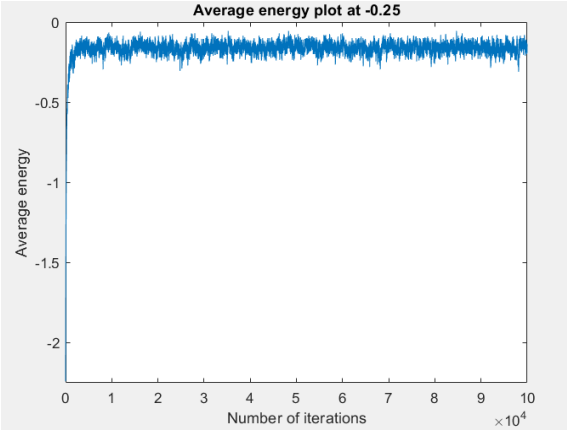
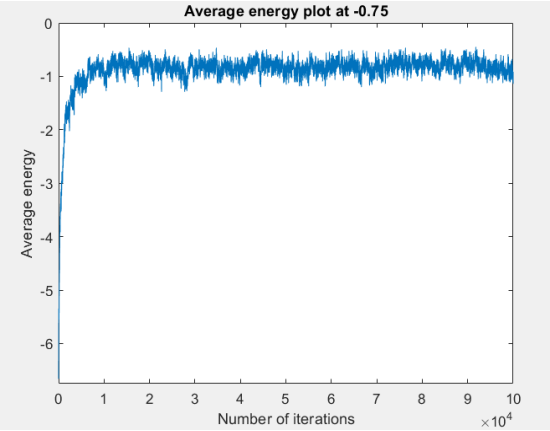
‘Protein Unfolding’ using Monte Carlo Simulation

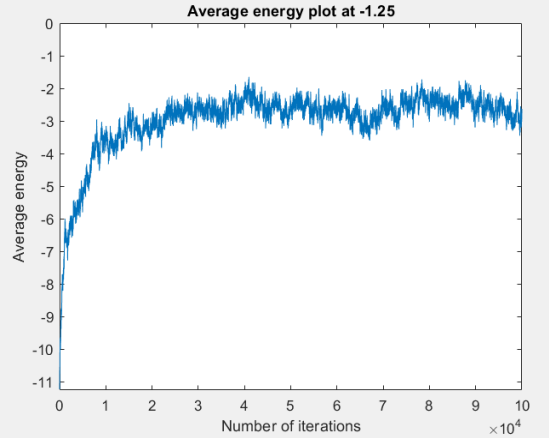
BT2042 - Assignment 1

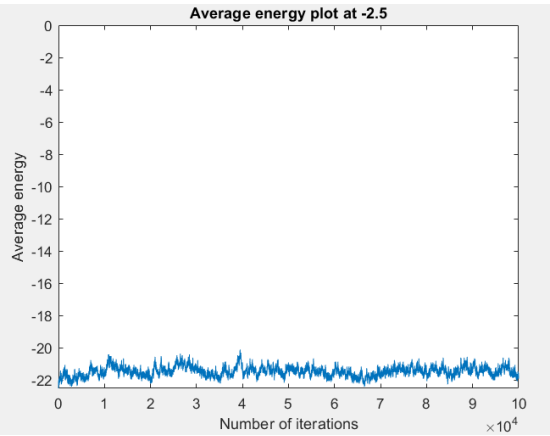
What differences do you observe across the different energy values? Rationalize them.

Energy value averaged over 50 runs for 100,000 iterations:



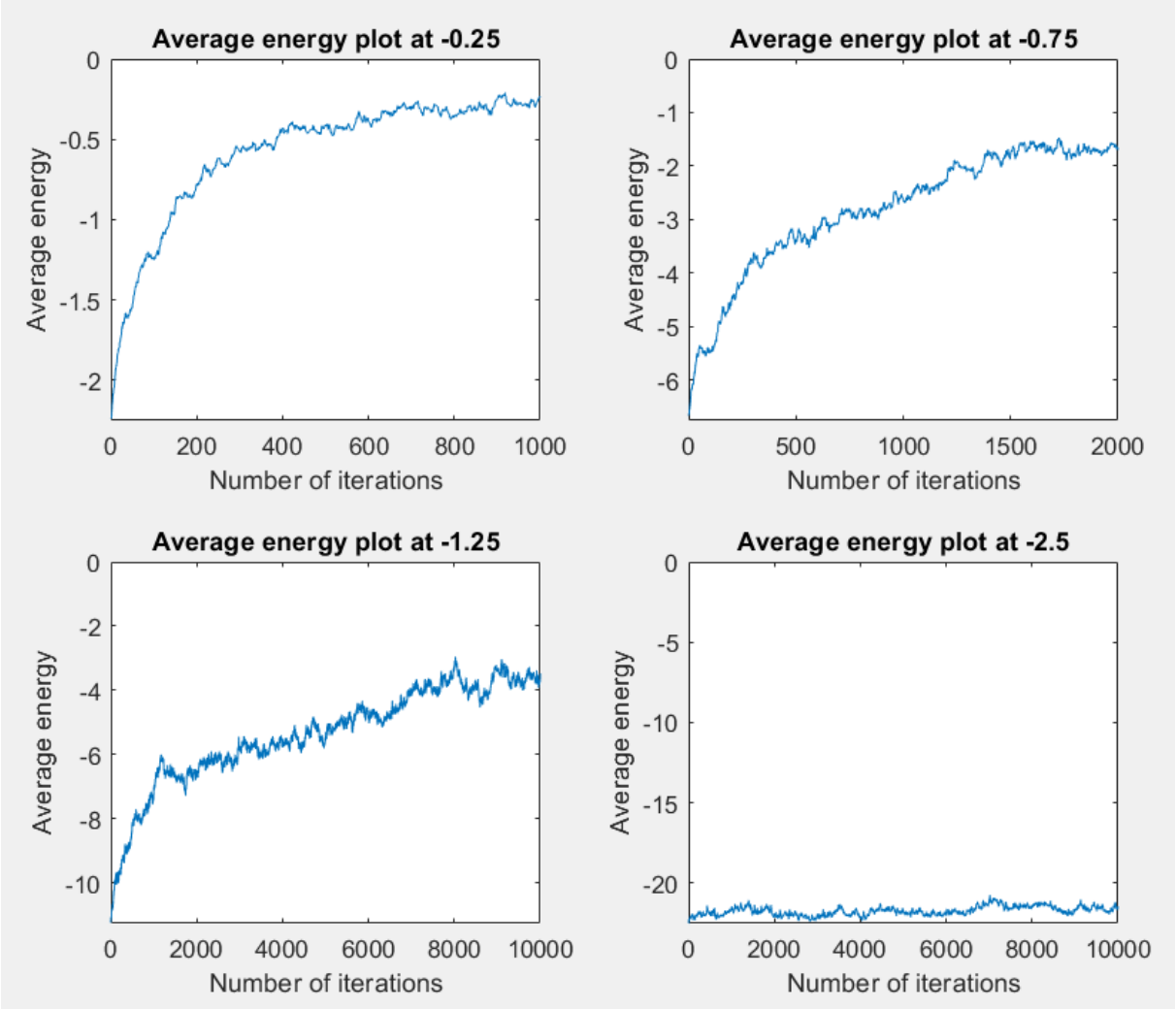






- At a low magnitude of interaction energies (eg. -0.25), many of the original interactions are easily broken down, and the polymer unfolds quickly.
- As we increase the magnitude of interaction energy, the number of iterations taken to break interactions increases; hence, the polymer takes longer to unfold.

- At very high energy levels (eg. -2.5), the thermal fluctuations are not strong enough to break the native interactions. The polymer roughly maintains its native folded state.

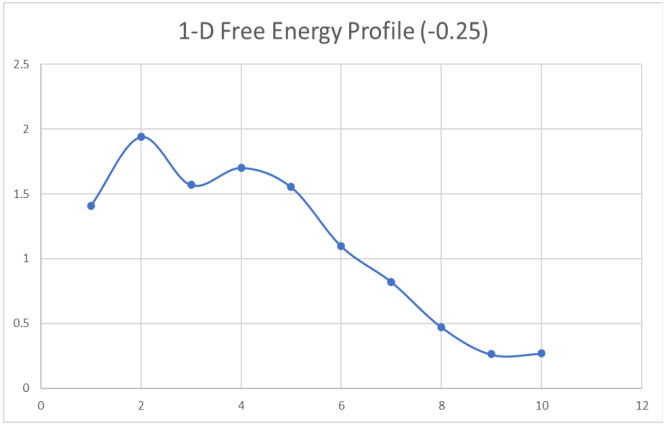


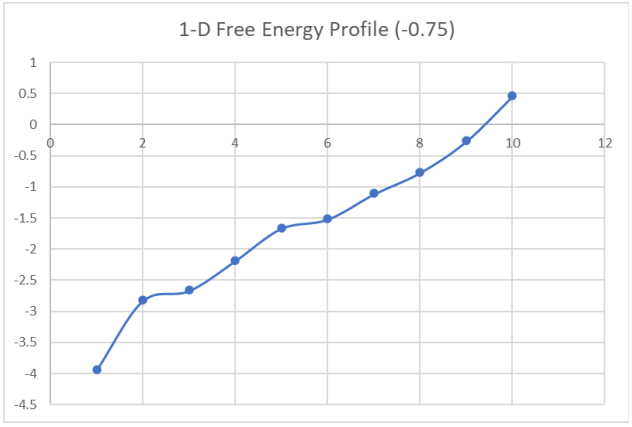
Here we have enlarged the four graphs by taking the average number of iterations it takes for the protein to unfold on the x-axis. Here we can clearly see that as the magnitude of energy increases, the protein takes more iterations to unfold. In the -2.5 energy graph, due to the high interaction energy, the protein does not unfold, and total energy remains relatively constant.

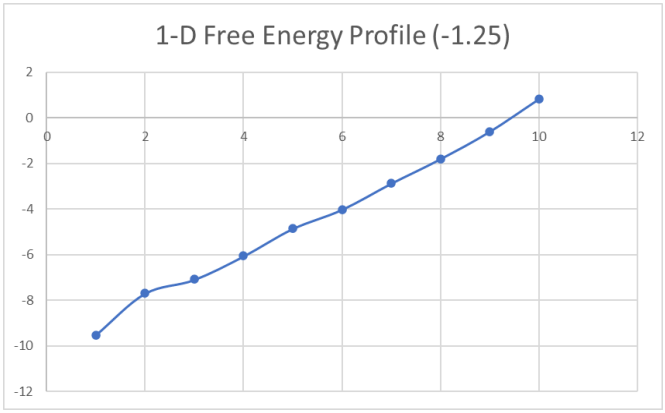
What reaction coordinate can you use to monitor the folding or unfolding of the protein? Using your own take on this, construct a one-dimensional free energy profile.

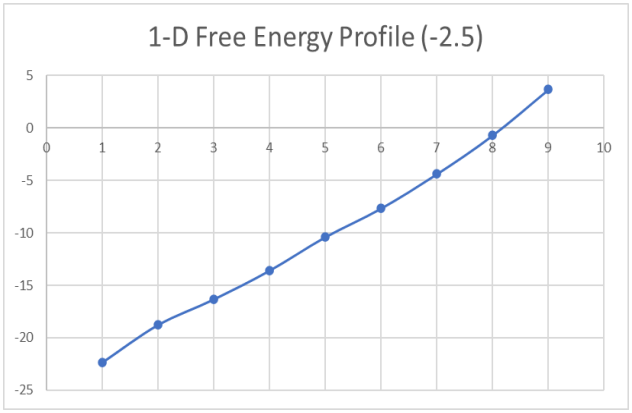
We can take the number of native interactions as the reaction coordinate and calculate free energy at each energy level as

$$F = U - TS$$
$$= (\text{Non-covalent interaction energy}) - (TK_B)*(\log_e([\text{Microstates(U)}/\text{Microstates}_{tot}]))$$







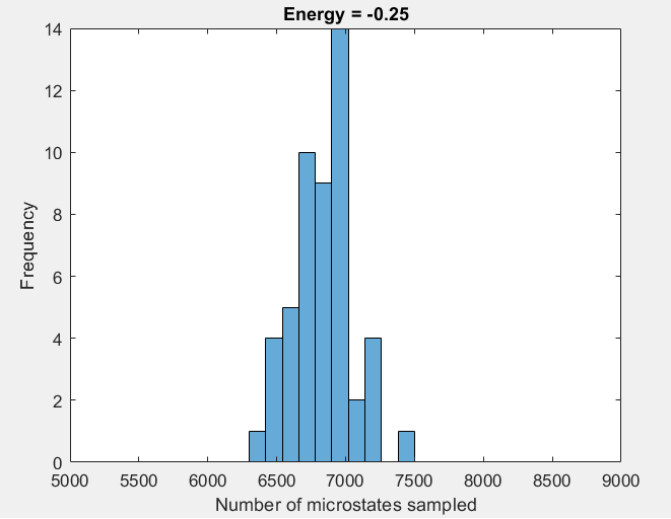


How many different microstates are sampled and how would you say a given molecule is unfolded?

The number of microstates is the different configurations that the protein exists in, over a run of 10^5 steps. We can take the radius of gyration values about the center of mass or the end-to-end distance values as a definition of different configurations. We are taking a 'score' value to sample a wider range of microstates.

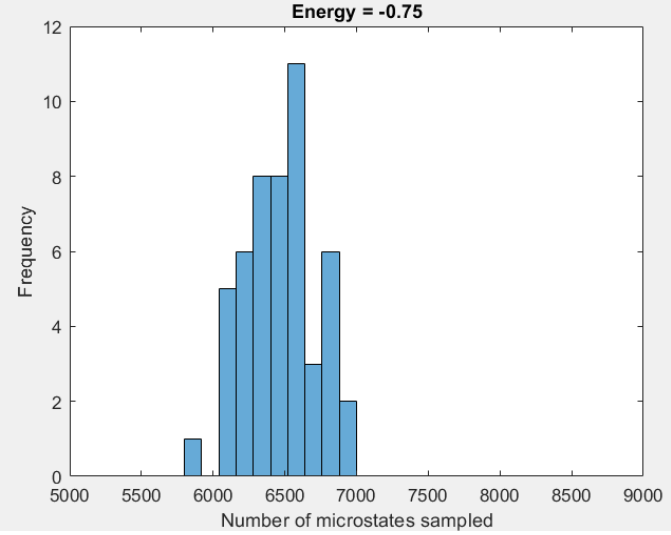
Score(Iteration) = **[***Radius of Gyration(Iteration)***]*[***End-2-end distance(Iterations)***]**

The number of unique score values over 10^5 iterations can be taken as the number of different microstates sampled over a run.



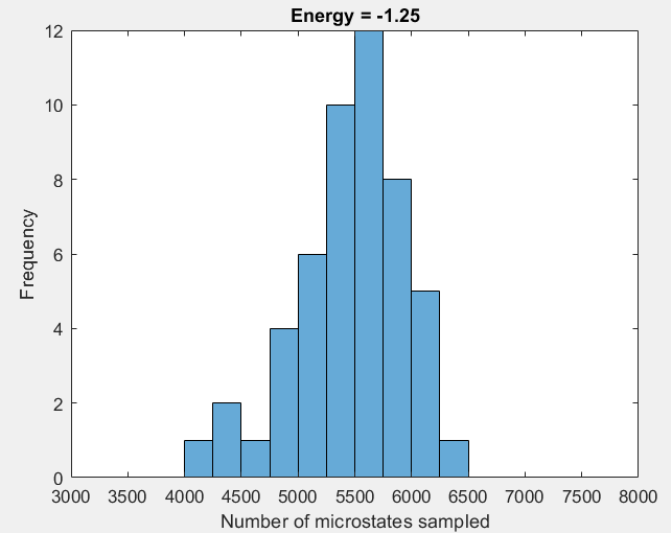
The distribution of the number of microstates across 50 runs at Interaction Energy = -0.25.

Mean no. of microstates sampled = 6384.



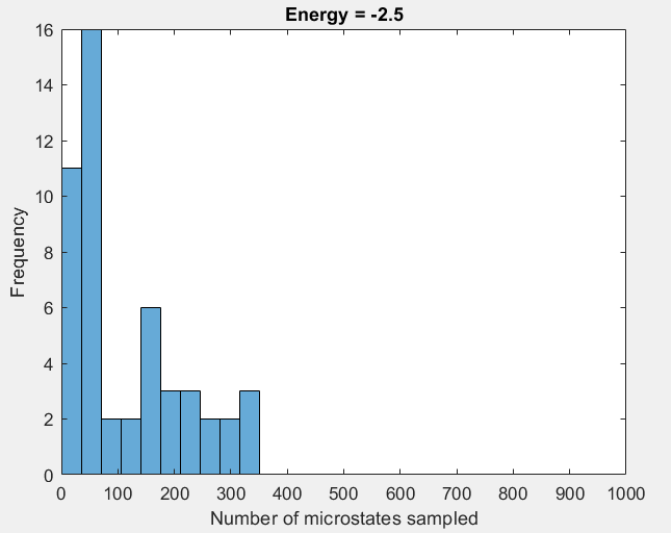
The distribution of the number of microstates across 50 runs at Interaction Energy = -0.75.

Mean no. of microstates sampled = 6460.



The distribution of the number of microstates across 50 runs at Interaction Energy = -1.25.

Mean no. of microstates sampled = 5460.

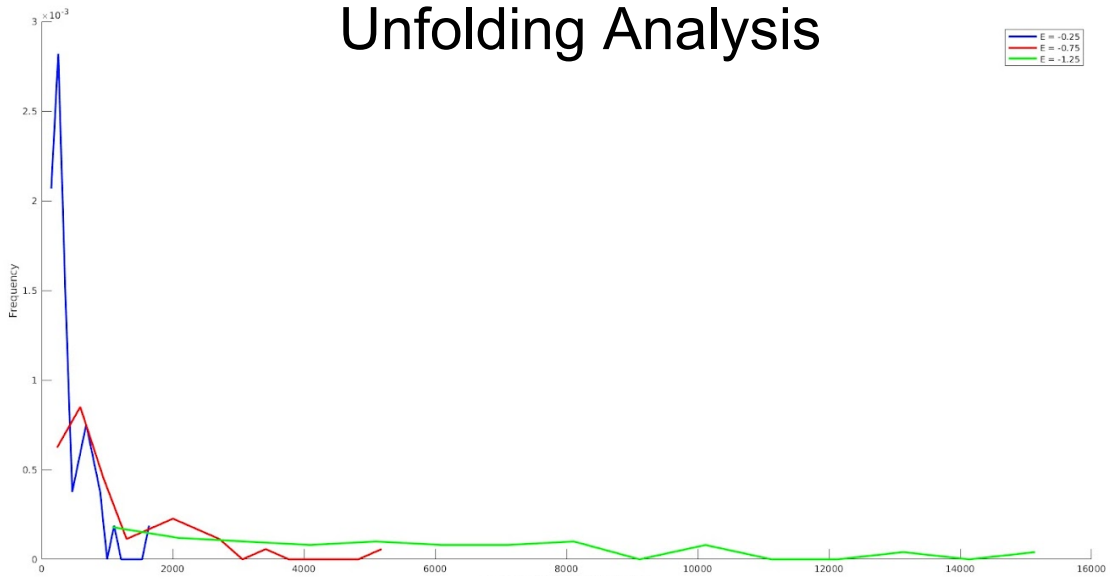


The distribution of the number of microstates across 50 runs at Interaction Energy = -2.5.

Mean no. of microstates sampled = 117.

We are assuming the protein to be unfolded completely when **all native interactions have broken** and the **interaction energy of the protein is 0.**

How long does the 16-mer take to unfold? Is there a distribution in the number of steps taken to unfold? If so, why?



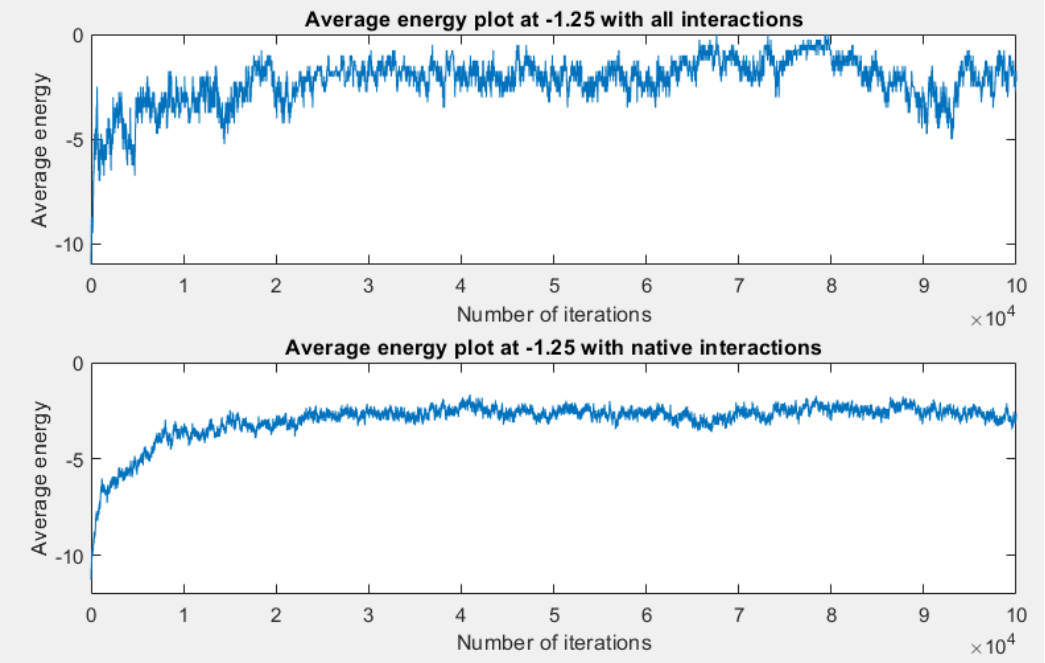
On increasing the magnitude of the binding energy, the protein becomes more stable, and there is a lesser chance of unfolding moves being accepted in the Metropolis Criterion.

- At binding energy = -0.25, the protein takes an average of 414 iterations to unfold completely.
- At binding energy = -0.75, the protein takes an average of 1177 iterations to unfold completely.
- At binding energy = -1.25, the protein takes an average of 5340 iterations to unfold completely.
- At binding energy = -2.5, the protein has extremely high binding energies and does not unfold completely even after 10^5 steps. It has an average energy of -8.35 or 3 to 4 native interactions left in the most unfolded microstate.

However, at each energy level, there is variation in the number of steps taken to unfold. This is because there are random elements while running the simulation.
(The residue being chosen to move is randomly chosen, and a randomly generated value is used in the metropolis criterion as well.)
Hence over 50 runs of 10^5 iterations, there will be a distribution in the number of steps taken to unfold.

What would happen if the beads in the polymer were assigned a favorable non-covalent interaction energy of -1.25 irrespective of whether it is present or absent in the folded structure? In other words, any bead can interact with another bead as long as it is not covalently linked. Would the molecule take longer to unfold? If so, by how many more steps?

Yes, the protein will take longer to unfold since now the protein has a higher number of possible interactions that can develop as it changes its configuration. Therefore, the protein will have to break more non-covalent interactions in order to unfold.



The plots show that the average energy over several runs is higher when residues can have all possible non-covalent interactions. The protein tries to maintain stability by binding to nearby residues. While all interactions are possible, there are more restrictions on movement as well.

According to the Data,

The average number of moves with all interactions is 12500.

The average number of moves with native interactions is 30500.

The average number of iterations taken to unfold is 5340 if only native interactions are present and 4570 if all possible interactions are present.

Members:
Pravar V BS21B025
Raghav Mathur BS21B026
Shubhan V BS21B032