

Project 2, TMA4320: Protein Folding

Nils Petter Jørstad, 480441, nilspj
Elisabeth Thrane, 478701, elisasth

February 8, 2021

1 Introduction

In this document, we present our answers and results to the questions in Project 2. Furthermore, we will discuss whether our results adhered to our expectations and are reasonable.

2 Visualization to ensure correct behavior

2.1 Question 1.1 and 1.2

This project consisted of twisting proteins with respect to the rules given at different temperatures. Furthermore, to make our explanations easier, we will refer to the proteins as polymers, and to their individual segments (amino acids) as monomers. The polymers have to adhere to three rules: the polymers are not allowed to split, the permutation of the monomers is not allowed to be changed and two or more monomers cannot occupy the same position. For simplification the polymers are limited to two dimensions and only clockwise or counterclockwise 90° twists. To ensure correct behavior of the polymers, a suitable way to visualize them had to be found. A simple and effective visualization method was to create a quadratic grid consisting of zeros and placing the polymer in the center as a series of integers, with the integers representing the different monomers in the polymer. To see if the behavior of the twists was correct a simple polymer was plotted before, after one and after two random twists which can be seen in Figure (1). No illegal structures appeared after two random twists. As to be extra sure of the certainty of the functionality, more twists were performed and none of the rules were broken.

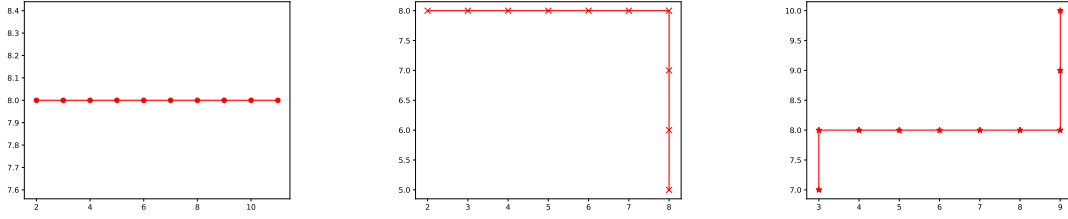


Figure 1: Visualization of a polymer consisting of 10 monomers, *Left*: plotted with no twists performed. *Middle*: plotted after one random twist. *Right*: plotted after two random twists.

3 Energy as a function of temperature

3.1 Question 2.1

With a correctly behaving polymer the next step was to take the energy and thermal fluctuations into consideration. Now the polymer would seek to minimize its potential energy when twisting, except when thermal fluctuations would allow for the opposite. The mean binding energy $\langle E \rangle$ could now be calculated for a great numbers of twists at different temperatures. Since the physics are slower for low temperatures the number of twists d performed had to be above 10000 twists. To ensure the correct amount of twists for each temperature, a function $d(T) = d_{max} \exp\{-sT\}$ was chosen. When $T = 0$ K, $d(T) = d_{max}$, so d_{max} is the maximum number of twists, d . For a polymer with 15 monomers, $d_{max} = 15000$ and $s = 4 \cdot 10^{-4}$ which is a parameter that defines the slope of $d(T)$. To find these values, we solved a system of linear equations with two unknowns s and d_{max} ,

$$d_1 = d_{max} \exp\{-sT_1\}, d_2 = d_{max} \exp\{-sT_2\}, \quad (1)$$

where we chose suitable d_1, d_2, T_1 and T_2 . This was also done for a polymer with 30 monomers. The results for 30 monomers are presented in section (3.5). Using these values, the mean energy $\langle E \rangle$ was then plotted as a function of temperature ranging from 0 – 1500 K. The resulting plot can be seen in Figure (2) for a polymer with 15 and 30 monomers.

3.2 Question 2.2

To compare the development of the energy for low and high temperatures, the energy was plotted as a function of the number of twists for the temperatures $T = 0$ K and $T = 500$ K. The plot for the temperatures can be seen in Figure (3). For the polymer at $T = 0$

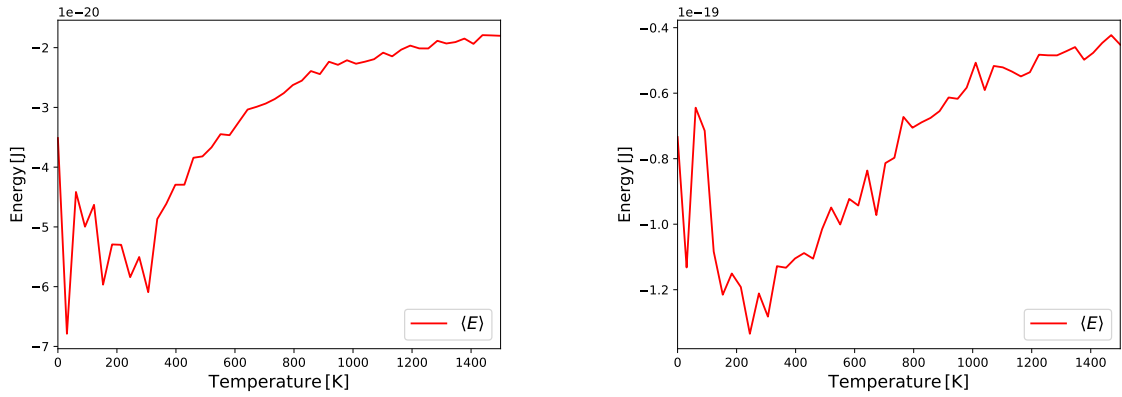


Figure 2: The mean binding energy $\langle E \rangle$ plotted as a function of temperature where the polymer was reset to horizontal position at each temperature change, *Left*: for a polymer with 15 monomers. *Right*: for a polymer with 30 monomers.

K, the energy drops after a few twists and stays constant for the remaining thousand twists. We see from the plot that 5000 twists was sufficient to approximate the probability distribution of states. After a few twists, we observe from the graph that the polymer got stuck in an energy level that it couldn't get out of because of the constant horizontal line in the plot. An explanation for this is that the temperature was so low, that there were few temperature fluctuations that allowed and pushed the polymer to twist into a different position. In addition, it could be stuck there because the polymer found its minimum potential energy and does not want to get out of it. Since the polymer always wants to decrease its potential energy, once it has managed that, it stays in this position.

For the polymer at $T = 500$ K, our first impression of the graph was that it is drastically different from the graph for $T = 0$ K. Instead of plummeting to an energy level and staying there, the energy of the polymer fluctuates and varies greatly. When the temperatures were low, it got stuck, but the opposite is true for $T = 500$ K, where the high frequency of the temperature fluctuations did not allow the polymer to remain at a low energy state for very long. However, this behaviour does not tell us much about the probability distribution of states. We can see a slight pattern in the plot because it looks like a wave, but otherwise there is not more we can see that describes the probability distribution. A good reason for this is that there is not sufficient enough twists to assess the probability distribution. When calculating the mean binding energy $\langle E \rangle$ of the states in a Markov chain, a test where we implement 10000 twists gives a more accurate value of $\langle E \rangle$, than a test where we implement 5000 twists. So to approximate the true probability distribution in this case, we need to perform more twists, such that we get a graph that tells us something more.

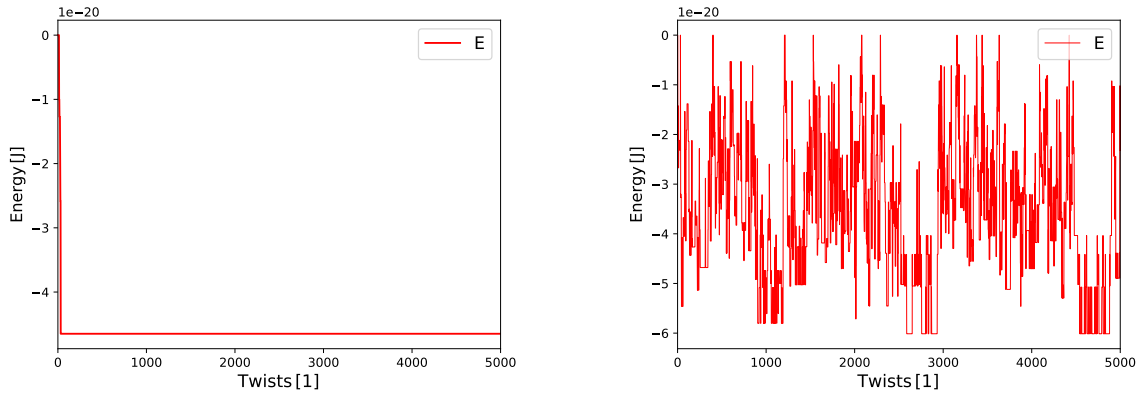


Figure 3: The binding energy E of a polymer plotted as function of the number of twists for a given temperature, *Left*: $T = 0$ K. *Right*: $T = 500$ K.

3.3 Question 2.3

From Figure (2), we see that for $T > 200 - 300$ K, the mean binding energy $\langle E \rangle$ increases as T increases. A possible physical interpretation for this is that for higher temperatures, the polymer has more tendency of twisting. In other words, it can twist more often. This could be because the polymer wants to decrease its potential energy and become a small cluster of monomers, but the temperature fluctuations allows it to increase its diameter and stretch out. So the polymer might be trying to reduce this effect by twisting more often to find a lower energy state. To explain the graph, it takes more energy for the polymer to twist at higher temperatures, in addition to that it has a higher tendency of twisting. Which means that the polymer has to acquire more energy to twist and more often, which leaves us with a higher mean value $\langle E \rangle$. That is why we see an increasing trend for $T > 200 - 300$ K.

Furthermore, from the graphs in Figure (2), we can see that for T greater than $200 - 300$ K, the slope $d\langle E \rangle/dT$ decreases as T increases. Since the graph closes in on a maximum point, the derivative approaches zero. An explanation for this could be that the temperature fluctuations become increasingly frequent and allow for more and more twists that do not lower the total energy of the polymers. After a high enough temperature the polymer will rarely twist and stay in a lower energy position, but always twist back to a higher energy position. The same will be true for all higher temperatures, which is why the slope $d\langle E \rangle/dT$ converges to zero. A physical interpretation would be that high temperature fluctuations stretch and loosen up the polymers and do not allow them to curl up into low energy structures. Another simple reason for that $d\langle E \rangle/dT$ goes towards zero, is that the

maximum energy a polymer can have is 0 J, because all the energies in this project are negative. So when the mean binding energy keeps increasing, it will eventually reach the maximum value of 0 J, and give us a maximum point.

3.4 Question 2.4

From Figure (4) and (2), we observe that the mean binding energy at $T = 0$ K and low T is very random. We expected low energy levels for low temperatures, but in both polymers of 15 and 30 monomers, the energy is higher than expected for the lowest temperatures since low frequency of temperature fluctuations should limit the polymer to only twist into low energy structures. An explanation can be that the polymers sometimes get stuck very early in a medium energy structure and can't rearrange into a lower one since the thermal fluctuations are too rare. While in Figure (2), we see that in the 100 – 400K range the thermal fluctuations are common enough to get unstuck and rearrange in even lower energy structures.

Another explanation for the randomness is that each monomer wants to decrease its potential energy such that the polymer tests different possible twists to search for this minimum state. From section (3.3), we know that at higher temperatures, it requires more energy to twist. This means that at low temperatures it doesn't require much energy, such that the polymer can test many different possibilities to find a minimum state. It could lead to the polymer ending up at random energy levels, giving us a high mean energy $\langle E \rangle$ after several thousand twists. In addition, at low temperatures, the polymer has less tendency to twist than at higher temperatures. Which leads us to speculating if the polymer got stuck at a high energy state unable to twist into something lower. Or that this causes it to twist only in a certain range of higher energy levels that gives a higher mean energy. There could be many reasons that the graphs are random in the beginning. Otherwise, the rest of the plots are very reasonable, the thermal fluctuations increase in frequency which prohibits the polymer to tangle up in low energy structures.

3.5 Question 2.5

For a polymer with 30 monomers, we used $d_{max} = 20000$ and $s = 1 \cdot 10^{-3}$ which gave us a sufficient convergent result, as we can see in Figure (2). We found these values by using the same method as in section (3) with the system of linear equations (1). From Figure (2), we can see that the graph for 30 monomers is more jumpy and has more fluctuations than the graph for 15 monomers. The 15 monomer graph is smoother, which means that it has better convergence than for a polymer with 30 monomers. So why does a longer polymer have poorer convergence? First of all, a polymer with 30 monomers is twice as long, which makes the calculations more complex. It increases the numerical complexity and this can

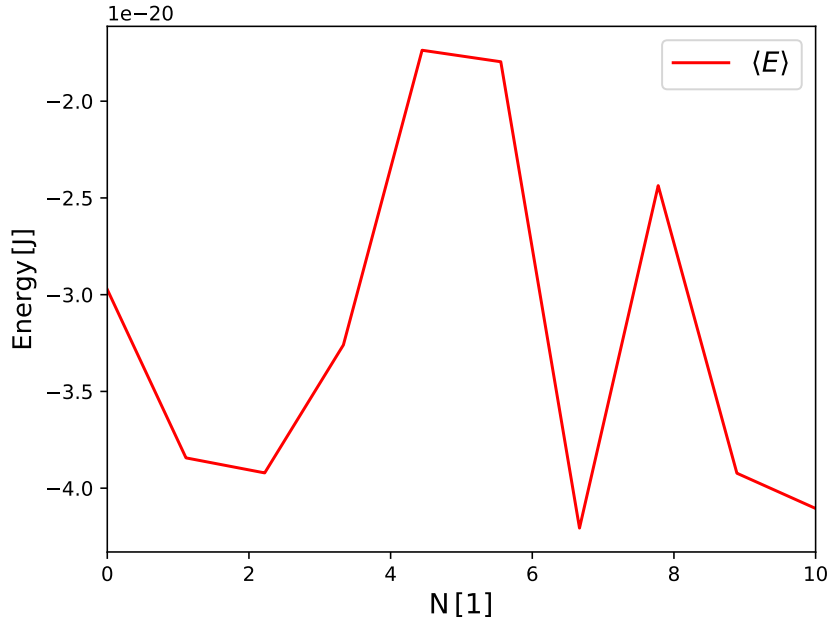


Figure 4: The mean binding energy $\langle E \rangle (T = 0 \text{ K})$ plotted after calculating it N times where the polymer was reset to horizontal position each time.

lead to less convergence. In addition, a longer polymer means that there are more possible positions that it can twist into. This means that to calculate an accurate mean binding energy $\langle E \rangle$, the polymer has to be subjected to more twists, which is why we have a higher d_{max} than for a polymer with 15 monomers. However, it is impossible for the polymer to undergo all the possible twists even with a high d_{max} , and that gives us a less accurate mean energy, and therefore poorer convergence than for a shorter polymer. So having more options to twist into, gives greater numerical complexity. Furthermore, another possible explanation for that the plot for a polymer with 30 monomers is less smooth than for a polymer with 15 monomers, is that each twist involves several more monomers, so it needs more energy to twist. Combining this with several more twisting possibilities could lead to greater energy changes, and therefore a jumpier graph. Otherwise the graphs look very similar, because the protein is essentially the same, just longer. It also behaves similarly for T less than 300 K, where the mean energy fluctuates considerably more.

4 Diameter as a function of temperature

4.1 Question 3

The plot of the mean diameter $\langle L \rangle$, as a function of temperature for a polymer with both 15 and 30 monomers can be seen in Figure (5). The results closely resemble the results of the mean energy $\langle E \rangle$ from Figure (2), which is not surprising since the energy of a polymer and its diameter is closely linked. The more curled up a polymer is, the lower energy it has. And the more energy a polymer has, the greater its diameter will be. These results from Figure (5) also compliment the previously stated physical explanations in section (3.3) and (3.4). For example one of the explanations was that at lower temperatures the polymer might often get stuck in a medium energy structure with a medium diameter. When the temperature increases up to about 300 K, it can find even lower configurations with smaller diameters because there are enough temperature fluctuations to prevent the polymer from getting stuck in a state. After a high enough temperature point, for example about 300 K, the temperature fluctuations begin to override the polymers desire to find a minimum energy state, and the polymer then begins to uncurl more and more as the temperature increases which increases the diameter and energy.

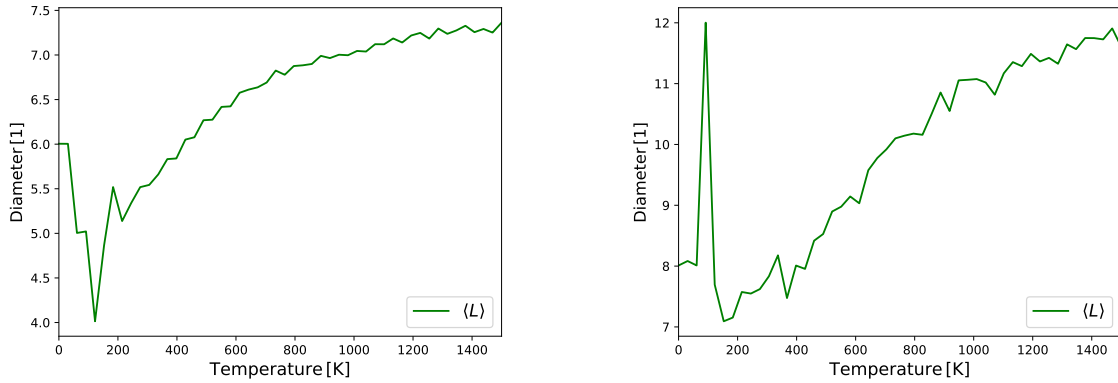


Figure 5: The mean diameter $\langle L \rangle$ plotted as a function of temperature where the polymer was reset to horizontal position at each temperature change, *Left*: for a polymer of 15 monomers. *Right*: for a polymer of 30 monomers.

5 Gradual cooling of a polymer

5.1 Question 4.1

In Figure (6), the energy E of a 15 and 30 monomer polymer was plotted as a function of number of twists while the temperature was gradually decreased from 1500 to 0 K with $\Delta T = 30$ K. The plots of both the polymers are very similar. At high temperatures the energy fluctuates very wildly. When the temperature begins to cool down the energy increasingly often stays in certain energy ranges for a number of twists and then go back to wildly fluctuating. At the lowest temperatures the energy drops into lower energy ranges since the polymer seeks the minimum potential energy. In the end it settles at a very low energy level since there are few temperature fluctuations that can change the energy. The energy does not settle at a global minimum only a local one, since there are often much lower energy positions at higher temperatures. The two polymers differ the most in the energy range they fluctuate in - the energy of the 30 monomer polymer ranges between energies larger than the other one's.

5.2 Question 4.2

After plotting the energy E of a polymer in the cooldown process, we plotted the mean binding energy $\langle E \rangle$ in the cooldown process as a function of temperature in Figure (7) for polymers of both 15 and 30 monomers. What we observe for $\langle E \rangle(T = 0 \text{ K})$ in the cooldown process is that it is flat and constant in comparison to the random behaviour of the polymer in section (3.4). This is at first unexpected because in Figure (2) we increase the temperature, while in Figure (7) we decrease the temperature. So what we would have expected at first is that they would look similar for all temperatures, but our plot in Figure (7) does not have the jagged ups and downs for low T . Instead, it settles into a local minimum. Our guess is that the main reason these plots are different is that we did not reset the polymer between each temperature change in Figure (7), while we started with a straight polymer at each temperature in Figure (2). In this section, the polymer starts horizontal and then gradually curls itself together as the temperature drops, searching for a minimum potential energy state. In section (3.4), the polymer starts horizontal at each temperature change, and tries to find a minimum state at that specific temperature, which makes it harder to find a minimum. The reason the polymer settles into a local minimum in Figure (7), is that it gets stuck in that state because of few temperature fluctuations. There are not enough to push the polymer into another energy state. In addition, the polymer's goal is to find a minimum energy state, so once it has settled into a local minimum, why would it want to get out of it? It only makes sense to stay there.

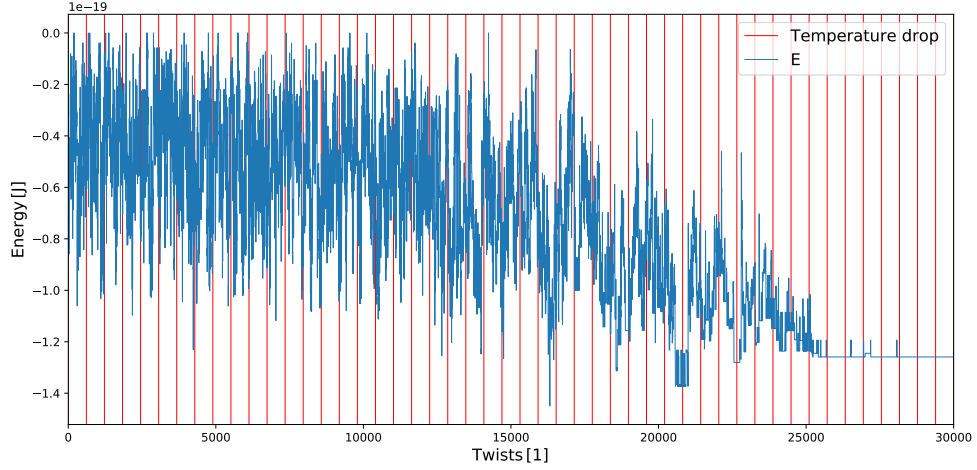
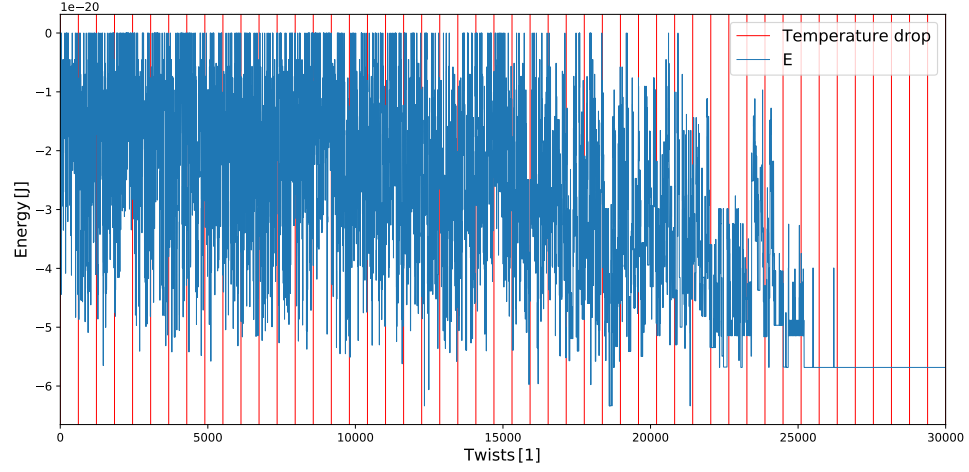


Figure 6: The binding energy E plotted as a function of the number of twists for a cooldown process, with a temperature decrease of $\Delta T = 30$ K every 600 twists, where the polymer was not reset to horizontal position at each temperature change, *Above*: for a polymer of 15 monomers. *Below*: for a polymer of 30 monomers.

When comparing the plots in Figure (7) for a polymer with 15 and 30 monomers, we see that they are very similar in shape. The mean energy range of a polymer with 30 monomers is larger than for the shorter polymer, but that is to be expected because it has more monomers and therefore greater potential energy. Also, the 30 monomer polymer

settles faster into a global minimum than the other one. A reason could be that the higher complexity allows it to get stuck quicker when the fluctuations decrease as it is easier to fold a longer polymer into a bundle than a short polymer.

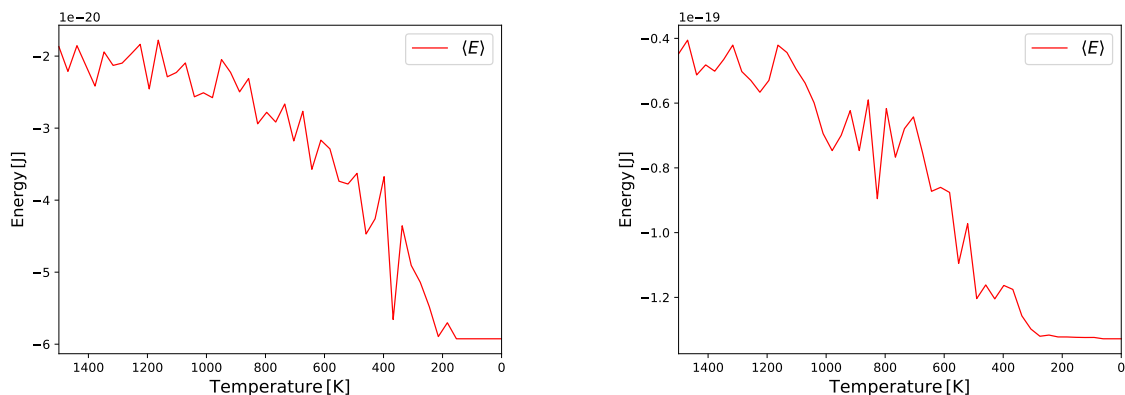


Figure 7: The mean binding energy E plotted as a function of temperature for a cooldown process where the polymer was not reset to horizontal position at each temperature change, *Left*: for a polymer of 15 monomers. *Right*: for a polymer of 30 monomers.

5.3 Question 4.3

The mean diameter $\langle L \rangle$ of the 15 and 30 monomer polymers was plotted as a function of temperature with the temperature gradually decreasing to zero. The resulting plots can be seen in Figure (8). The plots are similar to (7) and this is because of the same reason as in section (4.1): the diameter is related to the energy. The mean diameter $\langle L \rangle$ at the highest temperature begins slightly below half the starting diameter of the polymers. Then $\langle L \rangle$ gradually decreases with the temperature with some slight fluctuations and settles at a global minimum. Again this is because the polymers seeks to minimize the potential energy, and the temperature fluctuations decrease which "traps" it. Otherwise, the plots of both the polymers are very similar and mostly differ when it comes to the ranges of diameters on the y-axis, which is reasonable because the polymer with 30 monomers is obviously longer and therefore can have a larger diameter. The 30 monomer polymer diameter settles faster into a global minimum than the other one. A reason could be that the higher complexity allows it to get stuck quicker when the fluctuations decrease as it is easier to fold a longer polymer into a bundle than a short polymer.

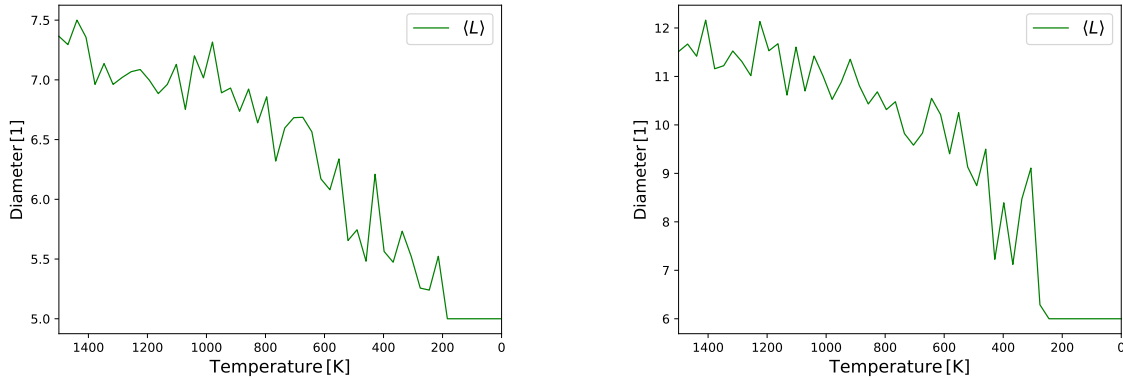


Figure 8: The mean diameter L plotted as a function of temperature for a cooldown process where the polymer was not reset to horizontal position at each temperature change, *Left*: for a polymer of 15 monomers. *Right*: for a polymer of 30 monomers.

5.4 Question 4.4

After doing several cooldowns for a polymer of both 15 and 30 monomers, we plotted what they looked like at the end where $T = 0$ K in Figure (9). As we can see from the figure, both polymers have twisted into a clustered-like position. The reason the polymers got "trapped" into this position has been explained in several of the sections above, but it is mostly because there are not enough temperature fluctuations that allow the polymer to escape this position. The polymer's goal is to find a low energy state, so when it has found this local minimum, it still wants to decrease its potential energy, so it twists some more, but there are not enough temperature fluctuations that allow the polymer to try a higher energy state which could lead to it finding an even lower energy state. That is why it is "trapped" like this.

When comparing the figures for a polymer with 15 monomers and a polymer with 30 monomers, we can see that the polymer with 30 monomers is more intricately twisted than the shorter polymer. This is obviously because it is easier to fold a longer polymer into a complex, tight structure than it is to fold a shorter one. In addition, there are more low energy positions that the longer polymer can take, many of which are much more complex than a shorter one can have. One possible reason the polymer with 15 monomers is not so intricately twisted is that it doesn't have as much energy when horizontal compared to a horizontal 30 monomer polymer. So that makes it easier for the shorter polymer to find a local minimum, because it has less energy to start with and therefore does not need to twist as much as the longer polymer.

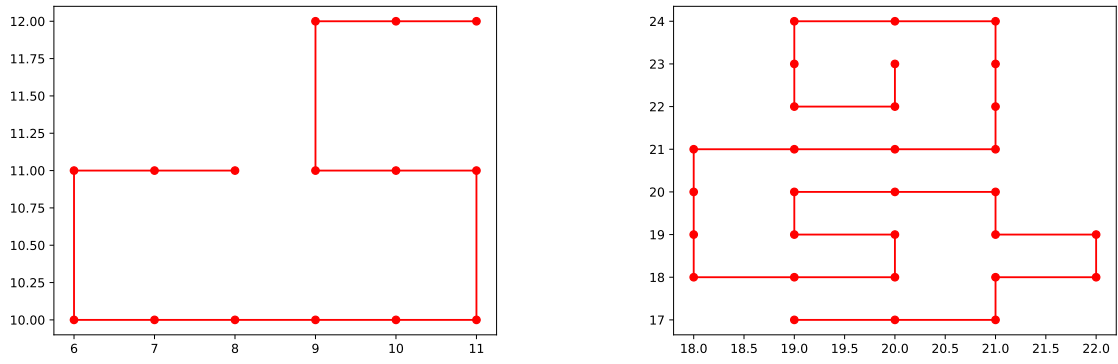


Figure 9: Visualization of a polymer at the end of a cooldown process ($T = 0$ K) where the polymer was not reset to horizontal position at each temperature change, *Left*: for a polymer of 15 monomers. *Right*: for a polymer of 30 monomers