simulating IDP networks structure changes
Computational Cell Biology

Introduction

Intrinsically disordered proteins (IDP) are proteins with no distinct and unique conformation, they instead fluctuate and behave more like flexible polymers. They can also form weak interactions between each other and thus condense, under certain circumstances, into a liquid phase droplet. The IDP inside the droplet form a network whose properties depend on their interaction sites. Inside a cell, these sites could be turned on and off by kinases to change, for example, the permeability of this network to small molecules. The aim of this project is to simulate liquid IDP droplets and see the effect of the activation of additional interaction sites.

Methods

polymers with attractive end caps and switchable attractive sites along their backbone are used to model IDP. The observable used is the end-to-end length : if a network is formed, the end-to-end length of its constituting proteins will be different from the end-to-end length of free proteins and will depend on the structure of the network. This observable gives an idea about the packing of the network, a small lattice will result in shorter end-to-end lengths.

The diffusion coefficient of the beads of the polymers inside the network is another possible observable and was also used at first. The idea is that a polymer in a droplet will have its movement somewhat restricted depending on the the structure of the network and that this would be visible through the diffusion coefficient. This method was quickly abandoned because it was not possible to distinguish the different networks of this project with it.

The simulations are conducted as follows. First, an initial network with only the attractive end caps active is simulated until equilibrium of the end-to-end length. From this point, the various possible intermediate attractive sites are turned on and the simulation is run long enough to have a stable average end-to-end length for each network. Finally, the intermediate attractive sites are turned off to compare the relaxation time and the condensation time of the different networks.

Simulation specifics

DPD simulation is used in this project. To simulate water, beads with a conservative force parameter of 25 between them are used. Two different polymers with different lengths are simulated but the attractive sites all have a conservative force parameter of 5 between them when active. The table below summarize the different conservative force parameters. W is water, B is the bead used for the backbone, S-on is the activated bead for the attractive sites and S-off is the deactivated bead.

	W	B	S-on	S-off
W	25			
B	25	25		
S-on	23	25	5	
S-off	23	25	25	25

The first smaller polymer has a 17 beads long backbone, two attractive end caps and one switchable attractive site in its middle. This results in two configurations : two active site with 17 beads between them and 3 active sites with 8 beads between them. This polymer is not very realistic : IDP are usually longer and have many interaction sites. It is used as an initial proof of concept and as a point of comparison for the second larger polymer simulated. For this polymer, the relaxation of the condensed network back into the initial one was not done.

The second polymer has a 31 beads long backbone, two attractive end caps and multiple switchable attractive sites resulting in 3 possible configurations :

1. 2 sites with 31 beads between them
2. 3 sites with 15 beads between them
3. 5 sites with 7 beads between them

This polymer is closer to an IDP than the first one but could be even longer. More possible sites would be better suited to characterize the structure changes in an IDP network but the simulations in this case would require a larger simulation box and thus take a lot of time.

Note : the attractive beads along the backbone can seem to be sticking out too much in the illustration but they are attracted to each other when active and will thus form more compact sites.

The simulation box for the smaller polymer is 25x25x25. The simulation box for the second polymer is 30x30x30. For additional details on the parameters used (Hookean spring bonds parameters, stiff bonds parameters, etc.), see the initial files attached at the end of the report.

Results and Discussions

Figure 1 shows the evolution of the simulation of the first polymer. The initial network (2 sites, 17 beads spacing) takes about 30'000 steps to form and stabilize from an initial random state. it results in a mean end-to-end length of 5.36 with a standard deviation of 0.17 (computed between time steps 30'000 and 60'000). After the activation of the intermediate site, the network changes its structure relatively rapidly : it takes about 5'000 time steps for the end-to-end length to stabilize around a new mean. The mean end-to-end distance of the new network (3 sites, 8 beads spacing) is 4.64 with a standard deviation of 0.16 (computed between time steps 70'000 and 100'000).

As can be seen on figure 2, the two networks are clearly different. It is visible that after the activation of the intermediate site, the lattice is smaller and the network is thus more compact.

Even if this polymer is not really close to an IDP, this experiment shows that in polymers forming a network, the activation of an attractive site can "rapidly" change the tightness of the network (it is not possible at this point to say if the change is really that fast, there is no comparison with real times, but compared to the formation of the initial network it is fast. Additionally, these times depend on the conservative force parameter between active beads). The change in end-to-end length is small but definitely noticeable (figure 1) even if the change in the density of the network is rather big (figure 2).

The curve in figure 1 can seem to be linearly descending, the two means characterizing the two networks are not so obvious. running the simulation for longer in each configuration would probably have made it clearer but previous proof of concept experiments showed very similar means for the same networks.

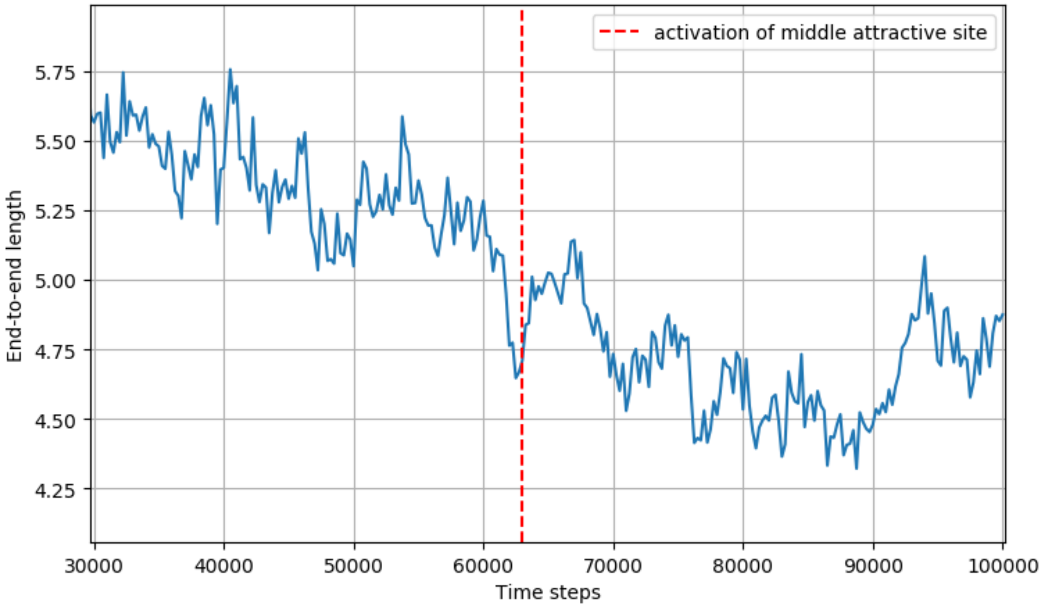
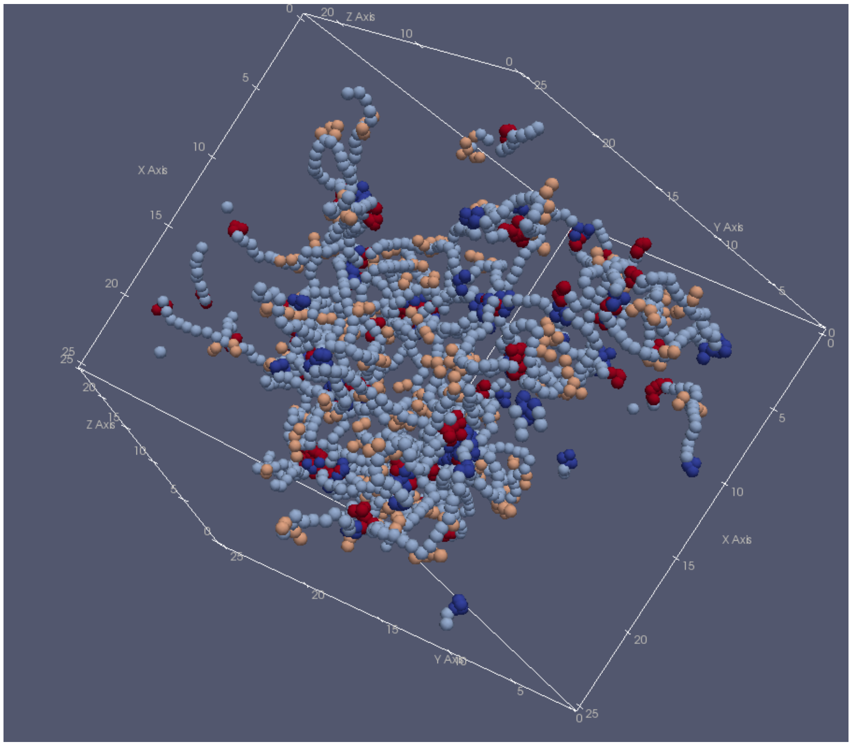
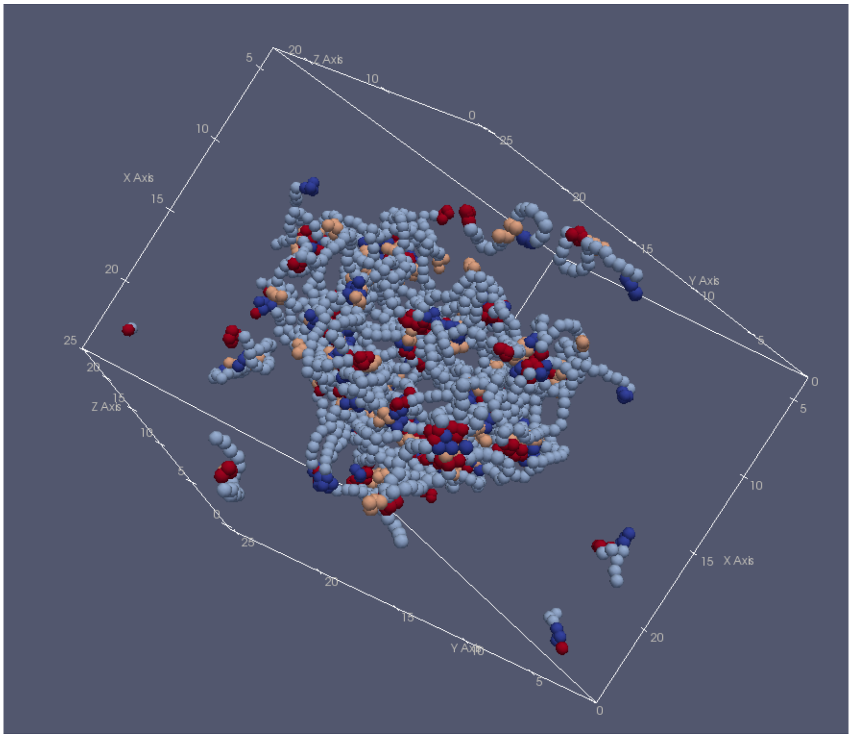


Figure 1: evolution of the end-to-end distance for the first polymer



(a) 2 active sites (17 beads spacing). Snapshot taken at time step 55’000.



(b) 3 active sites (8 beads spacing). Snapshot taken at time step 90’000.

Figure 2: droplets formed by the first polymer. red and blue : end caps, yellow : intermediate site, grey : backbone.

As can be seen on figure 3, the initial network takes more time to stabilize with the larger polymer (around 250'000 time steps). On figure 4 is displayed the end-to-end length for the different experiments :

1. no activation leading to only 2 active sites
2. middle site activation leading to 3 active sites
3. all intermediate sites activation leading to 5 active sites

After a stabilization time, the three networks have clearly different mean end-to-end lengths. Figures 6, 7 and 8 show the different droplets formed (snapshots taken at time step 300'000). the increase in number of sites (and thus decrease in distance between sites) clearly has an effect on the density of the networks.

The distance between the attractive sites probably has a direct influence on the density of the network, and thus its permeability to small molecules. On figure 5 is displayed a normalized end-to-end length as a function of the distance between the attractive sites. The normalization is done with the end-to-end length of the polymer with no active sites (not even the end caps) and is thus some kind of percentage of the inactive polymer end-to-end length.

The points coming from the data of the first polymer are really close to those coming from the second one. This could give more validity to the claim that the density of the networks is related mainly on the distance between sites. with two different polymers, the same distance between attractive sites gives rise to the same relative decrease in end-to-end length.

Another interesting trend that seems to show on this figure is that the variation in end-to-end length (that informs us about the packing of the network) is greater for small distances than large distances between attractive sites. This would mean that for short distances between sites, a small change in these distances would create a large change in the packing of the network. This would be especially interesting because it would mean that the activation a a few attractive site could significantly change the global structure of the network. Unfortunately, there is too few points in this graph to draw real conclusions on this relation. The relation could also be linear between two plateaus at a maximum packing and at 1.

The two sites initial network seems to equilibriate after around 250'000 time steps. This could be problematic since the activation of the intermediate sites takes place at 200'000 time steps : the starting network is not exactly an equilibrium 2 sites network. the simulations should probably have been run for longer before the activation. However this is not too problematic since the networks have then more time to equilibriate and stable means are visible for all three droplets. Nonetheless, This could introduce errors in the estimation of the condensation time (this is why this estimation is done in a very general way in the following section).

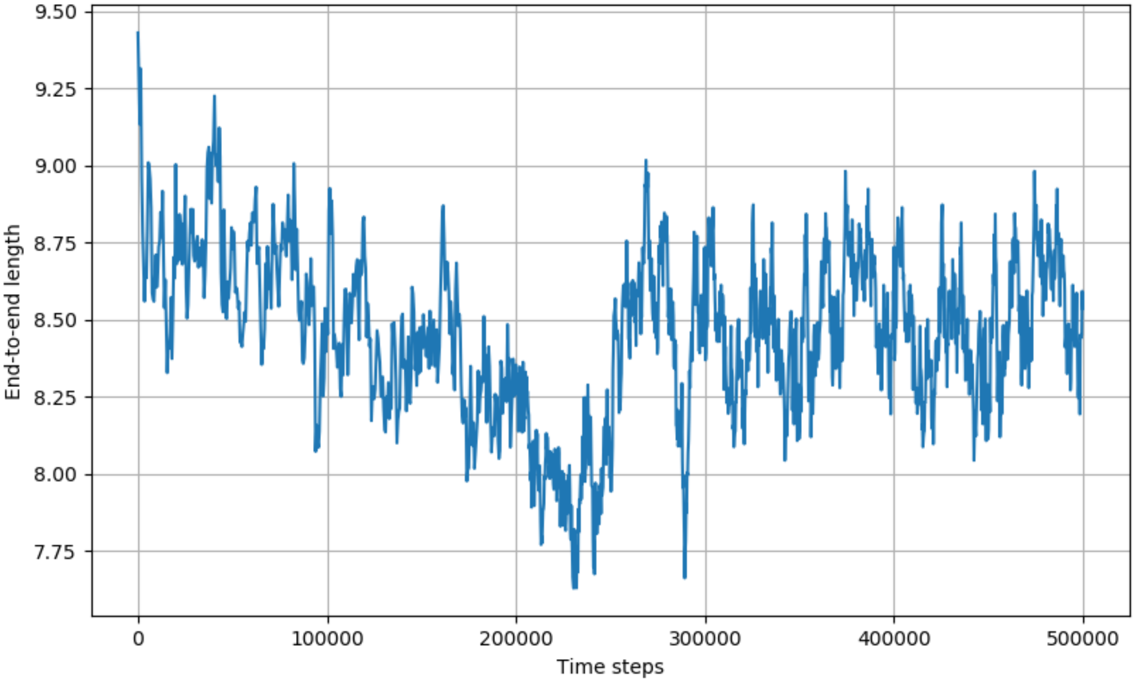


Figure 3: formation and stabilization of the 2 sites network (end-to-end distance as observable)

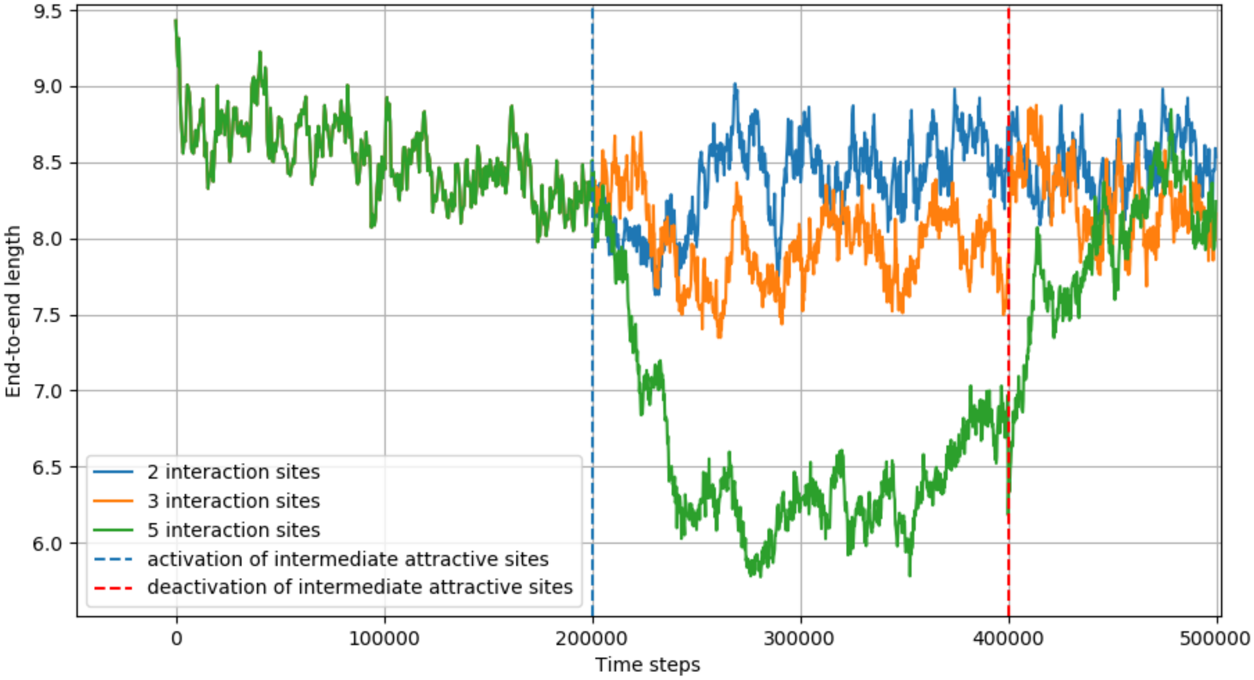


Figure 4: evolution of the end-to-end length throughout all the simulations

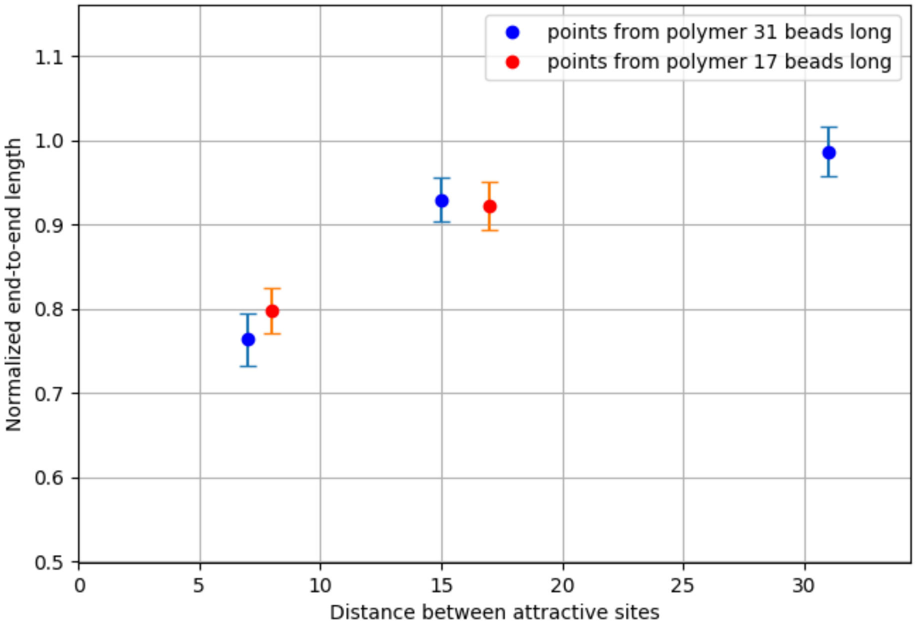


Figure 5: Normalized end-to-end length as a function of the spacing between attractive sites.

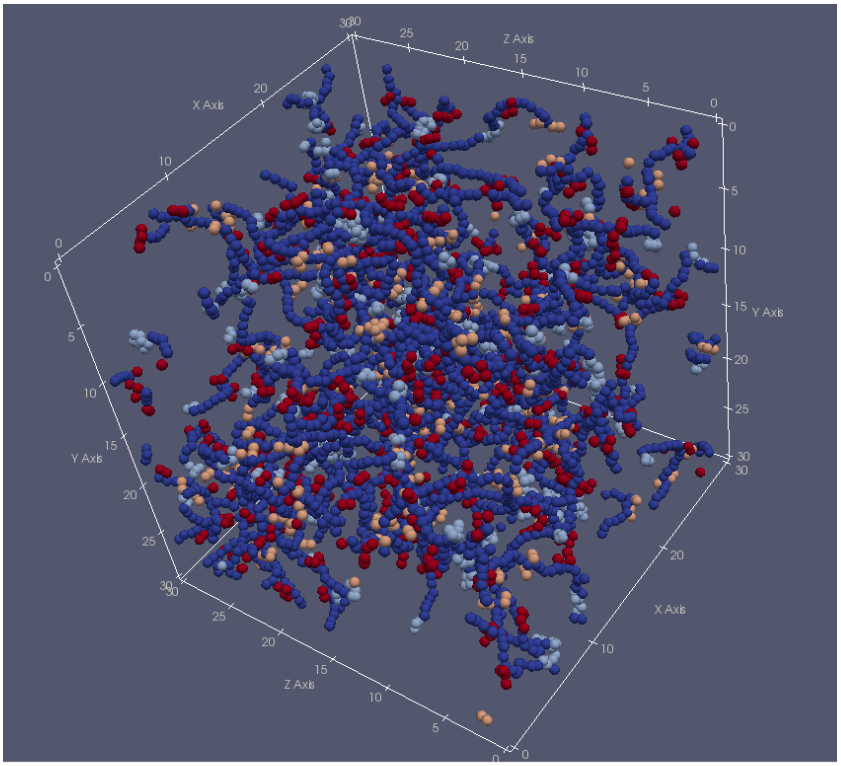


Figure 6: 2 sites, 31 beads spacing

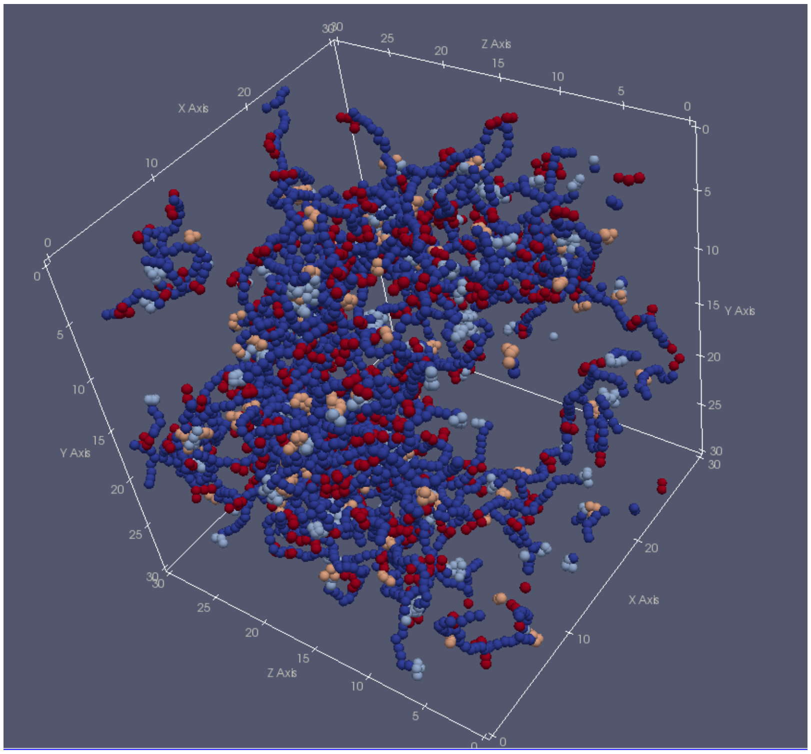


Figure 7: 3 sites, 15 beads spacing

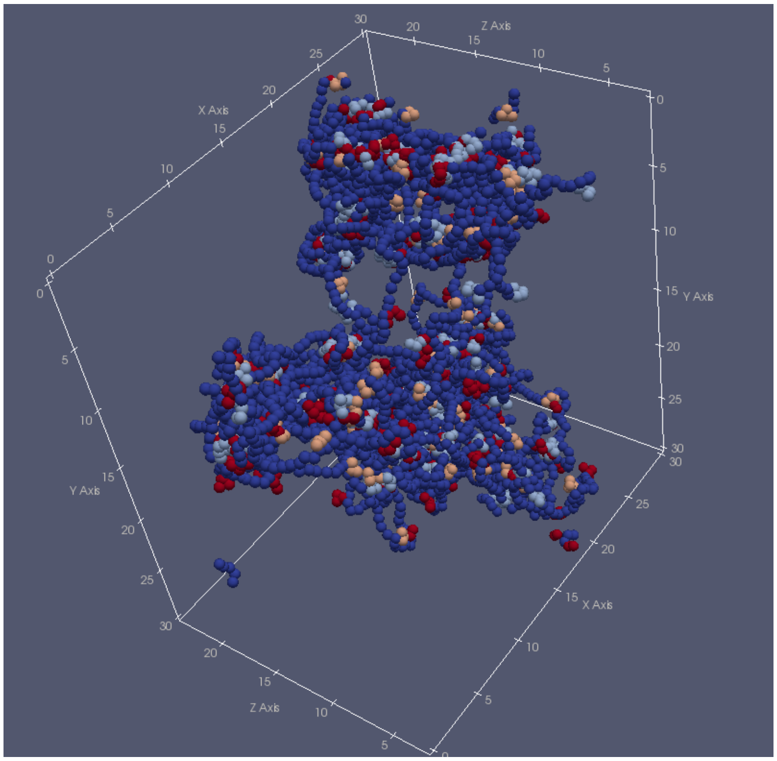


Figure 8: 5 sites, 7 beads spacing

Figure 9 shows the transitions between the initial two sites network and the other networks. The blue curve is a baseline : it is always the initial network with no changes to it.

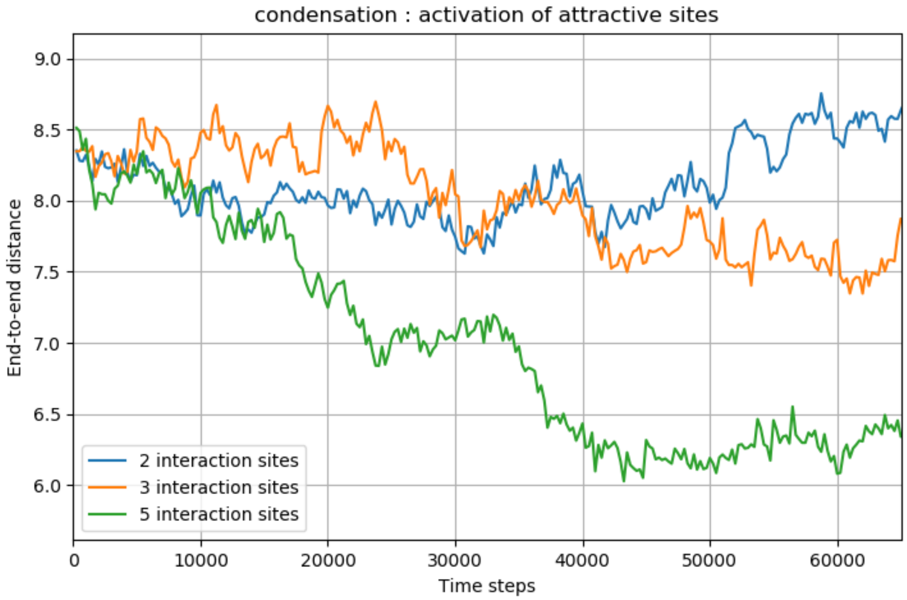
The formation of the networks seems to take around 40'000 time steps for both the 3 and 5 sites networks. The relaxation into the initial network also seems to take around 40'000 time steps for the 5 sites network but seems to be much faster for the 3 sites network : around 10'000 time steps. It would make sense for the relaxation to take less time. To form a denser network, the polymers must explore the space through diffusion and if by chance their active sites meet other active sites, they will stick. The network is stable when all sites have found other sites (or if the rate at which they break interactions is the same as the rate at which they form interactions). The relaxation into a less dense network is easier : diffusion itself drives the non sticky beads apart. The relaxation of a really dense network could take more time if the polymers are entangled in a complex way. This could explain why the relaxation of the 5 sites network takes as long as its condensation.

The condensation and relaxation times are relatively short compared to the time it take for the initial network to form. It is also noticeable that the formation and the transition times between the networks are longer for the larger polymer than for the shorter one.

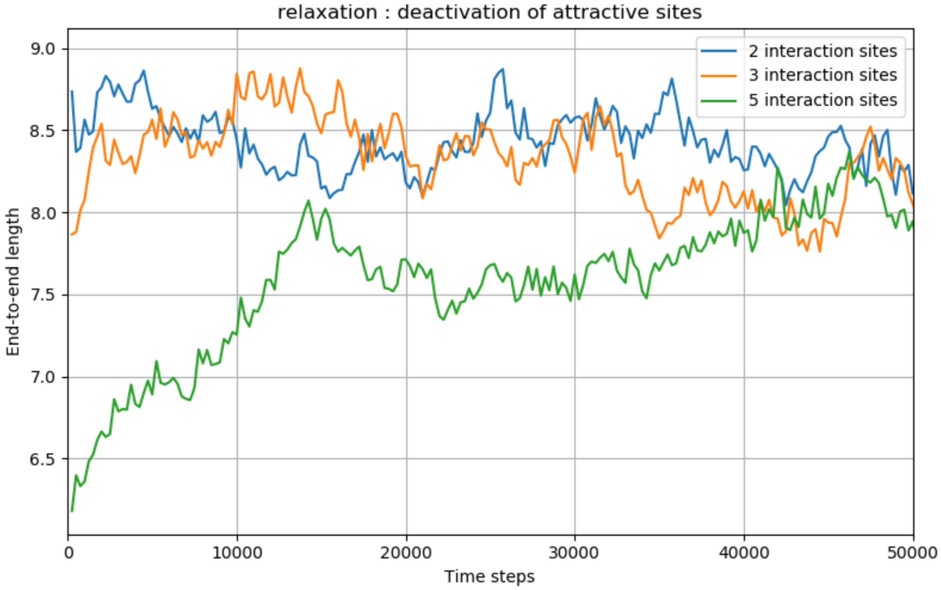
The graphs of figure 9 have a lot of fluctuations, it is difficult to say when the networks have formed or relaxed. Additionally, the time it takes probably depends on the state of the starting network (for example if the polymers are entangled a lot). Averaging the end-to-end length of multiple runs of the same experiment with different random seeds would remove variations and give more robust results to compare condensation and relaxation.

Conclusion

Flexible polymers with attractive sites (used to model IDP) tend to form networks (droplets). Their density can be changed drastically by activating additional attractive sites. Additionally, this density seems to depend mainly on the distance between the attractive sites. These networks can rearrange into each other through activation and deactivation of attractive sites in relatively short times (compared to the formation time of the initial network from a random state). These times seem to be longer for larger polymers.



(a) formation of the different networks



(b) return to the initial network (2 sites, 31 beads spacing)

Figure 9: transitions between initial network and condensed networks

Sources of errors

A source of statistical error introducing fluctuations in the computation of the end-to-end length inside the networks is the fact that not all polymers simulated are inside the droplet. Some of them are free and the resulting average end-to-end length doesn't exactly reflect the average end-to-end distance of the polymers inside the network. Using a visualization tool (paraview) showed that very few polymers were free in the experiments of this project. This is why no measures were taken to reduce this source of error.

Some configurations of networks that can seem at equilibrium but are not can also introduce errors. For example the 5 sites network in figure 8 is composed of 3 dense lobes instead of one roundish droplet. The end-to-end length of the polymers in this configuration could be different than if it was one single droplet. The rearranging seems to take a really long time in this case, the three lobes were visible throughout the whole 5 sites network simulation ([250'000 : 400'000]).

Overall, every result would be more statistically relevant if more simulations had been done and averaged

Extensions to this project

This section give improvement ideas on this project and other ideas to extend it.

Some results are not very robust. In particular, figure 5 has very few points characterizing the relation 'distance between active site / Normalized end-to-end length'. The problem is that to have more possible sites, the polymer simulated would need to be really long and the time required would increase a lot. To resolve this problem, the fact that the density of the networks seems to depend mainly on the distance between active sites and not the total length of the polymer could be taken advantage of. To confirm this assumption, multiple networks of polymers with different lengths but same spacing between active sites could be simulated. Their normalized end-to-end length could be compared that way. If they were equal, the assumption would be confirmed. Then, multiple networks of short polymers with various distances between their interacting sites could be simulated until equilibrium and their normalized end-to-end length could be used to have a better relation.

If the results were more robust in figure 9 and times of relaxation and condensation were established by averaging multiple simulations, a relation between condensation/relaxation times and difference between starting and resulting end-to-end length or distance between sites could be established. This would give a more complete idea about the time scales involved in the rearranging of those networks. Those times also depend on the conservative force parameters between active beads.

The dynamic nature of those network is also something that could be studied. Are the polymers forming and breaking interactions between them ? do they leave and re enter the droplet ? These questions could maybe be answered by tagging some polymers and watching their behavior with respect to the droplet.

Other observables could be used to better characterize the droplets, for example their radius of gyration.

It would be interesting to see how a small molecule diffuses through the different networks.

Calculations could be done to compare the simulated lengths and time scales to real ones. Trying to simulate a polymer closer to a real IDP could also be interesting.

input file

```
dpd

Title        " protein liquid liquid phase separation "
Date 02/12/19
Comment    " formation of intrinsically disordered protein liquid droplet "

State    restart
         RunId   m2s_in
         StateId   200000

Bead    W
        0.5
        25
        4.5

Bead    B
        0.5
        25    25
        4.5   4.5

Bead    E
        0.5
        23    25    5
        4.5   4.5   4.5

Bead    P
        0.5
        23    25    25   25
        4.5   4.5   4.5  4.5

Bead    S
        0.5
        23    25   25    25    25
        4.5   4.5  4.5   4.5   4.5

Bond    E  E   128  0.5
Bond    B  B   128  0.5
Bond    E  B   128  0.5
Bond    S  B   128  0.5
Bond    S  S   128  0.5
Bond    P  B   128  0.5
Bond    P  P   128  0.5

BondPair B B B   5.0  0.0

Polymer Water   0.9985  " (W) "
Polymer Rod     0.0015  " (E E (* E) (* E) B B B B B B B B (* (S S)) (* (S S)) B B B B B B B B (* (P P)) (* (P P))
B B B B B B B B (* (S S))(* (S S)) B B B B B B B E (* E) (* E) E)  "

Box        30  30  30            1 1 1
Density     3
Temp        1
RNGSeed     -22209
Lambda      0.5
Step        0.02
Time        100000
SamplePeriod        250
AnalysisPeriod      100000
DensityPeriod       100000
DisplayPeriod       50000
RestartPeriod       100000
Grid  1 1 1

Command ToggleBeadDisplay              1   W
Command SetCurrentStateCamera          1   0.5 -1.5 0.5  0.5 0.5 0.5
Command SetCurrentStateDefaultFormat   1   Paraview
```

input file

```
dpd

Title        " protein liquid liquid phase separation "
Date 18/11/19
Comment    " formation of intrinsically disordered protein liquid droplet "

State    random

Bead    W
        0.5
        25
        4.5

Bead    E
        0.5
        23    5
        4.5   4.5

Bead    B
        0.5
        25    25    25
        4.5   4.5   4.5

Bead    S
        0.5
        23    25    25   25
        4.5   4.5   4.5  4.5

Bead    P
        0.5
        23    5    25    25    5
        4.5   4.5  4.5   4.5   4.5

Bond    E  E   128  0.5
Bond    B  B   128  0.5
Bond    E  B   128  0.5
Bond    S  B   128  0.5
Bond    S  S   128  0.5
Bond    P  B   128  0.5
Bond    P  P   128  0.5

BondPair B B B   5.0  0.0

Polymer Water   0.998  " (W) "
Polymer Rod     0.002  " (E E (* E) (* E) B B B B B B B B (* (S S)) (* (S S)) B B B B B B B P (* P) (* P) P) "

Box         25  25  25            1 1 1
Density     3
Temp        1
RNGSeed     -673010
Lambda      0.5
Step        0.02
Time        100000
SamplePeriod    250
AnalysisPeriod  100000
DensityPeriod   100000
DisplayPeriod   5000
RestartPeriod   100000
Grid  1 1 1

Command ToggleBeadDisplay              1   W
Command SetCurrentStateCamera          1   0.5 -1.5 0.5  0.5 0.5 0.5
Command SetCurrentStateDefaultFormat   1   Paraview
Command SetDPDBeadConsInt              63000   S E 5
Command SetDPDBeadConsInt              63000   P S 5
Command SetDPDBeadConsInt              63000   S S 5
```