

# Chem 4050 Project 2 Report

Charlie Li

November 30, 2025

## 1 Introduction

Polymer materials are widely used in various applications due to versatility in their structures and properties. In particular, they serve such indispensable functions in the exploration of space as thermal insulation, radiation shielding, adhesives, and more. However, under extremely low temperatures in space, polymers are prone to transition from an unfolded to a folded state, which compromises their mechanical properties and can lead to severe consequences to the spacecraft. Therefore, it is imperative that we design those polymers to remain unfolded under low temperatures. In this study, we explore the folding of polymers under various temperatures and interactions *in silico* using molecular dynamics (MD). We found that the folding behavior is strongly influenced by the harmonic bond strength and chain length, but not by the repulsive potential. This provides valuable insight into the behaviors of phase transition of polymers, which could guide the future design and development of novel polymer materials for use in space.

## 2 Methodology

### 2.1 Simulation Schemes

Coarse-grain molecular dynamics (MD) of a linear homopolymer is performed in a cubic unit cell. The system is treated as a canonical ensemble with constant particle number, volume, and temperature.

To explore the phase transition of the polymer chain under different temperatures, simulations are performed using different spring constants  $k$  and repulsive well depths  $\epsilon_{repulsive}$ , and phase transition is observed at temperatures ranging from 0.1 to 1. A chain length of  $N = 20$  beads is used. Three duplicates are studied for each parameter set to take into account the effect of the random initialization of the particle positions.

To explore the effects of chain length on phase transitions, simulations are performed with  $N = 5, 10, 20, 40, 80$  with the same  $k = 1.0$  and  $\epsilon_{repulsive} = 1.24$ . The unit cell dimensions are chosen to be five times the chain length ( $a = 25, 50, 100, 200, 400$  for each  $N$ , respectively). Three duplicates are studied for  $N = 5, 10, 20$ , but only one duplicate is available for each of the  $N = 40$  and  $80$  groups due to the limit of resources on the personal laptop.

### 2.2 Polymer Model

The polymer is modeled with the bead-spring model as a linear chain of homogeneous beads, with each bead representing one monomer.

### 2.3 Interactions

Three types of interactions are included in our model, which could be classified as either bonded or non-bonded. The forces are derived by taking the negative gradient of their respective potential energies. Each

particle pair feels one and only one interaction. The bonded potential governs the particle pairs exactly one bond apart, the repulsive potential governs those exactly two bonds apart, and the attractive potential governs those further than two bonds apart. The total potential energy of the system is calculated as the sum of the potential energies between each particle pair.

- **Bonded Interaction:** the harmonic potential is used to model the bond energies between monomer.

$$U_{bond}(r) = \frac{1}{2}k(|\vec{r}| - r_0)^2 \quad (1)$$

$$F_{bond}(r) = -k(|\vec{r}| - r_0) \frac{\vec{r}}{|\vec{r}|} \quad (2)$$

Bonded interaction potential energy and forces, where  $k$  is the spring constant and varies in our study,  $\vec{r}$  is the displacement vector between particles, and  $r_0 = 1$  is the equilibrium bond length.

- **Non-Bonded Interaction**

- **Attraction:** The attractive interaction is modeled with the full Lennard-Jones potential.

$$U_{attractive} = 4\epsilon_{attractive} \left[ \left( \frac{\sigma}{|\vec{r}|} \right)^{12} - \left( \frac{\sigma}{|\vec{r}|} \right)^6 \right] \quad (3)$$

$$F_{attractive} = -\frac{24\epsilon_{attractive}}{|\vec{r}|} \left[ \left( \frac{\sigma}{|\vec{r}|} \right)^6 - 2 \left( \frac{\sigma}{|\vec{r}|} \right)^{12} \right] \cdot \frac{\vec{r}}{|\vec{r}|} \quad (4)$$

Attractive L-J potential energy and force, where  $\epsilon_{attractive} = 0.5$  is the energy well depth,  $\sigma = 1$  is the distance at which the potential energy is 0, and  $\vec{r}$  is the displacement vector between particles.

- **Repulsion:** The repulsive interaction is modeled with a truncated and shifted Lennard-Jones potential, which exhibits net repulsion when the distance is smaller than the equilibrium distance and a net 0 when the distance is larger.

$$U_{repulsive} = \begin{cases} 4\epsilon_{repulsive} \left[ \left( \frac{\sigma}{|\vec{r}|} \right)^{12} - \left( \frac{\sigma}{|\vec{r}|} \right)^6 + \frac{1}{4} \right] & |\vec{r}| < 2^{\frac{1}{6}}\sigma \\ 0 & |\vec{r}| \geq 2^{\frac{1}{6}}\sigma \end{cases} \quad (5)$$

$$F_{repulsive} = \begin{cases} -\frac{24\epsilon_{repulsive}}{|\vec{r}|} \left[ \left( \frac{\sigma}{|\vec{r}|} \right)^6 - 2 \left( \frac{\sigma}{|\vec{r}|} \right)^{12} \right] \cdot \frac{\vec{r}}{|\vec{r}|} & |\vec{r}| < 2^{\frac{1}{6}}\sigma \\ 0 & |\vec{r}| \geq 2^{\frac{1}{6}}\sigma \end{cases} \quad (6)$$

Repulsive L-J potential energy and force, where  $\epsilon_{repulsive}$  is the energy well depth and varies in our study,  $\sigma = 1$  is the distance at which the potential energy before shifting is 0, and  $\vec{r}$  is the displacement vector between particles.

## 2.4 Propagation

The positions and velocities of the particles are propagated using the **Velocity Verlet Algorithm**, which updates the positions by constant acceleration approximation and updates the velocity using the average accelerations before and after the position update.

$$\vec{r}(t + \Delta t) = \vec{r}(t) + \vec{v}(t)\Delta t + \frac{1}{2}\vec{a}\Delta t^2 \quad (7)$$

Position update in a small timestep is approximated by constant acceleration.

$$\vec{v}(t + \Delta t) = \vec{v}(t) + \frac{\vec{a}(t) + \vec{a}(t + \Delta t)}{2} \quad (8)$$

Velocity update in a small timestep is approximated using the average acceleration.

Acceleration on each atom is calculated using Newton's second law, where  $\vec{F}_{ji}$  represents the force on atom  $i$  by atom  $j$ .

$$\vec{a}_i = \frac{\sum_{j \neq i} \vec{F}_{ji}}{m_i} \quad (9)$$

## 2.5 Temperature Control

The velocity rescaling thermostat is used to keep the system at constant temperature. Temperature is computed from the equipartition theorem, where  $KE$  is the total kinetic energy of the system,  $N$  is the number of particles, and  $K_B$  is Boltzmann's constant.

$$T = \frac{2KE}{3NK_B} \quad (10)$$

Equipartition Theorem

$$\vec{v}' = \sqrt{\frac{T_{target}}{T}} \cdot \vec{v} \quad (11)$$

Velocity Rescaling Equation

## 2.6 System Preparations

The simulation system is prepared for production by undergoing the following steps in order: initialization of particle positions, damped MD, energy minimization (EM), heating, and equilibration.

1. **Position Initialization:** Particle positions are first initialized starting at the box center and randomly growing the linear polymer at equilibrium bond length.
2. **Damped MD:** Damped MD is performed after position initialization to avoid overlap of the particles. In this step, particle velocities start from 0, and their propagation magnitude is reduced by a coefficient  $\gamma$ , which is set to 0.0001 in our study. A maximum of 100 time steps is allowed in this step, unless the potential energy difference of the system drops under some convergence threshold (set to  $10^{-2}$  in our study), indicating a local minimum is reached.

$$\vec{v}(t + \Delta t) = \vec{v}(t) + \gamma \cdot \frac{\vec{a}(t) + \vec{a}(t + \Delta t)}{2} \quad (12)$$

In damped MD, velocity propagation is damped by a small coefficient  $\gamma$

3. **Energy Minimization:** Next, EM is performed using the `scipy.optimize.minimize` module. EM finishes when either the total potential energy converges on the threshold of  $10^{-3}$ , the norm of all forces converges on the threshold of  $10^{-6}$ , or a maximum step of 200 is reached. EM often reduces the total potential energy significant and helps to put the system into an optimal position.
4. **Heating:** After EM, the system is heated up to the target temperature by rescaling its velocities at each step. Heating is performed in our study by linearly increasing the temperature in 500 consecutive steps. A gradual heating process gives the right kinetic energy to the system without greatly disrupting the particle configurations.

**5. Equilibration:** Finally, equilibration is performed. Velocities are first re-initialized at the target temperature by drawing from the Maxwell-Boltzmann distribution. 1000 equilibration steps are then performed with the exact same procedure as in the production steps, but the trajectory is not recorded. Equilibration allows the system to undertake its favored configuration at the target temperature.

$$f(v_x) = \sqrt{\frac{m}{2\pi K_B T}} \cdot \exp\left(-\frac{mv_x^2}{2K_B T}\right) \quad (13)$$

The Maxwell-Boltzmann velocity distribution for a single dimension, where  $m$  is mass,  $K_B$  is Boltzmann's constant, and  $T$  is the temperature.

**6. Production:** Production steps are performed using the velocity verlet algorithm with velocity rescaling every 100 steps.

## 2.7 Output Materials and Integration

For every simulation, a topology file (.psf) and a trajectory file (.xyz) are produced. The topology file records the bonding information of the molecule, while the trajectory file records the positions of the particles. The trajectory is saved every 10 steps during production. The results are readily integrated into the MDAnalysis python package, which is a powerful molecular dynamics analysis tool. The molecular configurations can be visualized with the `show_mdanalysis` function of nglview once the trajectories are loaded into the Universe.

## 2.8 Analysis

### 2.8.1 Analysis preparations

Analysis is performed by reading coordinates from the trajectory files on-the-fly, which allows for concurrent analysis of the large number of trajectories with limited computer memory. The molecule is first unwrapped onto the infinite lattice space by iteratively placing all neighboring atoms onto their closest image to the central atom using depth-first search. The unwrapped molecule is then translated to the center of the box and re-wrapped. If the molecule is ill-extended and cannot be contained inside the box, its center of geometry will fail to match the box center once the re-wrapping is performed. Using this criterion, along with the metric that the max length of the molecule should not exceed half the length of the box diagonal, trajectories that violate the minimum image convention are picked out and excluded from further analysis.

All metrics presented are calculated on the centered molecule and are averaged across the last 500 trajectory frames.

### 2.8.2 Analysis Metrics

Analysis is performed by calculating three metrics: radius of gyration ( $R_g$ ), end-to-end distance ( $R_e$ ), and potential energy. Phase transition is identified by analyzing the trend of those metrics under varying temperatures.

- **Radius of gyration:** radius of gyration is the root mean squared deviation of each atom from the center of mass ( $\vec{r}_{CM}$ ) of the polymer.

$$R_g = \sqrt{\frac{1}{N} \sum_{i=1}^N |\vec{r}_i - \vec{r}_{CM}|^2} \quad (14)$$

- **Normalized radius of gyration:** the relative or normalized radius of gyration is defined as the radius of gyration divided by the theoretical maximum  $R_g$  of the linear polymer chain, assuming a bond length of  $r_0$ . This allows the comparison of  $R_g$  between polymers of different lengths. See equation 22 for derivation.

$$\text{Normalized } R_g = \frac{R_g}{\frac{r_0}{2\sqrt{3}} \sqrt{N^2 - 1}} \quad (15)$$

- **End-to-end distance:** end to end distance is the distance between the two end particles of the polymer.

## 2.9 A note on units

In this study, all simulation parameters are in reduced units of Lennard-Jones, where  $r^* = \frac{r}{\sigma}$  and  $E^* = \frac{E}{\epsilon}$  (reduced unit denoted by  $*$ ,  $\sigma$  is the distance at which the L-J potential energy is 0, and  $\epsilon$  is the L-J well depth). Mass  $m^*$  is set to 1 in units of single particle mass, and the Boltzmann's constant  $K_B$  is set to 1.

## 3 Results

Simulation is first performed to find optimal spring constants  $k$  and repulsive well depths  $\epsilon_{repulsive}$  that keeps the polymer unfolded under low temperatures. All simulations use a chain of 20 particles in a unit cell of side length 100. The fixed parameters in these simulations are as follows, all in reduced units:  $r_0^* = 1.0$ ,  $\epsilon_{attractive}^* = 0.5$ ,  $\sigma^* = 1$ ,  $m^* = 1$ ,  $\Delta t^* = 0.01$ . Velocity is rescaled every 100 frames, and trajectory is saved every 10 frames. The final metrics are averaged over the last 50 trajectory frames, and aggregated over the 3 duplicates.

Figure 1 shows the change of radius of gyration ( $R_g^*$ ) with temperature under different spring constants and repulsive well depths. Figure 2 shows the change of end-to-end distance ( $R_e^*$ ) with temperature, and figure 3 shows the change of the potential energy of the system with temperature. Some example configurations of the polymer chains are shown in figures 4 and 5. Note that the configurations do not reside in a constantly sized unit cell; instead, they are enlarged around the molecule. In contrast, The particle spheres are constantly sized with regard to the size of the unit cells. Therefore, larger spheres represent closer enlargement, and therefore a smaller total volume occupied by the molecule, and vice versa.

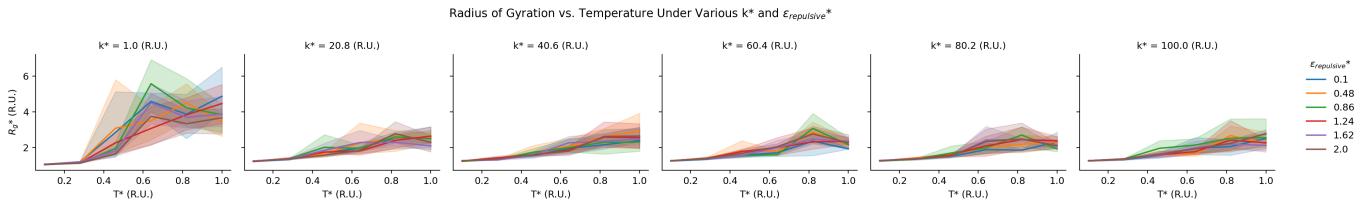


Figure 1: Radius of gyration plotted against temperature. Columns are grouped by spring constants. Lines colored by  $\epsilon_{repulsive}^*$ . Error bands show 95% confidence intervals for 3 duplicates. R.U. denote reduced unit.

Figures 6 and 7 show different metrics plotted against temperature for polymers of different lengths. Note that the  $N = 40$  and  $N = 80$  (where  $N$  denotes chain length) groups have only one duplicate per group due to limits of computing power. Some instances of the  $N = 40$  and  $N = 80$  simulations have produced invalid results that violate the assumptions of the minimum image convention, and are therefore excluded from the analysis. Normalization of  $R_g^*$  is performed by dividing it by the theoretical maximum  $R_g^*$  derived

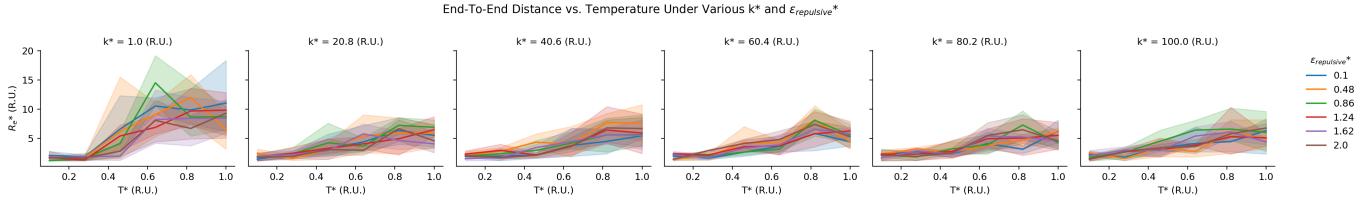


Figure 2: End-to-end distance plotted against temperature. Columns are grouped by spring constants. Lines colored by  $\epsilon_{repulsive}^*$ . Error bands show 95% confidence intervals for 3 duplicates. R.U. denote reduced unit.

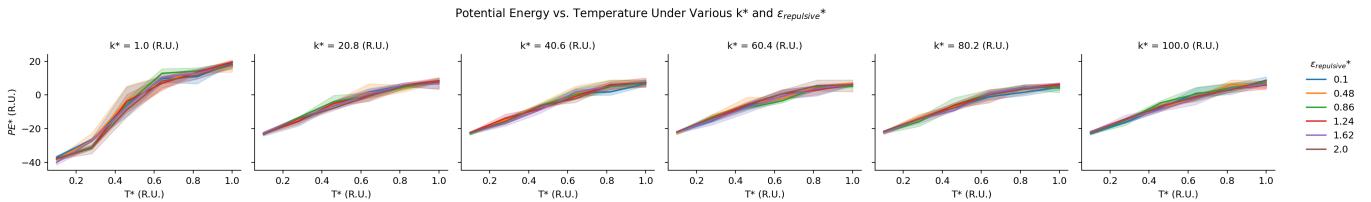


Figure 3: Potential energy of the system plotted against temperature. Columns are grouped by spring constants. Lines colored by  $\epsilon_{repulsive}^*$ . Error bands show 95% confidence intervals for 3 duplicates. R.U. denote reduced unit.

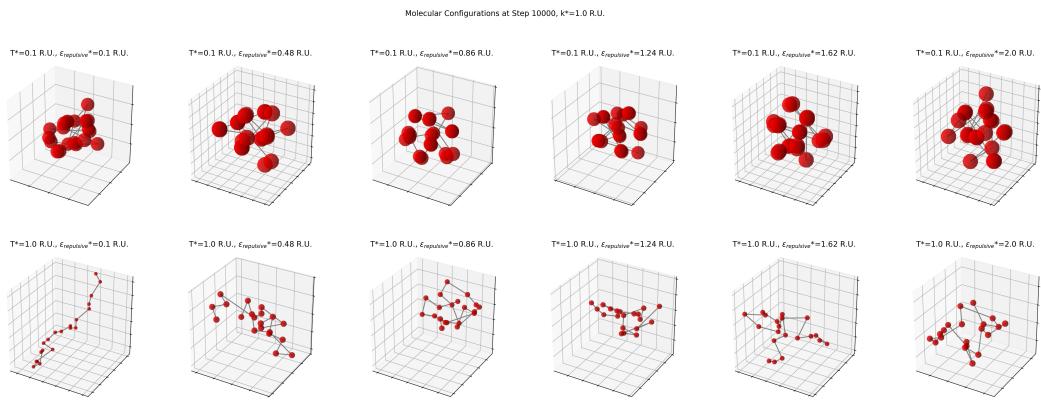


Figure 4: Visualizations of the polymer configuration at the last trajectory frame with  $k^* = 1.0$  and varying  $\epsilon_{repulsive}^*$ . Rows grouped by temperature; columns grouped by  $\epsilon_{repulsive}^*$ . Plots are enlarged around the polymer; the size of the particles is constant with regard to unit cell length. R.U. denote reduced unit.

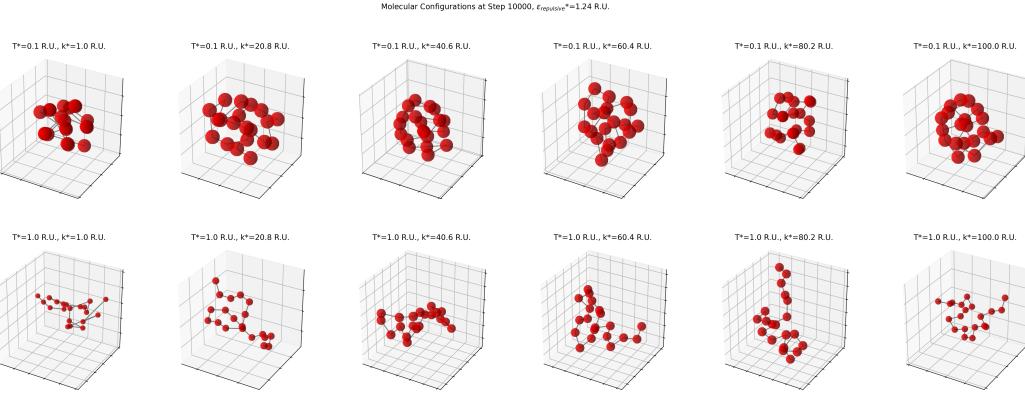


Figure 5: Visualizations of the polymer configuration at the last trajectory frame with  $\epsilon_{\text{repulsive}}^* = 1.24$  and varying  $k^*$ . Rows grouped by temperature; columns grouped by  $k^*$ . Plots are enlarged around the polymer; the size of the particles is constant with regard to unit cell length. R.U. denote reduced unit.

in equation 22. Some example configurations of the polymer chains of different lengths are shown in figures 8 and 9. The same notes apply.

An animation of the polymer chain with  $N = 40$  beads at  $T^* = 1.0$  and  $k^* = 100.0$  is included in the project folder as a .gif file.

## 4 Discussion

### 4.1 Justification of the parameter range

In this study, we explored a range of spring constants varying from 1.0 to 100.0. By dimensional analysis,  $[E^*] = [k^*][x^*]^2$  in reduced units, where  $[x]$  denotes the unit of quantity  $x$ . So  $\epsilon = [k^*]\sigma^2$ , which leads to  $[k^*] = \frac{\epsilon}{\sigma^2}$ , and  $k = \frac{\epsilon}{\sigma^2}k^*$ . A typical spring constant for a covalent bond is on the order of  $10^2$  to  $10^3 N/m$ , typical  $\sigma$  falls on the order of  $10^0$  to  $10^1 \text{ \AA}$ , and typical  $\epsilon$  on the order of  $10^{-3}$  to  $10^{-1} eV$ . Therefore, a typical spring constant in reduced units for a covalent bond can be on the order of  $10^2$  to  $10^7$ . However, as will be discussed, no significant behavioral difference is shown between the groups where  $k$  is greater than 20. Therefore, such a sampling range is well justified.

We begin by analyzing the effects of  $k^*$  and  $\epsilon_{\text{repulsive}}^*$  on the phase transition temperature, using results from the  $N = 20$  simulations.

### 4.2 Effect of Spring Constant on Phase Transition

- **Radius of Gyration:** A greater  $R_g^*$  indicates unfolding, while a smaller  $R_g^*$  is typical for globular, folded polymers. Using equation 22, the theoretical maximum  $R_g^*$  for a polymer chain of 20 beads is expected to be  $\frac{r_0^*}{2\sqrt{3}}\sqrt{20^2 - 1} = 5.77r_0^*$ . Figure 1 shows that the magnitude of the spring constant significantly affects the phase transition of the polymer chain. When  $k^* = 1.0$ , phase transition begins as early as  $T^* = 0.3$ , where  $R_g^*$  begins to increase, and the polymer completely transitions to the unfolded state at around  $T^* = 0.6$ .  $R_g^*$  reaches 4 to 5 when  $T^* > 0.8$  for all the systems, which is quite close to the theoretic maximum. This phase transition is clearly visualized in figure 4, where each column is a

### Different Metrics vs. Temperature For Different Chain Lengths At $k^* = 1.0$

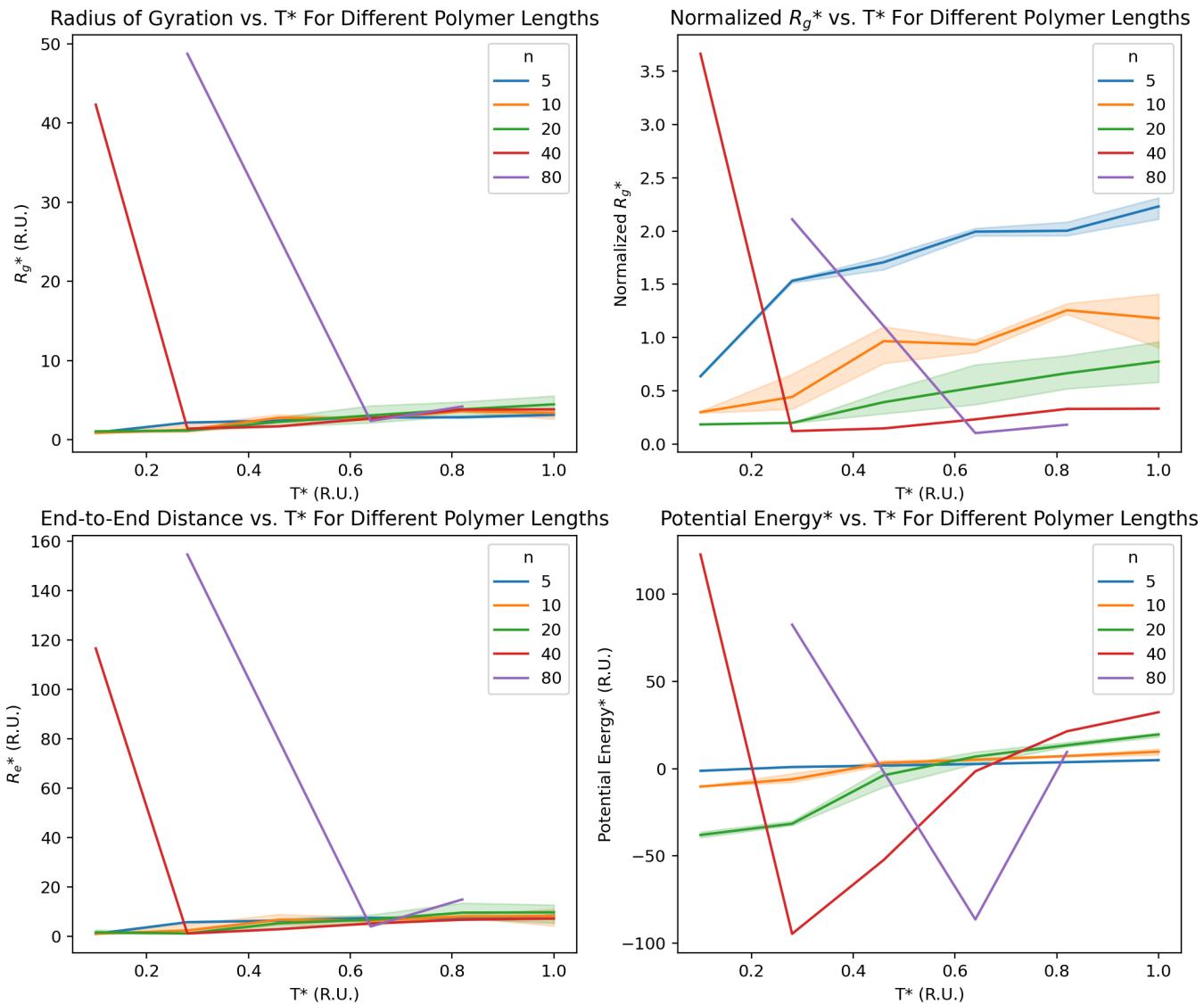


Figure 6: Different metrics plotted against temperature for different chain lengths at  $k^* = 1.0$  and  $\epsilon_{repulsive}^* = 1.24$ . Error bands show 95% confidence intervals. Note that systems of  $N = 5, 10, 20$  are aggregated across three duplicates, while systems of  $N = 40, 80$  have only one duplicate. Invalid results not plotted.

### Different Metrics vs. Temperature For Different Chain Lengths At $k^* = 100.0$

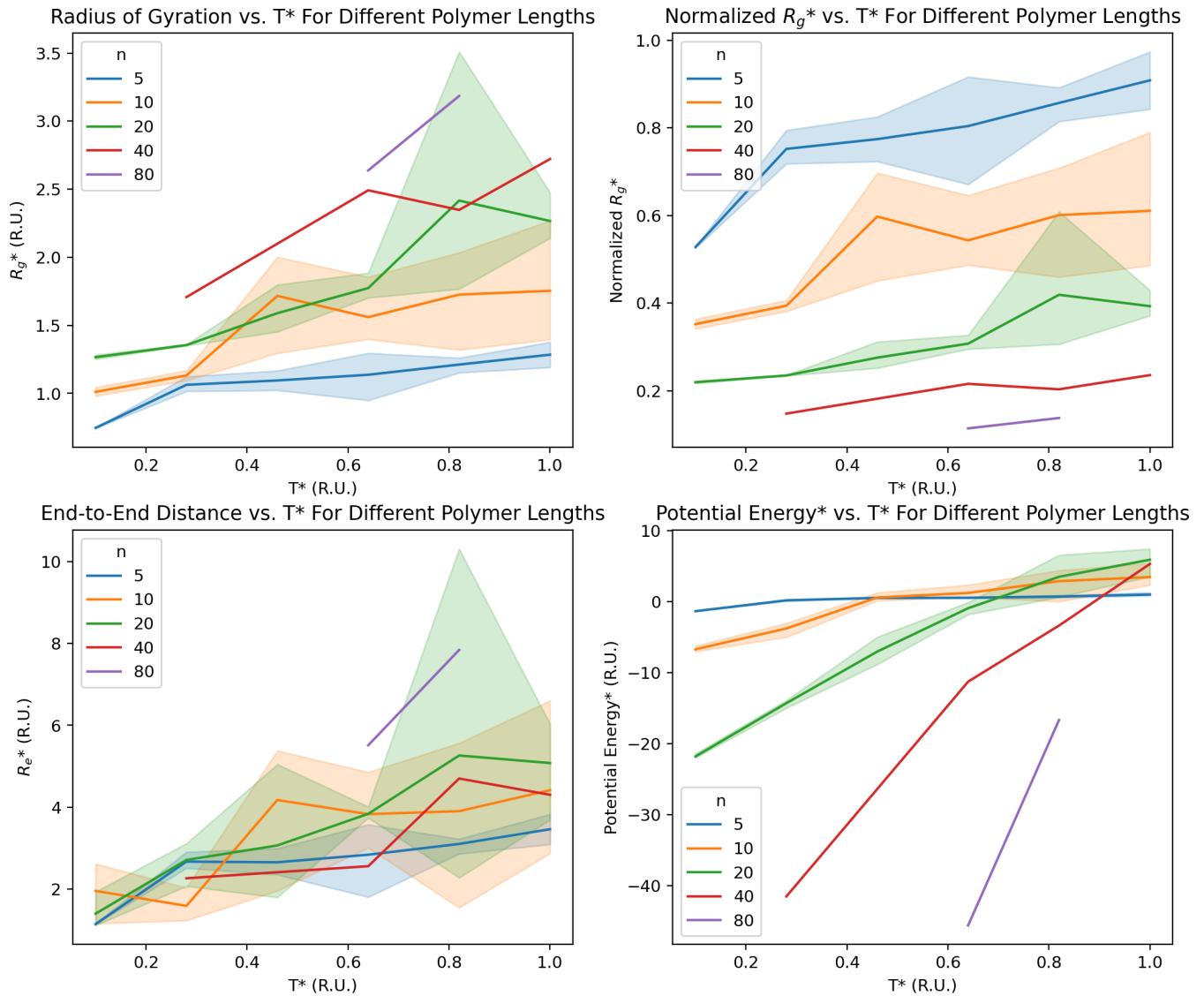


Figure 7: Different metrics plotted against temperature for different chain lengths at  $k^* = 100.0$  and  $\epsilon_{repulsive}^* = 1.24$ . Error bands show 95% confidence intervals. Note that systems of  $N = 5, 10, 20$  are aggregated across three duplicates, while systems of  $N = 40, 80$  have only one duplicate. Invalid results not plotted.

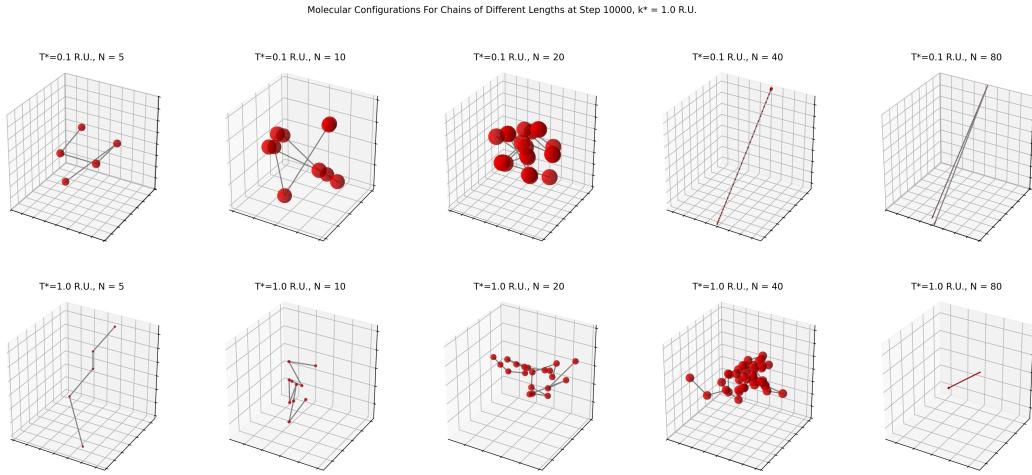


Figure 8: Visualizations of the polymer configuration at the last trajectory frame with  $k^* = 1.0$  and varying chain length  $N$ . Rows grouped by temperature; columns grouped by  $N$ . The unit cell dimensions are 25, 50, 100, 200 and 400, respectively from left column to right. Plots are enlarged around the polymer; the size of the particles is constant with regard to unit cell length. R.U. denote reduced unit.

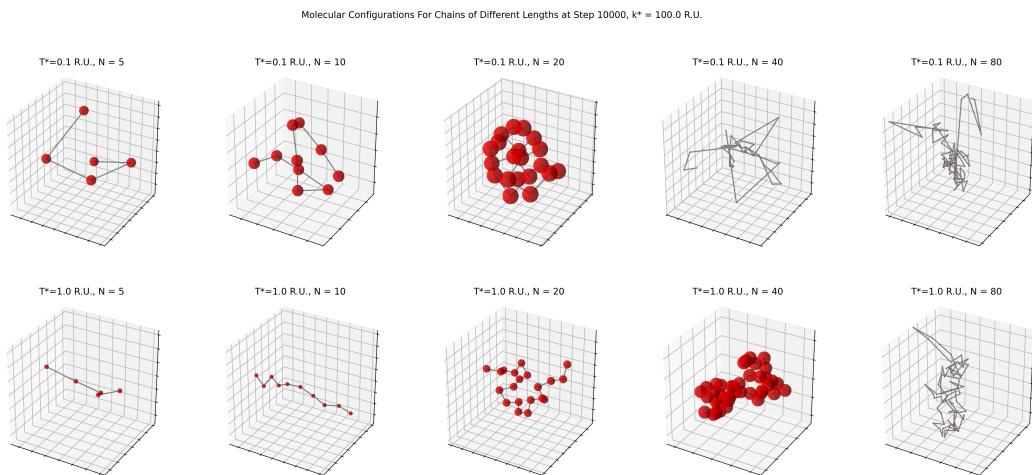


Figure 9: Visualizations of the polymer configuration at the last trajectory frame with  $k^* = 100.0$  and varying chain length  $N$ . Rows grouped by temperature; columns grouped by  $N$ . The unit cell dimensions are 25, 50, 100, 200 and 400, respectively from left column to right. Plots are enlarged around the polymer; the size of the particles is constant with regard to unit cell length. R.U. denote reduced unit.

pair of systems with  $k^* = 1.0$  and the same repulsive well depth. At low temperatures (top row), the polymer exhibit a globular phase. At high temperatures (bottom row), the polymer transitions to an unfolded, extended phase.

On the other hand, when  $k^*$  is greater than 20, there does not seem to be a difference between systems with different  $k^*$ . All five groups with  $k^* > 20$  begin to unfold around  $T^* = 0.4$  to 0.6, where  $R_g^*$  starts to significantly increase, and reaches the unfolded state at around  $T^* > 0.8$ , where  $R_g^*$  plateaus. However, the unfolded states in these systems have a radius of gyration of around 2 to 3, significantly lower than the theoretical maximum and the  $R_g^*$  of the groups with  $k^* = 1.0$ . This comparison can be visualized in figure 5, where each column is a pair of systems with the same  $k^*$ . The system with  $k^* = 1.0$  is most unfolded at  $T^* = 1.0$  (note the smaller spheres indicating the molecule is occupying a larger total volume, and therefore more extended), while the molecules in all the other systems with  $k^* > 20$  are less extended.

- **End-to-end distance:** The exact same trend is observed with  $R_e^*$ . For groups with  $k^* = 1.0$ , the transition from a folded state to an unfolded state starts at around  $T^* = 0.3$  to 0.5, and an unfolded state is reached at  $T^* > 0.6$ . The theoretical maximum  $R_e^*$  for a completely linear polymer chain with  $N = 20$  beads at bond length  $r_0$  is  $19r_0$ . The  $R_e^*$  of the unfolded chain for the  $k^* = 1.0$  groups ranges from 5 to 15, close to the theoretical maximum. This long end-to-end distance can also be observed in the bottom left plot in figure 4, where the polymer chain is almost completely linear.

Consistent with the trend of  $R_g^*$ , groups where  $k^* > 20$  show a significantly smaller increase in  $R_e^*$ . The phase transition begins at around  $T^* = 0.6$  where  $R_e$  starts to increase, and the unfolded state is reached at around  $T^* = 0.8$ . However, the  $R_e$  in all these groups are less than 10, significantly smaller than the  $R_e$  for the unfolded polymers with  $k^* = 1.0$ . By comparing the bottom plots in figure 5, it is clear that the polymers in systems where  $k^* > 20$  are still folding back on themselves even at  $T^* = 1.0$ , leading to a smaller  $R_e$  than the more extended polymer with  $k^* = 1.0$ .

- **Potential Energy:** When  $k^* = 1.0$ , the trend of  $PE^*$  with temperature is significantly different from the other groups. The  $k^* = 1.0$  group reaches a lower  $PE^*$  of around -40 at low temperatures and a higher  $PE^*$  at high temperatures. Phase transition happens when  $0.3 < T^* < 0.7$ , where a small increase in temperature translates to a large increase in potential energy, a signature of phase transitions. At  $T^* > 0.8$ , the potential energy plateaus at around 20.

The remaining five groups with  $k^* > 20$  exhibit similar behaviors with each other with respect to potential energy, but they have higher potential energies at low temperatures compared with the  $k^* = 0.1$  group and lower potential energies at high temperatures. A distinct phase transition is unrecognizable in these groups, where  $PE^*$  starts at -20 when  $T^* = 0.1$  and continues to rise gradually with temperature.  $PE^*$  plateaus at around  $T^* = 0.8$ , where it reaches around 10, which is presumably the temperature at which the polymer unfolds.

- **Summary and Interpretations:** In conclusion, the spring constant  $k^*$  could lead to two distinct behaviors of phase transitions. When  $k^*$  is on the order of 1, the system is packed closer at lower temperatures, as reflected by the lower  $PE^*$  and the slightly lower  $R_g^*$  and  $R_e^*$ , but unfolds faster with temperature and reaches a more extended state at higher temperatures. On the contrary, for all the values of  $k^* > 20$ , the same more gradual and more hindered phase transition is observed, where the polymer slowly unfolds at  $0.4 < T^* < 0.8$  and reaches an unfolded but still not as extended state at  $T^* > 0.8$ .

A higher temperature supplies thermo energy, in the form of kinetic energy coupled to velocity rescaling, to break the attractive potential of the molecule. Thermodynamically, a higher temperature favors the growth of entropy over enthalpy. Therefore, an increase in temperature is expected to cause a phase transition in the polymer. At low temperatures, the attractive forces dominate, which pack the molecule into a globule and minimizes potential energy. At high temperatures, there is enough kinetic energy in the molecule to overcome this attractive potential, leading to an unfolded state of an extended chain with a higher potential energy.

$k^*$  controls the stretchiness of the bonds. The kinetic energy of the system is given by  $KE^* = \frac{3}{2}NK_B^*T^*$ , which is supplied by thermo energy. If all of the kinetic energy turns into potential energy of the harmonic bonds, we have  $\Delta r^* = \sqrt{\frac{3K_B^*T^*}{k^*}}$ . In reduced units,  $K_B^* = 1$ . Therefore, at a temperature range of 0.1 to 1,  $0.55 \leq \Delta r^* \leq 1.73$  for  $k^* = 1$ , while  $0.12 \leq \Delta r^* \leq 0.39$  for  $k^* = 20$ . Therefore,  $k^* = 1$  makes the bonds extremely flexible, able to extend to much larger lengths than when  $k^* > 20$ . This allows the polymer to sample more configurations at both lower and higher temperatures. Therefore, at lower temperature, the molecule packs more tightly and therefore has a lower potential energy, while at higher temperature, it extends more readily and have a higher potential energy. When  $k^*$  increases further beyond 20, its effect on the stretch distances of the bonds barely changes, as the bonds are already stiff enough. Therefore, we don't observe any meaningful difference between the five groups with  $k^* > 20$ .

This result strongly supports the use of a space polymer with a weak covalent bond energy, as it is most extended under higher temperatures and remains unfolded under lower temperatures, therefore retaining its mechanical properties. However, it should be noted that weaker bonds could cause other problems, including the spontaneous disintegration of the material under the strong space radiations. It also may not be realistic to find a covalent polymer with such a weak bond. Another caveat is that such a polymer will coil more strongly under really low temperatures, which will significantly affect its properties and be detrimental to the instruments.

#### 4.3 Effect of Repulsive Well Depth on Phase Transition

Within the range of  $\epsilon_{repulsive}^*$  sampled in this study, there does not seem to be any correlation between the magnitude of  $\epsilon_{repulsive}^*$  with the phase transition temperature or how folded or extended the polymer is. The folding behaviors of the polymer seems to be affected only by the magnitude of  $k^*$  and temperature. With the same  $k^*$ , the 95% confidence intervals of the  $R_g^*$  curves,  $R_e^*$  curves and potential energy curves all overlap, indicating that there are no statistically significant difference between these sets. It is worth noting that the potential energy curves at the same  $k^*$  (figure 3) exhibit almost perfect overlap for different  $\epsilon_{repulsive}^*$ . The phase transition behaviors are the exact same as described in section 4.2.

The little influence of  $\epsilon_{repulsive}$  on the phase transition behavior of the polymer chain is surprising but not unexpected. In our model, the repulsive potential acts only on the particles exactly two bonds apart. Such a repulsion is very short-range and local-scaled, and will not have any effect on those particles that are distant and separated by multiple bonds on the primary structure but are brought closer to each other on the secondary structure of a folded polymer chain (see figures 5 and 4, where the globular polymers have multiple clusters of such distant particles that fold back together). By contrast, the attractive interaction is exerted on all atom pairs separated by three or more bonds, and is therefore long-range and could act on those folded-back atoms. Therefore, under the same  $k^*$ , temperature and the attractive well depth controls the folding of the polymer, while the repulsive well depth has no visible effect on the phase transition temperatures and how folded or extended the polymers become.

The deterministic feature of the potential energy curve under different  $\epsilon_{repulsive}^*$  demonstrates strongly the little influence it has on the folding of the polymer. The relative fluctuation of  $R_g^*$  and  $R_e^*$  could be explained by the fact that the same energy state can be achieved by a wide variety of spatial configurations.

#### 4.4 Effect of Chain Length

Two sets of simulations are performed with varying chain lengths, the first with  $k^* = 1.0$  and the second with  $k^* = 100.0$ . Both systems used  $\epsilon_{repulsive}^* = 1.24$ , although previous analysis has shown that  $\epsilon_{repulsive}^*$  does not impact the simulation results.

- $k^* = 1.0$  (**figure 6**): Despite the use of a unit cell with dimensions 5 times the chain lengths, invalid trajectories where the polymer chain extend longer than what the unit cell could hold have been encountered in the  $N = 40$  and  $N = 80$  cases, and they are excluded from the analysis. The small sample size of  $N = 40$  and  $N = 80$  could affect the interpretations of the results.

From the  $R_g^*$  and  $R_e^*$  plots, it can be observed that the chain lengths do not have a significant effect on  $R_g^*$  and  $R_e^*$  for  $N \leq 20$ . However, for  $N \geq 40$ , both metrics drop drastically from a large value at low temperatures to the average value at high temperatures, which is unexpected. An explanation for this anomaly is that at low temperatures, the polymers are trapped in a local energy minimum. From the visualizations in figure 8, we clearly see that the  $N = 40$  and  $N = 80$  polymers adopt a completely linear configuration with  $T^* = 0.1$ , which would minimize the repulsive potentials since the atoms are separated from each other as far as possible. The kinetic energy supplied by the thermo energy cannot surmount the local energy minimum for a long polymer chain. Therefore, what we observe at low temperatures for the long chains are kinetic products instead of thermodynamic products. When temperature gets higher, this repulsive potential can be overcome, and as soon as the molecules fold together, attractive potential dominates and keeps them together.

Interestingly, if not considering the anomalous kinetic configurations, both  $R_g^*$  and  $R_e^*$  of all the chains with different lengths seem to converge at the same value under a given temperature. The phase transition is observed still at around  $T^* = 0.4$ , and the full transition to an unfolded state seems to happen at  $T^* > 0.8$ . From the visualizations in figure 8, we could see the same trends. The shorter chains tend to be more extended, while the longer chains tend to be more folded, leading to a similar  $R_g^*$  and  $R_e^*$  for a given temperature. This trend is also reflected in the normalized  $R_g^*$  of the different chains (figure 6), where the shorter chains have a larger normalized  $R_g^*$ . This indicates that shorter chains extend more readily with temperature. However, the potential energy of longer chains seem to fluctuate greater, rising from a lower value under low temperatures to a higher value under high temperatures (figure 6).

- $k^* = 100.0$  (**figure 7**): The same note on invalid trajectories apply to this case. However, there are significant differences in the system behaviors in this situation compared with  $k^* = 1.0$ . The differences between chains of different lengths are amplified. Although longer chains still have smaller normalized  $R_g^*$ , indicating less unfolding, they have larger  $R_g^*$  in this scenario (figure 7). The end-to-end distances still seem to be unaffected by chain lengths. The same trend as in the  $k^* = 1.0$  case is observed with the potential energy of the system, where longer chains have lower potential energy at low temperatures, which rises faster with temperature. The phase transition is observed at the same temperatures, starting from  $T^* = 0.4$  and finishing at  $T^* = 0.8$ .
- **Summary and Interpretation:** Chain length is found to have no or limited effect on  $R_g^*$  and  $R_e^*$  under the same temperatures, and does not have a significant effect on the phase transition temperatures,

which range from 0.4 to 0.8. Potential energy fluctuates greater with longer chains, which have smaller normalized  $R_g^*$  and are therefore less extended.

These trends could be explained by the fact that the longer chains have much more possible folded configurations to sample. They could readily form local loops and fold-backs that lower their potential energy. These local clusters lead to significant stabilization of the folded configuration, and are difficult to completely open given the same thermo energy. As can be observed from figures 8 and 9, even under high temperatures, the longer chains retain multiple loops that bring particles together, while the shorter chains are extended completely. The rigid harmonic bond energy in the  $k^* = 100$  case might restrict the local fold-backs of the longer chains to some degree, leading to a small increase in  $R_g^*$ .

Overall, these results strongly suggest the use of shorter polymers in environments where folding is likely. Shorter polymers have less tendency to fold back either locally or globally, and could remain extended at lower temperatures.

## 5 Conclusion

In this study, we explored the phase transition behaviors of a linear homopolymer chain under various conditions using molecular dynamics. A consistent phase transition temperature range of  $T^* = 0.4$  to 0.8 is found, where the polymer transitions from a folded state at lower temperatures to an extended state at higher temperatures. We found that the harmonic bond spring constant,  $k^*$ , strongly influences the extent of folding, with a smaller  $k^*$  leading to less folding at low temperatures and more unfolding at high temperatures. The repulsive potential well depth,  $\epsilon_{repulsive}^*$ , on the other hand, is found to have no effect on the folding of the polymer. Finally, polymer chain length,  $N$ , also influences folding, with a smaller  $N$  encouraging a greater extent of unfolding as measured by the relative radius of gyration, although the absolute radius of gyration might not outperform a longer chain.

Such findings could guide the design of robust polymers for use in space, which are frequently exposed to extreme temperatures. The ability to remain unfolded is crucial to maintaining their mechanical properties. Our study points to the use of a weakly bonded, short-chain polymer. Of course, there are other realistic considerations into the design of polymers for use in space, for example their break down due to the high energy radiations breaking the weak linkages.

As we mentioned in our analysis, our repulsive interaction is highly localized and may not be a realistic modeling of the system, which might be the reason it is found to be inconsequential. In future studies, a more realistic simulation could be performed with a global repulsive potential that takes all neighboring particles into account, including those that are far apart on the primary structure but could get packed together spatially. Such a study may discover effects of  $\epsilon_{repulsive}$  on the folding behavior.

We modeled our system as a coarse-grain homopolymer. However, industrial polymers could be composed of different monomers, or there might be fine-scale interactions between certain atoms of the monomers that are not well modeled with a coarse-grain bead. Therefore, future study could be conducted on modeling the polymers with different repeating units or with an atom-scale model.

Finally, the use of the Lennard-Jones potential greatly models interactions between non-polar groups. However, polymers frequently have side chains that are capable of various directional interactions like hydrogen bonds and stronger, long range potentials like ionic interactions. Such interactions are poorly modeled by our system, and could be studied in future research.

## 6 Appendix: Derivation of the Normalized Radius of Gyration

The expected maximum radius of gyration for a completely linear polymer chain of  $N$  beads is derived as follows, assuming each bond is at its equilibrium distance  $r_0$ . Setting the first bead at position 0, and the last bead at position  $r_0(N - 1)$ . The center of mass of the homogeneous beads is at  $\frac{r_0(N-1)}{2}$ , and the  $i^{th}$  bead is at position  $r_0(i - 1)$ .  $R_g$  can be calculated as follows.

$$R_g = \sqrt{\frac{1}{N} \sum_{i=1}^N |\vec{r}_i - \vec{r}_{CM}|^2} \quad (16)$$

$$= \sqrt{\frac{1}{N} \sum_{i=1}^N \left( r_0(i-1) - \frac{r_0(N-1)}{2} \right)^2} \quad (17)$$

$$= \sqrt{\frac{r_0^2}{N} \sum_{i=1}^N \left( i - \frac{N+1}{2} \right)^2} \quad (18)$$

$$= \sqrt{\frac{r_0^2}{N} \left( \sum_{i=1}^N i^2 - (N+1) \sum_{i=1}^N i + N \left( \frac{N+1}{2} \right)^2 \right)} \quad (19)$$

$$= \sqrt{\frac{r_0^2}{N} \left( \frac{N(N+1)(2N+1)}{6} - (N+1) \frac{N(N+1)}{2} + N \left( \frac{N+1}{2} \right)^2 \right)} \quad (20)$$

$$= \sqrt{\frac{r_0^2}{N} \left( \frac{N^3}{12} - \frac{N}{12} \right)} \quad (21)$$

$$= \frac{r_0}{2\sqrt{3}} \sqrt{N^2 - 1} \quad (22)$$