S. J. Marrink, O. Berger, D. P. Tieleman and F. Jahnig, *Biophysical Journal*, 1998, **74**, 931-943.
31. H. Berendsen, J. Postma, W. Van Gunsteren and J. Hermans, Intermolecular forces: Proceedings of the Fourteenth Jerusalem Symposium on Quantum Chemistry and Biochemistry held in Jerusalem, Israel, 1981.
32. S. Miyamoto and P. A. Kollman, *J. Comp. Chem.*, 1992, **13**, 952-962.
33. J. P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comp. Phys.*, 1977, **23**, 327-341.
34. T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089-10092.
35. J. F. Nagle, R. Zhang, S. Tristram-Nagle, W. Sun, H. I. Petrache and R. M. Suter, *Biophysical Journal*, 1996, **70**, 1419-1431.
36. N. Sapay, W. F. D. Bennett and D. P. Tieleman, *Soft Matter*, 2009, **5**, 3295-3302.
37. M. Staykova, M. Arroyo, M. Rahimi and H. A. Stone, *Physical Review Letters*, 2013, **110**.
38. J. R. McDaniel, D. C. Radford and A. Chilkoti, *Biomacromolecules*, 2013, **14**, 2866-2872.
39. D. W. Urry, *Jounal of Protein Chemistry*, 1988, **7**.
40. O. T. Ukpebor, A. Shah, E. Bazov and G. S. Boutis, *Soft Matter*, 2014, **10**.
41. G. Qin, X. Hu, P. Cebe and D. L. Kaplan, *Nat Commun.*, 2012, **3**.
42. A. Tarakanova and M. J. Buehler, *Modelling Simul. Mater. Sci. Eng.*, 2013, **21**.
43. A. S. Tatham and P. R. Shewry, *Phil. Trans. R. Soc. Lond. B*, 2002, **357**.
44. A. S. Tatham and P. R. Shewry, *TIBS*, 2000, **25**.
45. G. Gronau, S. T. Krishnaji, M. E. Kinahan, T. Giesa, J. Y. Wong, D. L. Kaplan and M. J. Buehler, *Biomaterials*, 2012, **33**, 8240-8255.
46. R. E. Sallach, M. Wei, N. Biswas, V. P. Conticello, S. Lecommandoux, R. A. Dluhy and E. L. Chaikof, *J. AM. CHEM. SOC.*, 2006, **128**, 12014-12019.
47. M. Kim, S. Tang and B. D. Olsen, *Polymer Physics*, 2013, **51**, 587-601.
48. J. C. Johnson and L. T. J. Korley, *Soft Matter*, 2012, **8**.
49. M. Swierczewska, C. S. Hajicharalambous, A. V. Janorkar, Z. Megeed, M. L. Yarmush and P. Rajagopalan, *Acta Biomaterialia*, 2008, **4**, 827-837.
50. H. Lee, H. R. Kim, R. G. Larson and J. C. Park, *Macromolecules*, 2012, **45**, 7304-7312.
51. D. F. Lyons, V. Le, I. Gene L. Bidwell, W. H. Kramer, E. A. Lewis, D. Raucher and J. J. Correia, *Biophys. Journal*, 2013, **104**, 2009-2021.
52. S. R. MacEwan and A. Chilkoti, *Peptide Science*, 2010, **94**.
53. A. Ghoorchian, J. T. Cole and N. B. Holland, *Macromolecules*, 2010, **43**, 4340-4345.
54. D. W. Urry, *Biopolymers*, 1998, **47**, 167-178.
55. D. W. Urry, *Chemical Physics Letters*, 2004, **399**, 177-183.
56. D. W. Urry, in *Application of Thermodynamics to Biological and Materials Science*, ed. P. M. Tadashi, InTech, 2011.
57. R. Glaves, M. Baer, E. Schreiner, R. Stoll and D. Marx, *ChemPhysChem*, 2008, **9**, 2759-2765.
58. C. Nicolini, R. Ravindra, B. Ludolph and R. Winter, *Biophys. Journal*, 2004, **86**, 1385-1392.

59. D. H. Ardell and S. O. Andersen, *Insect Biochemistry and Molecular Biology*, 2001, **31**, 965-970.
60. D. A. Case, I. Thomas E. Cheatham, T. Darden, H. Gohlke, R. Luo, J. Kenneth M. Merz, A. Onufriev, C. Simmerling, B. Wang and R. J. Woods, *J. Comp. Chem.*, 2005, **26**.
61. E. Lindahl, B. Hess and D. v. d. Spoel, *Journal of Molecular Modeling*, 2001, **7**, 306-317.
62. W. Han and K. Schulten, *J. Chem. Theory Comput.*, 2012.
63. W. Han, C.-K. Wan, F. Jiang and Y.-D. Wu, *J. Chem. Theory Comput.*, 2012, **6**.
64. W. Han, C.-K. Wan and Y.-D. Wu, *J. Chem. Theory Comput.*, 2010, **6**, 3390-3402.
65. W. Han and K. Schulten, *J. Chem. Theory Comput.*, 2012, **8**, 4413-4424.
66. S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman and A. H. d. Vries, *J. Phys. Chem.*, 2007, **111**, 7812-7824.
67. H. J. C. Berendsen, J. Postma, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684-3690.
68. C. Jarzynski, *Physical Review Letters*, 1997, **78**, 2690-2693.
69. M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia and M. Parrinello, *Computer Physics Communications*, 2009, **180**, 1961-1972.
70. J. S. Hub, B. L. d. Groot and D. v. d. Spoel, *J. Chem. Theory Comput.*, 2010, **6**, 3713-3720.
71. R. M. Krupka, *Experimental Physiology*, 1998, **83**, 243-351.
72. R. S.-C. Su, Y. Kim and J. C. Liu, *Acta Biomaterialia*, 2014, **10**, 1601-1611.
73. M. Burrows, S. R. Shaw and G. P. Sutton, *BMC Biology*, 2008, **6**.
74. A. Amadei, A. B. M. Linssen and H. J. C. Berendsen, *Structure, Function, and Genetics*, 1993, **17**, 412-425.