Coacervation of 10-mer polyU and RRASLRRASL-peptide

Abstract

Over the past two decades molecular dynamics has been used to simulate a vast number of different biomolecular systems. In this report molecular dynamics is used to replicate wet lab experiments from an experimental paper about liquid-liquid phase separation and formation of coacervates. Liquid-liquid phase separation has been hypothesized to be important in the formation of coacervates because they provide stable compartmentalization without the need of a membrane. In the experimental paper is studied the formation of coacervates through electrostatic interaction between anionic polyUs and cationic peptides. They observed a critical coacervation concentration of 250 μM, this means that coacervation was not observed when lower peptide concentrations were used. They also found a critical salt concentration of 100 mM, this means that coacervation was not observed when higher salt concentrations were used. In our simulations we found that peptides aggregate in a range from 250 μM to 1250 μM peptide concentration but that coacervation still took place at a 100 mM salt concentration.

Introduction

Liquid-liquid phase separation has been a topic of interest for many years now in science. One of the first articles reporting about this phenomenon dates back all the way to the 1930s [1]. They suggested that this phase separation existing of macromolecules was the first step in the origin of life. It is believed that this was important in the development of protocells, which are proposed to be crucial for the origin of life. Although this is a reason why a lot of research has been done on this topic over the years, this research has also known some quiet periods. But now it has found renewed interest. In 2009 research on this topic was reawakened by Brangwynne et al., they suggested that the formation of coacervates by liquid-liquid phase separation plays a role in the development of membraneless organelles [2]. They propose that these organelles are similar to coacervate droplets and for these coacervate droplets it has been shown that they can be formed by a RNA/peptide complex. The idea behind this is that within these droplets the concentration of RNA and peptides is much higher than when these molecules would be solvated in the solution of the cytoplasm, which means the chances of a reaction between such molecules is much higher. Since then, in the past couple of years papers have been published showing more new insights in the field of liquid-liquid phase separation.

In this project we tried to reproduce some of the experimental findings shown in an experimental article by Aumiller et al, using Molecular Dynamics [3]. The system that was studied in this article was a phosphorylation-mediated RNA/peptide complex coacervation model. In this experimental study it has been shown that the assembly and disassembly of coacervate droplets can be controlled by phosphorylation and dephosphorylation of the peptide. In this case they used polyuracyl as a RNA molecule and a double sequence of RRASL as peptide. The formation of coacervate droplets is reported when phosphatase LPP is added and the disassembly of these coacervates when PKA, a kinase, is added to the solution. In this project we tried to reproduce one of the experiments performed in this study. We studied solutions with each a different peptide and polyU concentration and with a different salt concentration in order to observe under which conditions formation of coacervates takes place

and under which conditions it does not. In total five different systems were generated for this project.

Methods

The all-atom representation of the peptide sequence RRASLRRASL was built using VMD software and the Coarse-grain model of the peptide was generated using the martinize.py script [4]. Coarse-grain molecular dynamics simulations were conducted using the Martini v.3.0.b.3.2 force field using the GROMACS 2018.1 software [5,6]. Martini is a Coarse-grain force field which is used for producing simulations of biomolecular systems. The model started by using four-to-one mapping, which means that on average four heavy atoms and their associated hydrogens are represented by one interaction center [7]. In Martini 3 three-to-one and even two-to-one mapping can be used if this better represents a molecule or certain parts of a molecule. There are four different types of interaction and each type has different subtypes, which allows for an accurate representation of the chemical nature of the underlying atomic structure.

In order to investigate the effect of the salt and peptide concentration on the formation of coacervates, several systems were investigated in which those two parameters were changed. To study the effect of peptide concentration on the formation of coacervates we changed the peptide concentration while keeping the salt concentration constant at 50 mM. For this purpose we prepared systems with 4 (250 µM), 16 (1000 µM), 20 (1250 µM) and 50 (3125 µM) peptide/polyU into a 30x30x30 nm^3 periodic box with a 1:1 ratio of peptide:polyU. Moreover, we also studied the effect of salt concentration by preparing another system with 16 peptide/polyU into a 30x30x30 nm^3 periodic box with a 1:1 ratio of peptide:polyU but with a salt concentration of 100 mM. The four systems with 50 mM salt concentration are used to observe at what concentration of peptide/polyU coacervation takes place. These systems were chosen based on the experimental article where a critical coacervation concentration of 250 µM was observed for peptides. In addition, another system with 16 peptides was built based on the experimental findings where they report no formation of coacervates in solutions with salt concentrations of 100mM and above.

The following formulas were used to calculate the number of peptide and salt molecules that needed to be added to the five systems to get the correct concentrations:

$n = V*C$
$N = n*NA$

System setup

As described earlier, five different systems were created where the peptide/polyU concentration was increased. In particular, we built systems with 250 µM, 1000 µM, 1250 µM, 3150 µM where 4, 16, 20 and 50 peptides/polyU were added respectively in order to reach these concentrations. The systems got solvated with normal Martini water beads and for both systems enough minimization steps were run (10 femtoseconds per step) until the results were satisfactory. Energy minimization was used to make sure the system has no steric clashes or inappropriate geometry.

After this step the ions were added to the systems and because life does not exist in a net charge, the systems need to be electroneutral. Na or Cl counterions were added in order to make the systems neutral. Again, the systems were minimized (10fs time step) and equilibrated(10ns long with a 10fs time step). These relatively small timesteps are used because the polyU seems to be unstable and when larger timesteps are used, the system seems to crash more. Equilibration was done to relax the system and get the molecules in low-energy configurations. For equilibration Berendsen pressure coupling was used with reference temperature 298K and reference pressure 1.0. After equilibration, first a production MD is run for 1 μs with a 20fs time step on the local computers. For the production MD run Parrinello-Rahman pressure coupling was used with reference temperature 298K and reference pressure 1.0 [8]. First, we ran the system first for 1 μs because generally systems give the most errors and then crash in the beginning and after that they are more stable. After completion of the 1 μs run, a test production MD was run in the peregrine cluster for 30 minutes to check if any errors occurred. If so, these errors could be managed before the long production MD was run. After all the errors were fixed a production MD for 4.2 μs was run using the peregrine cluster. The production MD run was done to collect data about how the system behaves under the set conditions for a certain amount of time.

Analysis

Analysis was done on the collected data to confirm that the systems are converged and aggregation of peptides/polyU was observed. This was done in two different ways. First, it was tested if the systems are converged or not. We did this by calculating the Coulombic and LJ energies of three different interactions by using the mdrun rerun gromacs tool, the protein-protein, protein-RNA and RNA-RNA interactions. This information was then plotted. A ndx file was created to separate the beads of the peptide, polyuracyl, sodium and chloride into four groups. Then we rerun the trajectory and calculate for all these groups the Coulombic and Lennard Jones contributions to the energy.

We used the radial distribution function (RDF) to confirm that we had peptides-polyU or polyU-polyU aggregation. RDF defines the probability of finding a bead at a certain distance r from another bead. The idea behind this is to calculate the RDF of, for example, two beads and check if the results confirm the expected distance between the molecules or not. The RDF can be calculated by using the gmx rdf tool, this tool calculates the radial distribution functions of one set of reference points to one or more sets of reference points. Because the radius of a normal bead is approximately 0.235 nm and RDF uses the center of a bead, we expected to see a peak in the graph at around r is 0.45 when a bead of the peptide and the polyU are posted against each other. In the case where a polyU splits two beads of the peptide we expected to see a peak at r is around 0.9.

Results

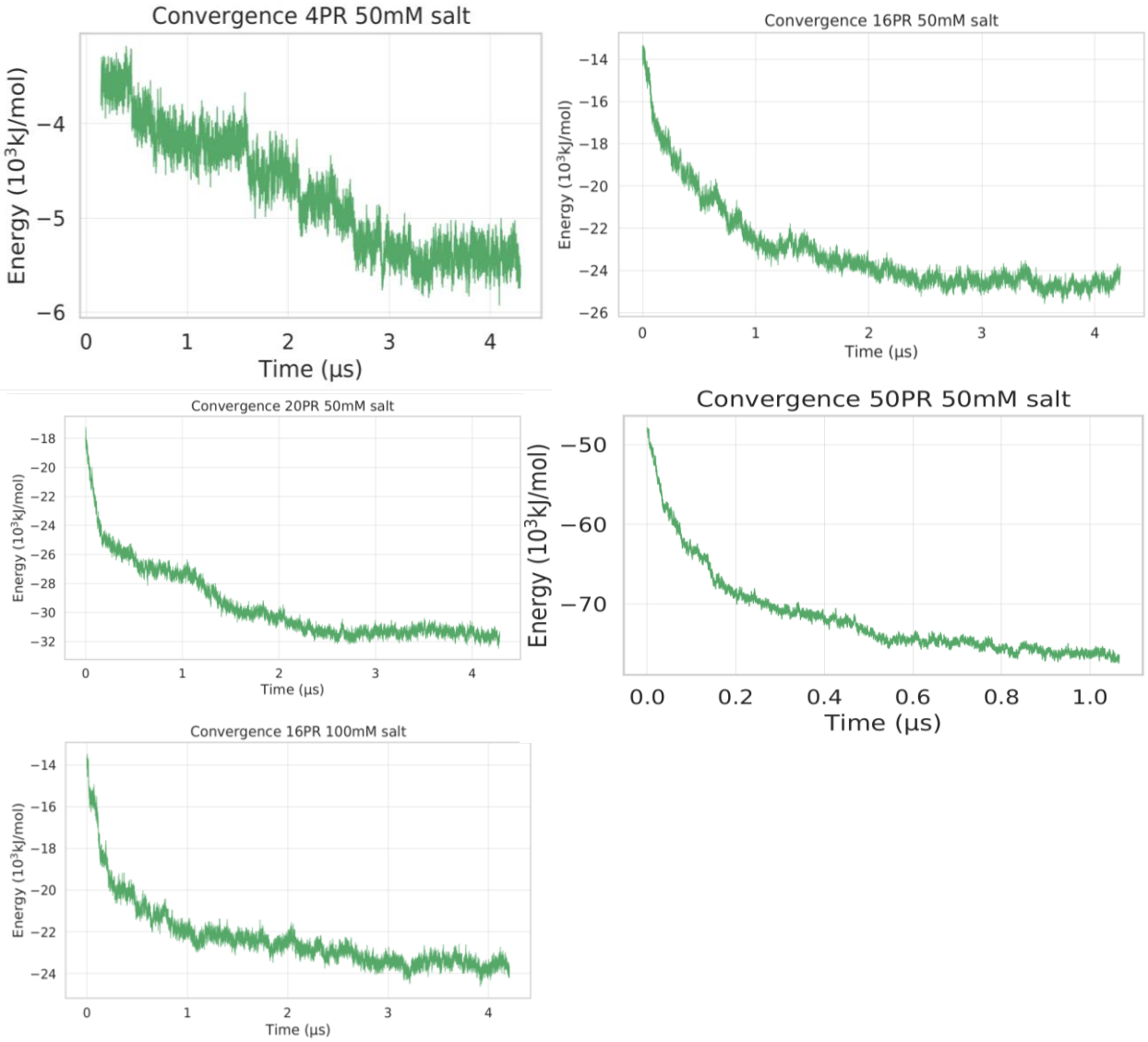

Figure 1. Convergence graphs in different systems.

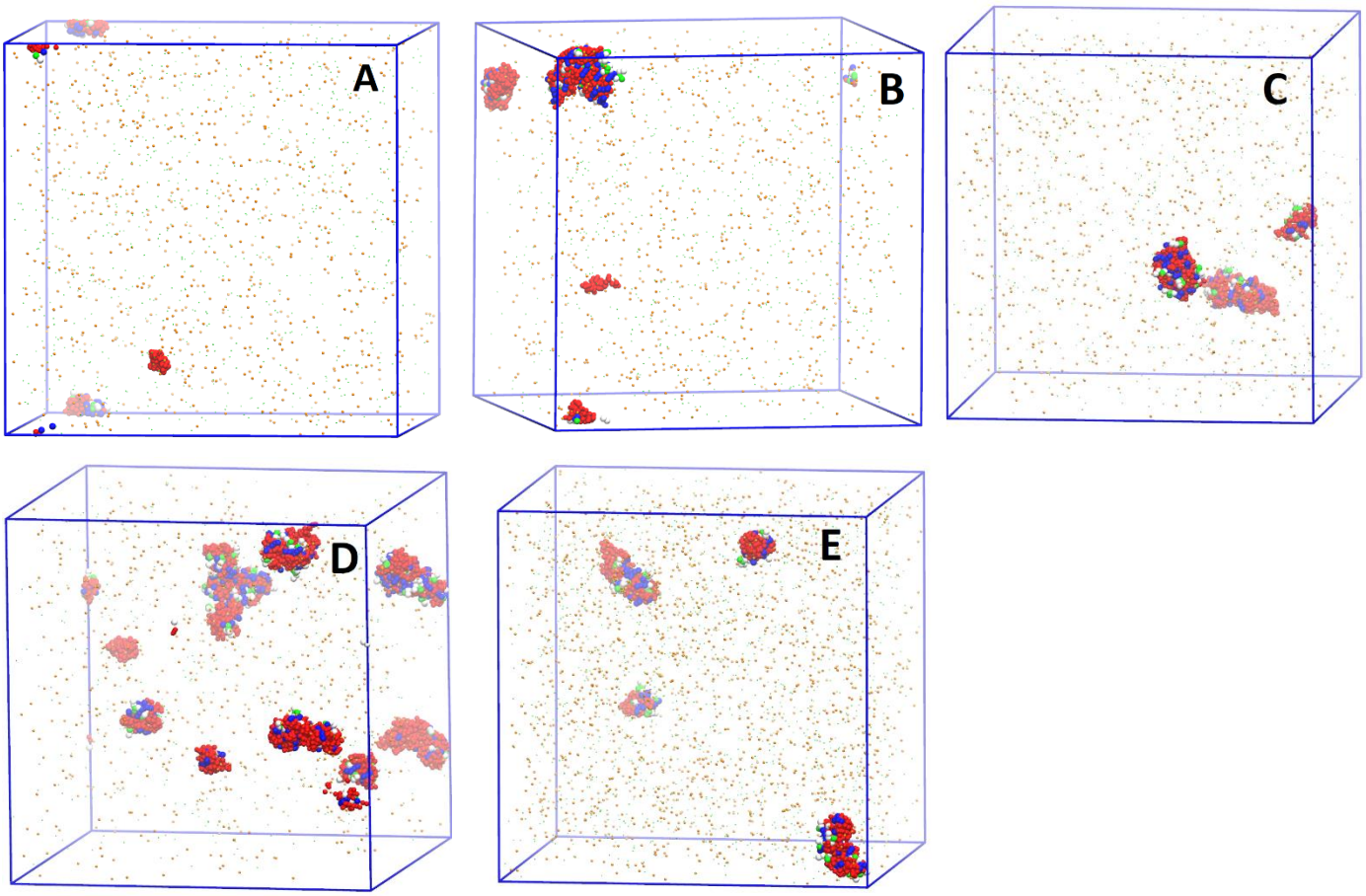

Figure 2. Snapshots of the final configurations. 2A) 4 peptide/polyU 50 mM salt concentration after 4.2 µs. 2B) 16 peptide/polyU 50 mM salt concentration after 4.2 µs. 2C) 20 peptide/polyU 50 mM salt concentration after 4.2 µs. 2D) 50 peptide/polyU 50 mM salt concentration after 4.2 µs. 2E) 16 peptide/polyU 100 mM salt concentration after 1 µs.

When the analysis started, all the systems had run for 4.2 µs, except for the system with 50 peptides. This system had only run for a little over 1 µs because the system had more molecules which made it more complicated to simulate. We checked if the systems were converged at this stage. After the gmx mdrun rerun tool was used the convergence of every system was plotted (figure 1). For the systems that ran for 4.2 µs, the system's energy of the systems with 16 and 20 peptides seemed to converge after 2.5 µs. This means that these systems probably were converged and that their data was usable for further analysis. For the systems with 4 peptides the energy did not seem to be fully stabilized but that was to be expected because the lower concentration causes the peptides to need to cover more distance before they can interact with each other. For the system that only ran for 1 µs, it seemed that the system's energy stabilized after 1 µs but this period is too short to say with certainty that the system is converged.

Aggregation of peptides/polyU

After observing the systems after 4.2 µs, all of them showed the formation of dimers existing of a peptide and a polyU (figures 2-6). It was also observed in all systems that in some cases these dimers interacted with multiple other dimers to form trimers, tetramers or even larger

networks. When observing the systems with 50 mM salt concentration, it seemed that when the peptide concentration increased, the larger the networks of peptides and polyUs became. This was probably also partly due to the fact that when the peptide concentration is increased, it is more likely for the peptides and polyUs to meet and interact with each other. When we compared both of the systems with 16 peptides, it seems that the difference in salt concentration did not make a difference in the formation of networks of peptides and polyUs for the concentrations that were used in this project. Interestingly, the systems also showed multiple polyUs interacting with each other without a peptide being involved. There also were single peptides and polyUs observed in multiple systems.

Peptide-polyU RDFs

To quantify these observations RDF was used to check if the molecules were indeed close enough to each other to interact. After the RDFs were calculated and the graphs were plotted they showed two peaks, one at around 0.45 nm and an unexpected peak at around 0.3 nm (figure 7). The peak around 0.45 nm confirmed our observed dimers. However, the peak at 0.3 nm and the fact that the radial distributions function numbers are relatively high was most likely due to something else. This peak at around 0.3 nm and these high numbers were due to the fact that when calculating the distances between the backbone beads of the polyU and peptides, also the distances between each peptide/polyUs and itself were calculated and used for plotting. In the peptide, backbone beads next to each other show overlap as can also be seen in the snapshot from VMD (figure 8). This means that the distance between the center of mass of these beads is smaller compared to the distance between the centre of mass of beads from separate molecules which do not overlap and this explains the peak at around 0.3 nm. This peak was removed by removing the calculated RDFs between the same peptides when plotting the graphs (figure 9 & 10). When inspecting the peptide-polyU RDFs for the systems with 50 mM salt concentration (figure 9), it seemed that after the peak at 0.45 nm, the lines declined until they reached a value of 1 at 2.5 nm. However, the system with 4 peptides also showed a peak at around 0.9 nm which is probably caused by two dimers located at that distance from each other. Moreover, there was a big difference in the RDF values between the systems with 4 peptides and the systems with 16 and 20 peptides. This is probably due to the fact that in a system with 4 peptides, each peptides has a greater impact on the calculations of the RDFs because it accounts for 25% of the total peptides whereas in a system with 16 or 20 peptides 1 peptide only makes up for 6.75% and 5% of the total peptides respectively. When inspecting and comparing both systems with 16 peptides (figure 10), there seems to be no big difference in formation of dimers and other networks between a 50 mM salt concentration and a 100 mM

salt concentration. The graphs are very similar, they show a peak at around 0.5 nm and then decline.

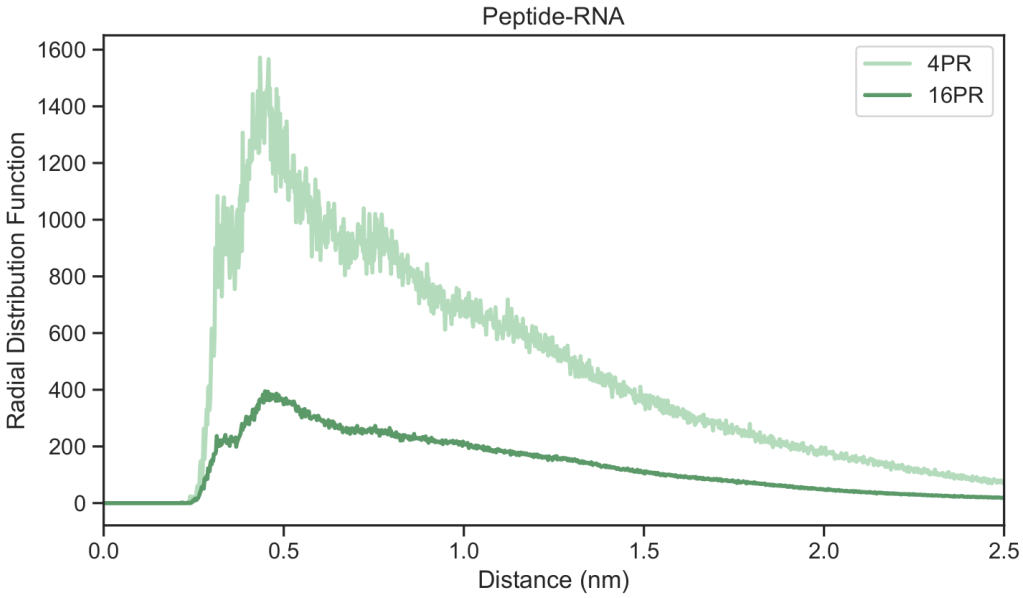

Figure 7. Plotted RDFs of the systems with 4 and 16 peptides and 50 mM salt concentration.

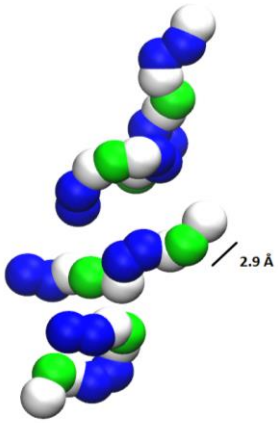

Figure 8. VMD snapshot showing only backbone beads.

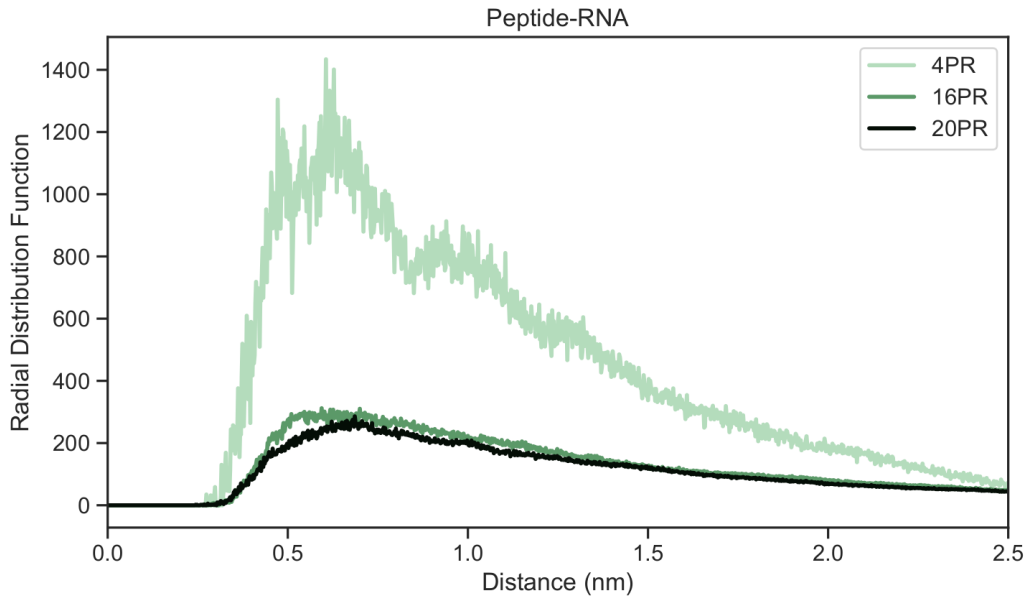

Figure 9. Plotted RDFs between peptides and polyUs of the systems with 50 mM salt concentration.

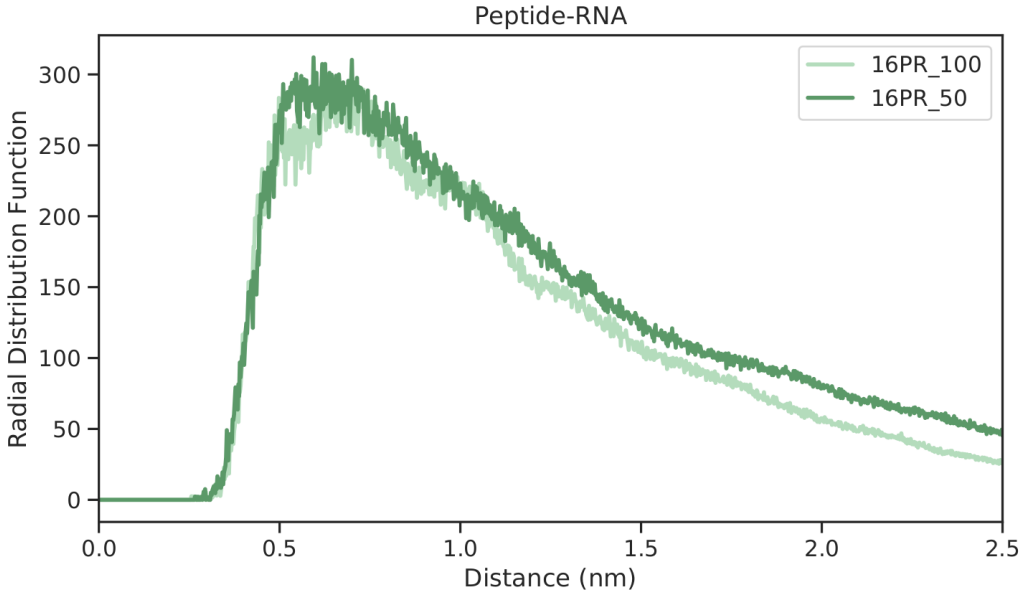

Figure 10. Plotted RDFs between peptides and polyUs of the system with 16 peptides for both 50 mM and 100 mM salt concentration.

PolyU-polyU RDFs

We also calculated and plotted the radial distribution functions between the polyUs only to confirm the formation of the larger networks that were observed (figure 11 &12). When inspecting these RDFs for the systems with 50 mMsalt concentration, the RDFs between the separate polyUs showed a peak at 0.5 nm and at around 0.9 nm. The peak at 0.5 nm was due to the fact that we calculated the RDFs between each RNA and itself as well. In this case the value was 0.5 instead of 0.3, as seen with the RDFs between peptides and RNA, because the

backbones of the RNA do not show overlap. The peak at 0.9 nm confirmed the formation of polyU-peptide-polyU networks that were observed in our snapshots. The peak at 0.9 nm for systems with 16 and 20 peptides was not clearly visible. This was probably due to the fact that the molecules that are part of a larger network were smaller in numbers compared to the RDFs of the RNAs that were compared with themselves. When inspecting figure 10, here it seemed as well that the salt concentration did not seem to influence the formation of networks much because the graphs are once again very similar. Here, the peak at around 0.9 nm was also not clearly visible.

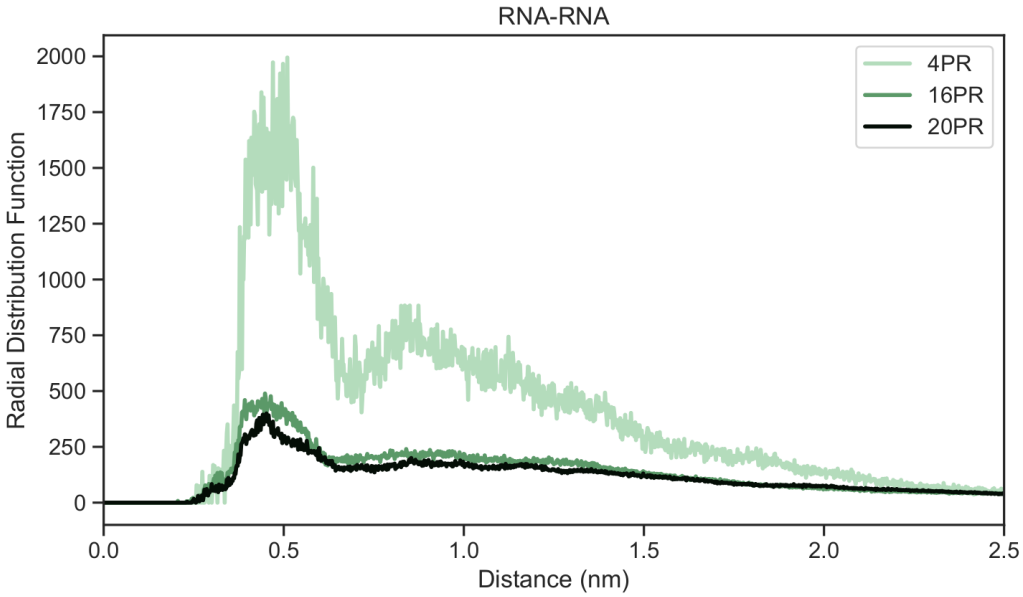

Figure 11. Plotted RDFs between the RNAs of the systems with 50 mM salt concentration.

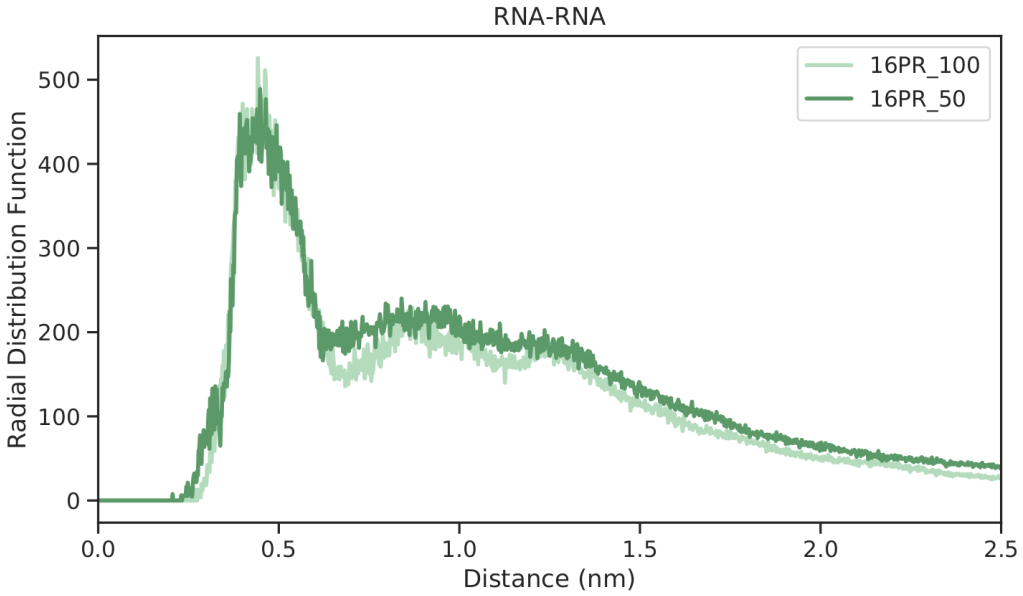

Figure 12. Plotted RDFs between the RNAs of the system with 16 peptides for both 50 mM and 100 mM salt concentration.

Discussion

The systems that were generated for this project provide some support for the formation of coacervates by peptides and polyUs. It was shown that for all concentrations of peptides and polyUs, these molecules interacted and formed dimers. On top of that, it was shown that for all of these systems also trimers and larger networks were formed. However, the data from the RDF did not clearly support this for the systems with 16 and 20 peptides. To really confirm this more types of analysis needs to be performed. One is for example the calculation of the density of the different molecules to see if certain types of molecules are aggregated or not. When we compare the results of this project with the results of Aumiller & Keating [3], we see that some of our results support their findings and conclusions but that there also are results that contradict their findings and results. In line with their findings is the critical coacervate concentration. It is shown that the lowest concentration of peptides for which coacervates were formed was 250 μM. Our system with 250 μM peptide concentration (4 peptides/polyU) also showed formation of coacervates at this concentration. However, we did not use a system with a lower peptide concentration because the number of peptide molecules would become really small. To do this perhaps a larger system must be built in which a lower concentration of peptides still contains a sufficient amount of peptide molecules. In the experimental paper the critical salt concentration was found at 100mM. For higher salt concentrations, no formation of coacervates is observed. In our case we observe the formation of networks very clearly in the snapshot of the final configuration and in the graphs of the plotted RDFs for the system with 100 mM salt concentration. In the research paper it is shown that the peptide and polyU concentrations are significantly higher in the droplets than outside the droplets. From our simulations we can assume the same, but to fully confirm this claim more analysis is needed. This can be done by density calculations, as mentioned before, but also by diffusion calculations because we expect that water for example diffuses less fast in the coacervates phase than in the aqueous phase. Another aspect that needs to be considered when looking at our data is the running time of the simulation. Our simulations ran for 4.2 μs, which was enough for the system to be converged. But after convergence the systems are still changing and ideally a simulation should be run for around 20 μs to collect the best data. All in all, this project has shown that it is definitely possible to replicate certain aspects of coacervation very well in molecular dynamics, but that it still needs improvement in some departments in order to be able to use its full potential.

Reference

1. Oparin AI, Morgulis S. 1938. The Origin of Life. New York: Macmillan. 270 pp.
2. Brangwynne CP, Eckmann CR, Courson DS, Rybarska A, Hoege C, et al. 2009. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. Science 324:1729–32
3. Aumiller, W. M., & Keating, C. D. (2015). Phosphorylation-mediated RNA/peptide complex coacervation as a model for intracellular liquid organelles. Nature Chemistry, 8(2), 129–137.
4. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. J. Mol. Graph. 1996, 14, 33 – 38.
5. Telles de Souza, P. Martini 3.0.b.3.2 Release Notes http://cgmartini.nl/index.php?start=36 2018
6. Abraham, M. J.; Murtola, T.; Schulz, R.; P´all, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations Through Multi-level Parallelism from Laptops to Supercomputers. SoftwareX 2015, 1-2, 19 – 25.
7. S.J. Marrink, H.J. Risselada, S. Yefimov, D.P. Tieleman, A.H. de Vries. The MARTINI forcefield: coarse grained model for biomolecular simulations. JPC-B, 111:7812-7824, 2007.
8. Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling Through Velocity Rescaling. J. Chem. Phys. 2007, 126, 014101.