Molecular simulation of the self-aggregation of coacervates

M.T. Greif,[1] C. Malajczuk[2] and R. Mancera[3]

Coacervates are macromolecular aggregates composed of monovalent charged polymers which form stable phase-separated microenvironments. They provide organisation and molecular crowding for early metabolism making them a promising protocellular candidate for the origin of life. Not only can they selectively absorb molecules and metal ion catalysts but they form spontaneously from polymers believed to have existed on abiotic Earth. Analysis of the mechanism of coacervate formation is vital towards comprehending the formation of coacervates on prebiotic Earth and the fundamental processes underpinning the origin of life. This study utilised atomistic molecular dynamics (MD) simulations to investigate the primary stages of coacervate formation in a representative system composed of oligolysine and ribonucleic acid molecules. The resulting oligolysine-ribonucleic acid aggregates exhibited properties consistent with experimentally characterised coacervic systems. In this paper we discuss the mechanism of early coacervate formation in terms of solvent accessibility, inter/intramolecular hydrogen bond interactions as well as the effect of RNA conformation on the rate of coacervate formation. These insights will inform further MD simulation studies to further elucidate the atomistic mechanisms of coacervate formation.

1 *School of Molecular and Life Sciences, Curtin University, Bentley, Australia; E-mail: michael.greif@student.curtin.edu.au*
2 *Supervisor, School of Pharmacy and Biomedical Sciences, Curtin University, Kent St, Bentley, Australia*
3 *Supervisor, School of Pharmacy and Biomedical Sciences, Curtin University, Kent St, Bentley, Australia*

Introduction

The origins of life on Earth has been a contested topic across human history but it was the work of Oparin (1924) which brought it firmly into the realm of natural philosophy (1). Since, it has been no less contested with three dominant theories in the field: polymerisation first (also known as 'the RNA world') whereby a random sequence of ribonucleic acid (RNA) forms the first replicase ribozyme (2); metabolism first ('the sulfur-iron world'), wherein inorganic metabolic cycles evolve (3), and; compartmentalisation first ('the lipid world'), whereby concentration is essential to polymerisation and metabolism (4). However, all schools agree that the first step was the formation of the organic molecules necessary for life.

One of the most well-known experiments in the origin of the chemicals of life was Miller's production of a wide range of amino acids from a simple mix of methane, ammonia, water and hydrogen gas exposed to UV radiation and electrical discharge; representing the Hadean, prebiotic, Earth (5). Later, an abiotic mechanism of amino acid polymerisation in wet–cool/dry–hot cycles in hydrothermal fields, as well as, metal ion catalysed, carbonyl sulphide mediated condensation at hydrothermal vents were proposed, suggesting peptides existed on prebiotic Earth (6,7).

Experiments have also provided mechanisms of abiotic synthesis of fatty acids from carbon dioxide and hydrogen gas around hydrothermal vents via a Fischer-Tropsch-type reaction (8), one among multiple abiotic mechanisms for different amphiphiles in both hydrothermal vents and fields (4).

The mechanism of abiotic formation of RNA on Earth remains challenging at all stages: from the formation of component parts to polymerisation (2,9,10). Some attribute the source of RNA to an extra-terrestrial origin during the bombardment, 10^8 kg yr^{-1}, of the Hadean Eon (7,11). While others suggest, as RNA is a precursor to DNA, proto-RNA existed during the early stages of life but has since become extinct (9,12).

Work in compartmentalisation, as the first step necessary to life, has largely focused on fatty acid layers and vesicles for their semi-permeability, such that important prebiotic chemicals could accumulate to necessary concentrations (10). They could also isolate the resulting biochemistry while providing some exchange with the environment and provide chemical and electrochemical gradients (4).

Oparin's 'prebiotic soup theory' focused on the formation of self-proliferating coacervates: oppositely charged polyelectrolytes which form large aggregates with liquid-liquid phase separation and unique structural properties. Interest in the theory was abandoned due to the perception that coacervates are unable to accumulate and retain molecules (13). However, this has been challenged, with evidence that short chain peptide-nucleotide coacervates can selectively sequester organic molecules and inorganic nanoparticles capable of catalysis (14).

Coacervates can form spontaneously where anionic and cationic polymers exist, making them an attractive candidate for compartmentalisation in protocells (1,15). Their ability to selectively absorb molecules and metal ion catalysts presents a crowded system with the

potential to contain reagents and catalysts (15–17). In this way, coacervates are a conceivable protocellular candidate for the origins of life.

Coacervate properties, such as phase behaviour and structure, are predominantly determined by polymer sequence, polymer length, polymer asymmetry, polymer concentration, chain flexibility, ionic strength, ionic identity, pH, and temperature (18,19).

Mann and co-workers have recently reinvigorated interest in the presence of coacervates in the formation of life with a hybrid protocell model of a lipid encapsulated coacervate. Their work suggests that coacervates provide the structured, molecularly crowded and enriched environment associated with modern organisation and metabolism (15). Using three representative pre-biotic molecules - oligolysine (olys), RNA and oleate - Tang et al. (2014) synthesised oleate-coated coacervate droplets that exhibit unique molecular uptake/exclusion properties.

Understanding the initial mechanisms of formation is vital to comprehending the formation of coacervates on prebiotic Earth and the fundamental processes necessary at the origins of life. Unfortunately, studying the structure and dynamics of coacervate formation at reasonable time-scales and resolutions is not possible using current experimental techniques. High resolution spectroscopic techniques such as synchrotron light time-resolved small-angle X-ray and neutron scattering have dead and exposure times on the milisecond scale, which are too slow to observe mechanisms of formation (20–24).

Molecular dynamics (MD) simulation is the study of molecular and atomic interactions, across time scales of femtoseconds up to tens of microseconds, by the use of particle-based, inter-atomic models (force fields). MD simulations can be used to accurately describe and elucidate the molecular interactions and mechanisms of aggregation for coacervates at an atomic time and length scale. Using a combination of field-theoretic and MD coarse-grained simulations, Danielsen et al. (2019) described the relationship between charge pattern and liquid-liquid phase behaviour, finding correlation between coacervate electrostatic strength and increasing polymer length (25). Lytle et al (2018) used MD simulation to describe diffusion properties in coacervate systems at varying ionic strengths and the effect on interfacial tension by representing polymers as sequences of monovalent charged beads, characterising the relationship between ionic strength and decreasing coacervatic stability (19). Finally, Bos and Sprakel investigated the ratio of anionic and cationic polymer lengths on coacervate kinetics through a Langevin dynamics, coarse-grained simulation, finding a relationship between electrolyte ratio and coacervate stability (18). However, these simulation studies do not describe coacervate formation on an atomistic level.

The present work aims to build on Tang et al.’s (2014) findings by investigating, at an atomistic level, the mechanism of coacervate self-assembly of oligolysine-RNA coacervates using MD simulations. In future this could be used to better understand molecular uptake/exclusion properties of the hybrid protocell model and the mechanism of lipid encapsulation.

Methodology

The work of Tang et al. (2014) used oligolysine as a hydrobromide salt and commercial RNA (Sigma Aldrich) is a salt extracted from yeast (Supplementary Information). The RNA salt was presumed to be coordinated with magnesium ions as it originated from a bioligical system. Chlorine and Magnesium ions were used to neutralise the charged molecules. The pH of the experimental conditions was approximately 7.0, whereby the oligolysine was fully charged and RNA had a charged phosphate backbone (27).

As mechanisms to produce long polymers are not thought to exist in abiotic environments, the range of polymer lengths were chosen to represent conditions on probiotic Earth (5-8). The ratio and length of oligolysine and RNA was chosen on the specifications of molecular weight ranges of the products used by Tang et al. (2014). No distribution or further information was provided by the manufacturer or supplier (Supporting Information). As such, peptide and RNA strand lengths were chosen within the range of molecular weights: oligolysine: 16 units, RNA: 59 units. The unit ratio of lysine to RNA were specified by Tang et al. (2014) as 4:1, so the starting system included 30 x 16 unit oligolysine, 2 x 59 unit RNA. This system was limited to 32 molecules plus explicit water and ion molecules (totalling approximately 1,400,000 atoms).

The AMBER99sb-ildnp force field (26) was chosen to describe oligolysine. Batys et al. (2020) used experimental date to validate α-helix content and β-sheet fraction of a 15-unit oligolysine system predicted using this force field (27). The Amber ROC-RNA force field (28) was chosen to describe RNA as it is compatible with AMBER99sb-ildnp and is considered to accurately represent RNA folding and interactions (29). Water molecules were described using the TIP3P water model (30), consistent with the parameterisation of both AMBER99sb-ildnp and ROC-RNA force field parameter sets. MD simulations were conducted using Gromacs 2018.3 (31).

Simulations were conducted using the Nose-Hoover thermostat (32) (τT = 0.1 ps) and the Parrinello-Rahman (33) barostat (τp = 2.0 ps) at 298 K and a pressure of 1.0 atmosphere for consistency with the experiments of Tang et al.

The initial conformation of oligolysine was predicted by iTASSER (34). Temperature replica exchange MD (T-REMD) was then conducted to obtain a representative set of conformations in a simulation totalling 16 μs across 59 replicas distributed between 298 and 410 K. Clustering analysis in Gromacs identified 367 unique conformations from 1606 frames using a 0.15 nm RMSD cut-off. All structures were inspected visually to confirm that the predictions represented oligolysine at that pH and temperature: elongated with low levels of beta turns (27). Thirty representative structures were proportionally selected from the clusters to be used for the starting configurations of oligolysine in the system.

A RNA sequence (CJ999184) was selected from the NCBI database (35) obtained from budding yeast mRNA and then ‘degraded’ at both ends to 59 units. This procedure emulates

Sigma Aldrich's RNA extracted from lysate with unknown solvent (Supplementary Information). The mRNA fits the reported length of the RNA, and the mRNA likely degraded partially during extraction and transport.

Two structures were chosen for two separate simulations: one compact RNA strand and one extended RNA (exRNA) strand. The compact (structured) RNA conformation was predicted using the SimRNAweb server (36), resulting in a condensed structure. A second fully extended RNA conformation was constructed using Discovery Studio (Dassault Systèmes) (37). This conformation was solvated, neutralised and then energy minimised using an in-house script (38) which followed a multi-step methodology of steepest descents and Polak-Ribiere conjugate gradients, prior to a 0.5 ns unrestrained equilibration run.

The first system was built by sequentially and randomly placing two structured RNA strands and oligolysine in 30 representative conformations with a minimum surface-to-surface distance of 0.105 nm away from each other, which allowed for molecules to be placed within 500 attempts in Gromacs module gmx insert-molecules. This resulted in a 5,516 nm3 simulation cell with a dodecahedral periodic boundary. The system was solvated and the total system charge was neutralised by addition of 58 Mg2+ and 480 Cl- using the Gromacs module gmx genion.The system energy was minimised using a multi-step methodology of steepest descents and Polak-Ribière conjugate gradients over 350,000 steps. The system was equilibrated for 1.3 ns until the total energy of the system stabilised at time steps of 1 and 2 fs and then simulated for 200 ns.

A second system was built as per the previous approach; however, in this system the structured RNA conformations were substituted for the exRNA strands and the initial simulation cell size was specified as the length of the RNA plus 1 nm at both termini. This resulted in a 13,591 nm3 box with a dodecahedral periodic boundary. The system energy was minimised (38) and equilibrated at steps of 0.5, 1, and 2 fs for 1.3 ns, at which point the total energy of the system stabilised.

The concentration of oligolysine in the structured RNA simulation was 9.1 mM and the concentration of RNA was 0.60 mM. For the extended RNA simulation, the concentration of oligolysine was 3.7 mM and the concentration of RNA was 0.24 mM. These values are both within the range of concentrations used by Tang et al. (2014): oligolysine 0.5 – 2.0 kDa at 5 mM and RNA 3 – 80 kDa at 5 mM. It is not known how the concentration was determined due to the range in molecular weights and a lack of information provided by the manufacturer (Supplementary Information).

The systems were benchmarked (Supporting Information) against the number of nodes and the production run used a conservative number of nodes, before linear divergence became significant, to minimise compute waste (39).

Trajectory files were reduced to one frame per ns. Hydrogen bonds were defined using geometric criteria: 30 deg and a maximum donor-acceptor distance of 0.35 nm. Solvent accessible surface area (SASA) analysis was performed using the Gromacs analysis suite (40,41). The periodic boundary condition ‘no jump’ was used for trajectory preparation prior to all analysis (41). Python 3.7 (42) with MDAnalysis 1.0.0 (43) were used to plot the resulting data implementing a moving average of 10 frames (38).

Results and Discussion

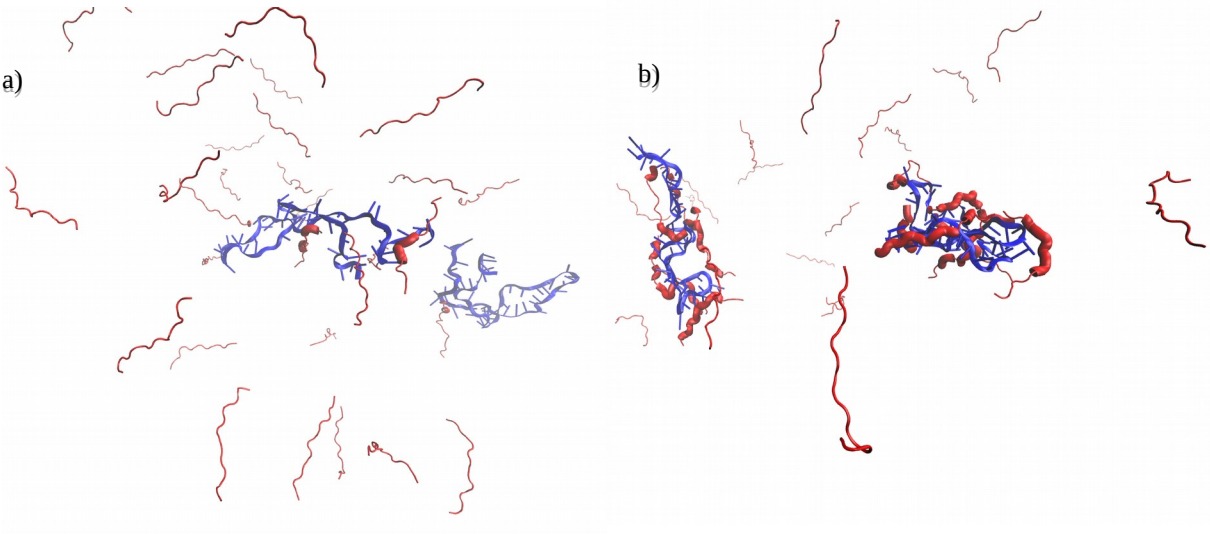


Figure 1. Visualisation of molecular dynamics simulation of initial coacervate formation at **a)** 0 ns and **b)** 200 ns (blue – RNA, red – oligolysine, thick red – oligolisine within 5 Å of RNA).

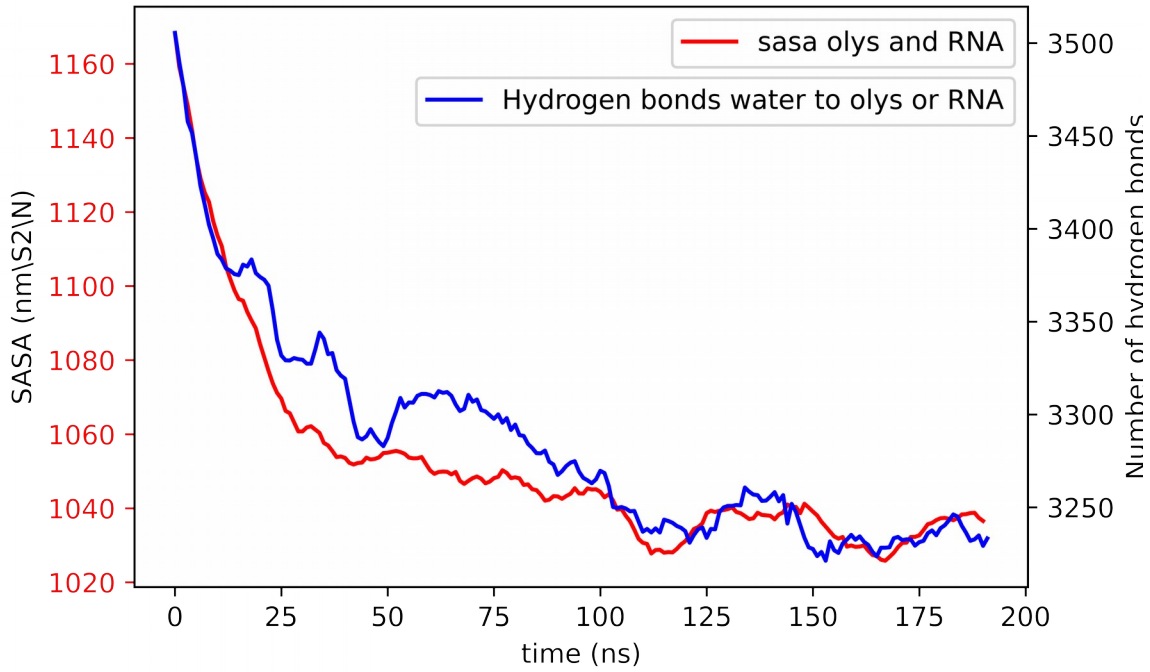


Figure 2. Solvent accessible surface area (SASA) and hydrogen bonding analysis of oligolysine and RNA coacervate formation

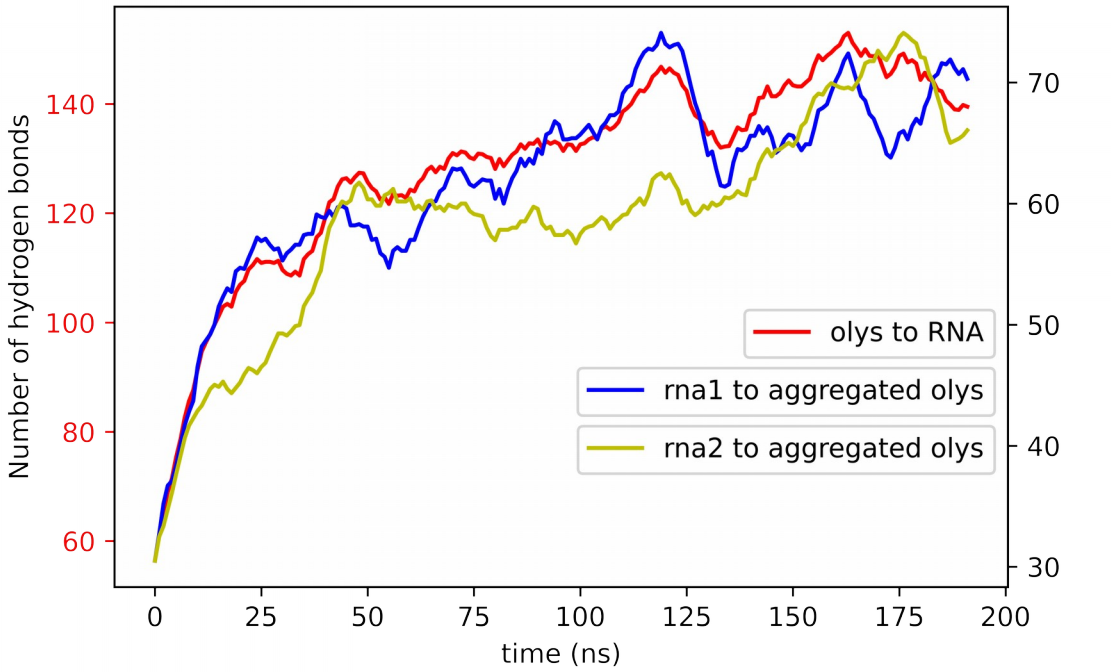


Figure 3. Hydrogen bond analysis of oligolysine to RNA in the initial formation of a coacervate

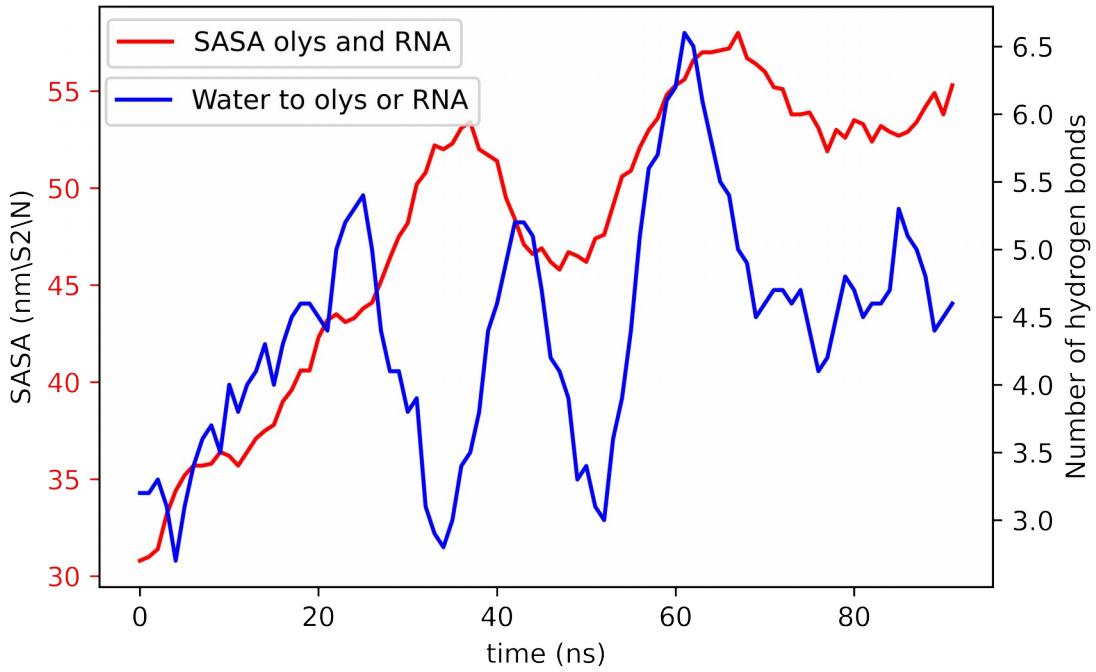
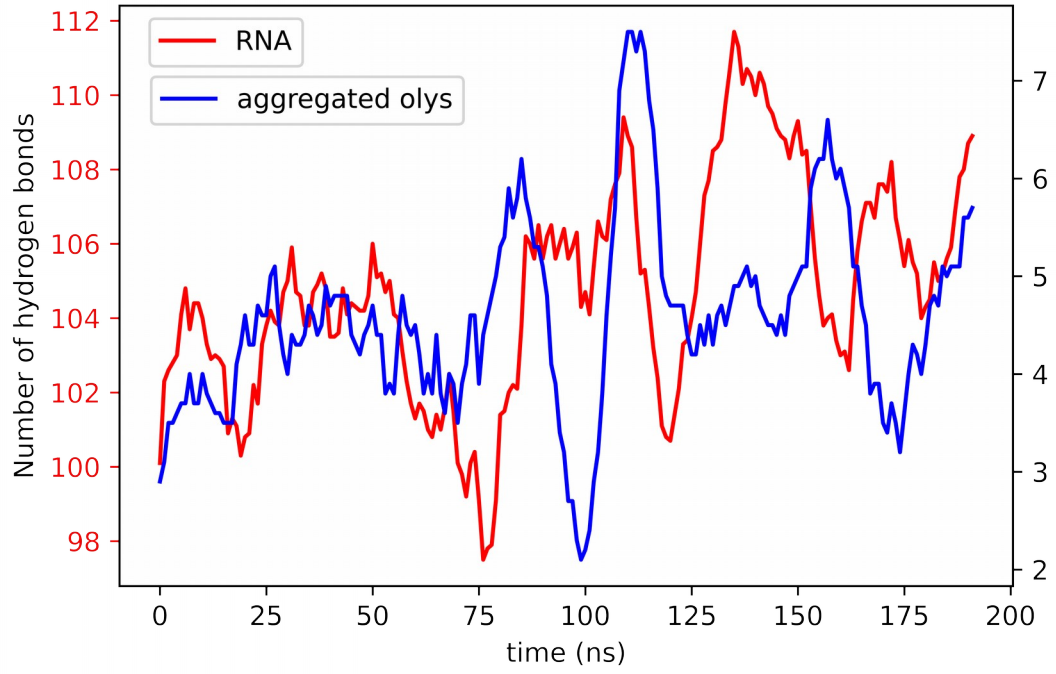


Figure 4. Solvent accessible surface area (SASA) and hydrogen bonding analysis of oligolysine and RNA coacervate formation

a)



b)

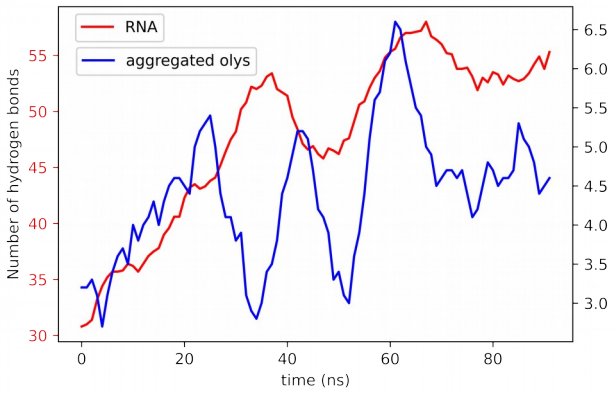


Figure 5. Intramolecular hydrogen bond analysis of oligolysine and RNA coacervate formation using **a)** structured RNA, and **b)** extended RNA

Self-aggregation of an oligolysine-RNA system was observed (Figs 1b, 2); however, a full coacervate was not observed to form in the timescales simulated in this study. Interactions between, and within, molecules were fluid and transient. The rate of aggregation was not consistent, initially occurring quickly and then slowly increasing with abrupt changes.

Hydrogen bond and SASA analysis were used to measure the formation of aggregates. While coacervates primarily rely on ionic interactions for aggregation, hydrogen bonds were used as a proxy to analyse solvent accessibility. As the molecules begin to aggregate, their intermolecular contacts are accompanied by a reduction in contacts with water. Consequently, solvent interactions have an inverse relationship with aggregation and can be used to measure aggregation.

The second system, using extended RNA, was not representative of RNA conformations and resulted in an unrealistic simulation. Differences in SASA and hydrogen bonding were not observed in exRNA (Fig 4). This was due to oligolysine coordinating to the RNA immediately, without sampling of conformations. Use of the extended conformation also led to significantly more oligolysine molecules interacting with each RNA: 6-8 interacting with each RNA within 100 ns on average. However, this was not represented in changes in SASA or intermolecular hydrogen bonding (Fig 2, 4), suggesting that it did not form a more condensed aggregate. This system may ultimately be useful as a negative control but the insufficient simulation time scale and the use of non-representative RNA conformation precluded its further consideration.

In the first simulation system containing structured RNA, aggregation occurred quickly within the first 30 ns and then continued slowly across the next 70 ns, after which there were periodic fluctuations for the remainder of the simulation (Fig 2). By 200 ns, each RNA was interacting with 5-6 oligolysines (Fig 1b) and no interactions between these two aggregates were observed. Therefore, a stable, phase-separated single coacervate had not formed.

Interactions between the individual components were analysed to better understand the trends observed. Hydrogen bonding analysis was conducted between each RNA and its interacting oligolysine. There were to RNA in the system, as such they will be referred to as RNA1 and RNA2. Molecules were determined to be interacting if they were within 10 A at 200 ns (Fig 1b). The RNA2 aggregate had a long period of stability (100 ns), between 30 and 130 ns, before a significant increase in hydrogen bonds with oligolysine occurred between 130 ns and 175 ns (Fig 3). The RNA1 aggregate had a period of stable increase of hydrogen bonds with oligolysine between 20 and 120 ns, before a period of structural fluctuations of a magnitude of 20%. It is therefore likely that the system could continue with cycles of structural fluctuations, slow and sudden growth, and temporary stability in structural fluctuations until, and after, a stable, phase-separated coacervate forms. Even then, it is likely that intermolecular hydrogen bond formation within a coacervate should be transient in nature (19–23).

Figure 5 suggests the intramolecular hydrogen bonds are not stable, with many fluctuations in the smoothed data. The molecules are in a fluid state, not forming stable conformations, even as they aggregate. This is also consistent with the semi-structured characteristics associated with coacervate components (19–23). Across the simulation time period, a small positive trend in intramolecular hydrogen bonds for RNA as well as oligolysine was

observed. This was not the case for RNA in the exRNA simulation, which had a significant positive trend (Fig 5). This was to be expected as the exRNA is not stable or realistic and it continued to approach a stable conformation, forming more intramolecular hydrogen bonds. The soft positive trend of the first simulation may be due to a similar stabilisation process or a result of changing conformations due to aggregation, whereby molecules are forced to interact with themselves. Additional simulation time may elucidate the significance of this trend.

Figure 2 suggests that the trends in SASA and hydrogen bonding analysis correlate to suggest that both are appropriate proxy measures for coacervate self-assembly. The difference in the 50 to 100 ns period emphasises the complementary nature of these two analyses: while the magnitude of SASA provides a clear delineation between the initial, fast, aggregation and the number of hydrogen bonds suggest an equilibration occurring within this period, as oligolysine and RNA gradually reduce their interactions with water. This is further emphasised by the extended RNA simulation, wherein a large amount of surface area for aggregation occurred more readily, not having to find a a lower energy conformation (Fig 4).

Conclusion and future work

This study utilised atomistic MD simulations to investigate the initial stages of coacervate self-aggregation in a representative system. The resulting aggregates exhibited properties consistent with experimentally characterised coacervate systems. During coacervate formation aggregates underwent an initial stage of fast aggregation before moving in cycles of structural fluctuations, slow and sudden growth, and temporary stability. These insights will inform further MD simulation studies to further elucidate the atomistic mechanisms of coacervate formation.

Future work is needed to further characterise the process of coacervate formation in order to provide more information on the cycles of structural fluctuations, slow and sudden growth, and temporary stability seen in the present simulations. Nonetheless, the process of coacervate formation could progress over the course of milliseconds (more than 500 times longer than the present work). This will require substantially higher compute efficiency and / or a more efficient approach. A hybrid coarse-grained MD simulation approach (44) may thus be appropriate. Alternatively, the progressive reduction in simulation cell size via the removal of water molecules would minimise superfluous water-water interactions, which largely dominate compute efforts.

In addition, exploration of the effect of pH, temperature, ionic strength, and protein-nucleic concentration could result in a different rate of aggregation and so might be explored as a means to minimise compute resources. Once a target system has been identified a large number of replicates could be run to reduce the bias of the starting configuration of the system. This approach would significantly increase resources required.

Further analyses of the existing data may consider the diffusion of ions and molecules, and the ionic interactions between them. Analysis of the radial distribution functions of polymer ionic groups to ions was attempted but the results were unclear and could not be confirmed visually (Supplementary Information). Equally, a better form of measuring aggregation could be used, such as principal component analysis, which was indeed attempted but did not produce meaningful results (38). Gromacs dihedral principal component analysis could not be completed due to time constraints. Such analysis might provide in-depth information on aggregation that could not be gleamed from analyses of SASA and hydrogen bonds. Analysis of the hydrophobic and hydrophilic SASA and of the exposure of charge, hydrophilicity, and hydrophobicity could also contribute to understanding the mechanism of aggregation.

Acknowledgements

This work was supported by resources provided by the Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia, as well as resources and services from the National Computational Infrastructure (NCI), which is supported by the Australian Government. With thanks to the Curtin Institute for Computation, Curtin Health Inovation Research Institute, the School of Molecular and Life Sciences, and the School of Pharmacy and Biomedical Science. Building of RNA structures and t-REMD sampling of oligolisine was performed by C. Malajczuk.

Supporting Information

Code

All code referenced in this paper can be found at the repository:
https://github.com/BMMG-Curtin/CoacervateMD

Benchmarking

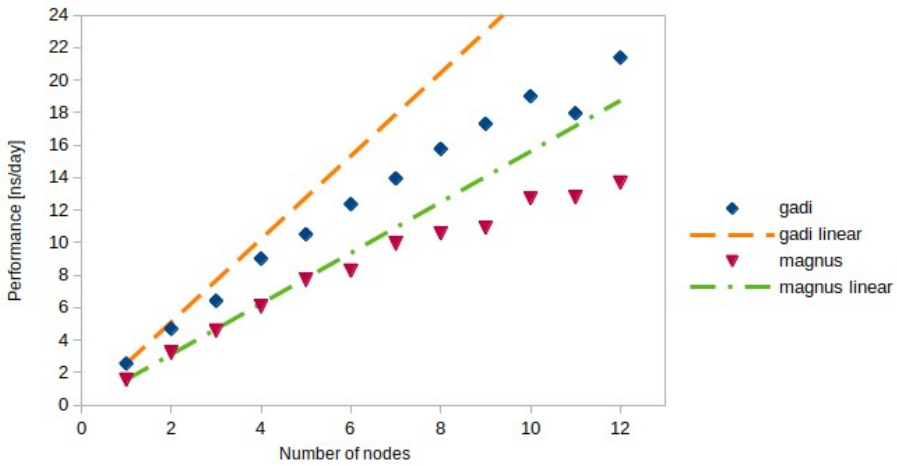


Figure S1. Performance by node of HPC's Gadi and Magnus using aa coacervate system.

Gadi has 96 virtual cores per node, Magnus has 24 physical cores per node. HPC nodes are not equal, this results in benchmarks with high, non-predictable, error. Nodes are compared, instead of cores, as a large factor in the drift from linear improvement is communication between nodes. To reduce resource waste four nodes were used on Gadi, seven nodes on Magnus (Fig S1).

Radial Distribution Function

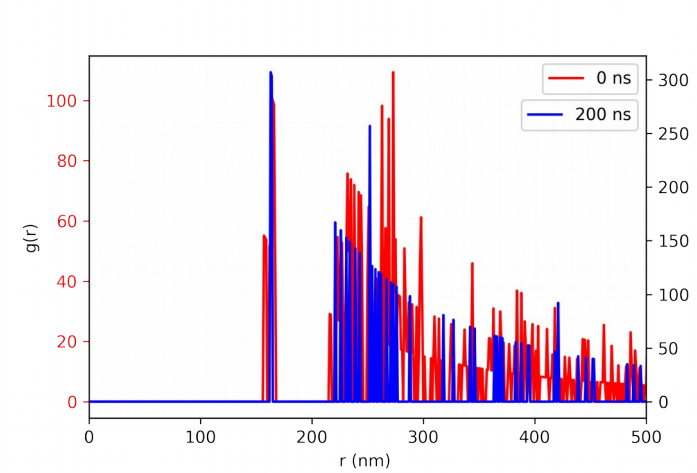


Figure S2. Radial distribution function of lysine amine to chlorine over a 5 ns period of 1ns frames.

An analysis of the radial distribution function (RDF) of both phosphate to magnesium and lysine amine to chlorine produced results inconsistent to those expected. The first coordination sphere was at 1800 Å for both groups (Fig s2); this suggests no interaction between ions and ionic groups. The shape of the RDF is also concerning; lacking the characteristic decreasing wave function curve. Visual inspection was used for confirmation but found many ion-ionic interactions within 10 Å. Gmx rdf was used over different time ranges; index groups and method was confirmed to be in line with specifications. More investigation is required to determine the source of this inconsistency.

Product information / materials

Correspondence with Merck Pty Ltd.

Product information: R6625 RNA. #C8508253.

"We are extracting RNA from yeast with certain solvents. Unfortunately detailed information about the solvent is confidential. One thing we can say is that we do not use RNase in the manufacturing process. I am afraid that we do not have the biochemical information regarding Our RNA product.

If you have any further queries, please feel free to write to us. Thank you."

Correspondence with Tang et al.

"….In addition, prior to starting our simulations we would greatly appreciate any clarification that you can possibly provide to us with regards to the materials and reagents that you used in your 2014 study.

We wanted to confirm that the polylysine that was used by you and your colleagues was poly-L-lysine.

Yes, this is poly-L-Lysine

And if so, was it purchased as a salt (poly-L-lysine hydrochloride) or as a pure solution?

Purchased as the hydrobromide salt

Along the same lines, was the RNA in a pure or salt form,

This was purchased as a solid so the salt form.

and which type of Torula yeast RNA was used for the particular study (type VI, type II-C)?

Type VI

Furthermore, notwithstanding the salt-induced micro-drop fusion and growth part of your study, do you remember which (if any) specific salt solution was used in your initial coacervate formation experiments to set an ionic strength?

yes, NaOH and HCl was used to set the pH/ ionic strength. Depending on the batch then predominantly NaOH is used. a rough ball pack estimate is that the NaOH would be around between 10^-3 and 10^-4 M. Please only use this as an

estimate.

And finally, was a buffer solution used to control for pH? This level of information will help us to construct our initial simulation systems in a way that is as consistent as possible with your previous work.

No buffer was used- only the NaOH / HCl specified above.”

References

1. Oparin AI, Synge A. The origin of life on the earth. 3nd, rev.enl. ed. ed. New York : Academic Press,; 1957. 532 p. Available from: https://www.biodiversitylibrary.org/item/23465

2. Gilbert W. Origin of life: The RNA world. Nature. 1986 Feb;319(6055):618–618.

3. Wächtershäuser G. Evolution of the first metabolic cycles. Proc Natl Acad Sci. 1990 Jan 1;87(1):200–4.

4. Segré D, Ben-Eli D, Deamer DW, Lancet D. The Lipid World. Orig Life Evol Biosph. 2001 Feb 1;31(1):119–45.

5. Miller SL. A Production of Amino Acids under Possible Primitive Earth Conditions. Sci New Ser. 1953;117(3046):528–9.

6. Forsythe JG, Yu S-S, Mamajanov I, Grover MA, Krishnamurthy R, Fernández FM, et al. Ester-Mediated Amide Bond Formation Driven by Wet–Dry Cycles: A Possible Path to Polypeptides on the Prebiotic Earth. Angew Chem Int Ed. 2015;54(34):9871–5.

7. Kitadai N, Maruyama S. Origins of building blocks of life: A review. Geosci Front. 2018 Jul 1;9(4):1117–53.

8. McCollom TM, Ritter G, Simoneit BRT. Lipid Synthesis Under Hydrothermal Conditions by Fischer-Tropsch-Type Reactions. Orig Life Evol Biosph. 1999 Mar 1;29(2):153–66.

9. Hud NV, Cafferty BJ, Krishnamurthy R, Williams LD. The Origin of RNA and “My Grandfather’s Axe.” Chem Biol. 2013 Apr 18;20(4):466–74.

10. Jowett P, Rayne R, Tomas S. Myth and fact in the origins of cellular life on Earth. Biosci Horiz Int J Stud Res. 2017 Jan 1;10.

11. Pearce BKD, Pudritz RE, Semenov DA, Henning TK. Origin of the RNA world: The fate of nucleobases in warm little ponds. Proc Natl Acad Sci. 2017 Oct 24;114(43):11327–32.

12. Robertson MP, Joyce GF. The Origins of the RNA World. Cold Spring Harb Perspect Biol. 2012 May;4(5).

13. Monnard P-A, Deamer DW. Membrane self-assembly processes: Steps toward the first cellular life. Anat Rec. 2002;268(3):196–207.

14. Koga S, Williams DS, Perriman AW, Mann S. Peptide–nucleotide microdroplets as a step towards a membrane-free protocell model. Nat Chem. 2011 Sep;3(9):720–4.

15. Dora Tang T-Y, Rohaida Che Hak C, Thompson AJ, Kuimova MK, Williams DS, Perriman AW, et al. Fatty acid membrane assembly on coacervate microdroplets as a step towards a hybrid protocell model. Nat Chem. 2014 Jun;6(6):527–33.

16. Astoricchio E, Alfano C, Rajendran L, Temussi PA, Pastore A. The Wide World of Coacervates: From the Sea to Neurodegeneration. Trends Biochem Sci. 2020 Aug 1;45(8):706–17.

17. Lu T, Spruijt E. Multiphase Complex Coacervate Droplets. J Am Chem Soc. 2020 Feb 12;142(6):2905–14.

18. Bos I, Sprakel J. Langevin Dynamics Simulations of the Exchange of Complex Coacervate Core Micelles: The Role of Nonelectrostatic Attraction and Polyelectrolyte Length. Macromolecules. 2019 Nov 26;52(22):8923–31.

19. Lytle TK, Salazar AJ, Sing CE. Interfacial properties of polymeric complex coacervates from simulation and theory. J Chem Phys. 2018 Jul 16;149(16):163315.

20. Takahashi R, Narayanan T, Sato T. Growth Kinetics of Polyelectrolyte Complexes Formed from Oppositely-Charged Homopolymers Studied by Time-Resolved Ultra-Small-Angle X-ray Scattering. J Phys Chem Lett. 2017 Feb 16;8(4):737–41.

21. Amann M, Diget JS, Lyngsø J, Pedersen JS, Narayanan T, Lund R. Kinetic Pathways for Polyelectrolyte Coacervate Micelle Formation Revealed by Time-Resolved Synchrotron SAXS. Macromolecules. 2019 Nov 12;52(21):8227–37.

22. Narayanan T, Konovalov O. Synchrotron Scattering Methods for Nanomaterials and Soft Matter Research. Materials. 2020 Jan;13(3):752.

23. Wu H, Ting J, Yu B, Jackson N, Meng S, de Pablo J, et al. Spatiotemporal Formation Kinetics of Polyelectrolyte Complex Micelles with Millisecond Resolution. 2020 Jul 27;

24. Zhao D, Wang E, Lodge TP. Hybridization of a Bimodal Distribution of Copolymer Micelles. Macromolecules. 2020 Aug 31;

25. Danielsen SPO, McCarty J, Shea J-E, Delaney KT, Fredrickson GH. Molecular design of self-coacervation phenomena in block polyampholytes. Proc Natl Acad Sci. 2019 Apr 23;116(17):8224–32.

26. Aliev AE, Kulke M, Khaneja HS, Chudasama V, Sheppard TD, Lanigan RM. Motional timescale predictions by molecular dynamics simulations: Case study using proline and hydroxyproline sidechain dynamics. Proteins Struct Funct Bioinforma. 2014;82(2):195–215.

27. Batys P, Morga M, Bonarek P, Sammalkorpi M. pH-Induced Changes in Polypeptide Conformation: Force-Field Comparison with Experimental Validation. J Phys Chem B. 2020 Apr 9;124(14):2961–72.

28. Aytenfisu AH, Spasic A, Grossfield A, Stern HA, Mathews DH. Revised RNA Dihedral Parameters for the Amber Force Field Improve RNA Molecular Dynamics. J Chem Theory Comput. 2017 Feb 14;13(2):900–15.

29. Zhao J, Kennedy SD, Berger KD, Turner DH. Nuclear Magnetic Resonance of Single-Stranded RNAs and DNAs of CAAU and UCAAUC as Benchmarks for Molecular Dynamics Simulations. J Chem Theory Comput. 2020 Mar 10;16(3):1968–84.

30. Price DJ, Brooks CL. A modified TIP3P water potential for simulation with Ewald summation. J Chem Phys. 2004 Nov 11;121(20):10096–103.

31. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX. 2015 Sep 1;1–2:19–25.

32. Evans DJ, Holian BL. The Nose–Hoover thermostat. J Chem Phys. 1985 Oct 15;83(8):4069–74.

33. Saito H, Nagao H, Nishikawa K, Kinugawa K. Molecular collective dynamics in solid para-hydrogen and ortho-deuterium: The Parrinello–Rahman-type path integral centroid molecular dynamics approach. The Journal of chemical physics. 2003 Jul 8;119(2):953-63.

34. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. Nat Methods. 2015 Jan;12(1):7–8.

35. Sayers EW, Agarwala R, Bolton EE, Brister JR, Canese K, Clark K, et al. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 2019 Jan 8;47(D1):D23–8.

36. Magnus M, Boniecki MJ, Dawson W, Bujnicki JM. SimRNAweb: a web server for RNA 3D structure modeling with optional restraints. Nucleic Acids Res. 2016 Jul 8;44(W1):W315–9.

37. Sutter J, Li J, J. Maynard A, Goupil A, Luu T, Nadassy K. New Features that Improve the Pharmacophore Tools from Accelrys. Curr Comput - Aided Drug Des. 2011 Sep 1;7(3):173–80.

38. Greif M. Molecular simulation of the self-aggregation of coacervates [Internet]. 2020. (CoacervateMD). Available from: https://github.com/BMMG-Curtin/CoacervateMD

39. Gecht M, Siggel M, Linke M, Hummer G, Köfinger J. MDBenchmark: A toolkit to optimize the performance of molecular dynamics simulations. J Chem Phys. 2020 Oct 9;153(14):144105.

40. Richmond TJ. Solvent accessible surface area and excluded volume in proteins: Analytical equations for overlapping spheres and implications for the hydrophobic effect. J Mol Biol. 1984 Sep 5;178(1):63–89.

41. gmx sasa. In: User Manual Gromacs 2018 [Internet]. University of Groningen, Royal Institute of Technology, Uppsala University; [cited 2021 Oct 11]. Available from: http://manual.gromacs.org/documentation/2018/onlinehelp/gmx-sasa.html

42. Van Rossum G, Drake F. Python reference manual. Centrum voor Wiskunde en Informatica Amsterdam; 1995.

43. Michaud Agrawal N, Denning EJ, Woolf TB, Beckstein O. MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. J Comput Chem. 2011;32(10):2319–27.

44. Rzepiela AJ, Louhivuori M, Peter C, Marrink SJ. Hybrid simulations: combining atomistic and coarse-grained force fields using virtual sites. Physical Chemistry Chemical Physics. 2011;13(22):10437-48.