

# POLITECNICO DI TORINO

Master's Degree in Physics of Complex Systems

A.a. 2022/2023



**Politecnico  
di Torino**

## Towards Understanding The Physico-Chemical Principle Of Molecular Sequestration By Coacervates Protocells

Supervisors

Dr. ir. Evan SPRUIJT

Prof. Andrea GAMBA

Merlijn van HAREN

Candidate

Riccardo SCHIROLI

October 2023



# Summary

Coacervates are micron-sized liquid droplets, formed by liquid-liquid phase separation directed by attractive and segregative molecular interactions, that have attracted considerable interest as protocells. In addition to their role in vital processes of modern-day cells, the ability of coacervates to encapsulate reactions, preserve the products against environmental fluctuations, and possibly allow for the emergence of more complex molecular structures, all suggest that they may have played a role in the formation of cells at the origins of life. A key role in protocell models, as well as in modern-day biological processes, is played by the interactions of coacervates with the surrounding environment and their ability to exchange molecules with the dilute phase. Given the complexity of interaction networks arising from multi-component systems and the large number of molecules involved, predicting client partitioning in host coacervates, and understanding the effect of sequestered molecules on phase separation is still a significant scientific challenge. In this study, we investigated the physico-chemical principles underlying coacervate assembly and sequestration of client molecules, constituted by small peptides or RNAs. We applied a combined experimental and numerical approach to unravel their behavior and analyzed the influence of ions on WGR-4 and RP3/polyU coacervates by constructing their phase diagrams. We employed Molecular Dynamics simulations using a 1BPA-3BPN coarse grained model with implicit ions and solvent description to reproduce coacervates systems. A concentration region in between 5-10 mM WGR-4 in which simulations agree with experiments has been found and results were used to take the initial steps in the study of the partition properties of small client molecules in coacervates. We developed methods to analyze confocal 3D images by extracting the partition coefficient of clients and compared them to simulations results. We highlights the interplay of host and guest in coacervates formation observing phase separation of WGR-4 coacervates enhanced by arginine peptides and small RNAs. We found that coacervates made up by small and simple components were able to sequester a wide range of prebiotically relevant molecules and that these molecules can improve their stability. These results allow for further studies on modeling coacervates in both biological research and origins of life studies.



# Acknowledgements

I would like to express my gratitude to my supervisors, Professor Andrea Gamba and Dr. ir. Evan Spruijt, for their invaluable guidance and support throughout this research. Special thanks to Merlijn van Haren for providing daily supervision, guiding me during experiments and growth in the lab environment. I extend my thanks to the entire Spruijt Lab for the inspiring environment in the laboratory, which made my time in Nijmegen truly enjoyable.

I am grateful to Professor Patrick Onck and Dr. Mark Driver for their assistance with numerical simulations work and for the time I spent in Groningen.

I am also deeply grateful to my family for their support and encouragement during my academic pursuits. Finally, I extend my thanks to my dear friend Davide, whose inspiring conversations and shared enthusiasm have significantly shaped my ideas and perspectives.

This work would not have been possible without the encouragement, mentorship, and love of all of them, and for that, I am sincerely thankful.

# Table of Contents

|  |      |
|--|------|
| <b>List of Tables</b>  | VIII |
| <b>List of Figures</b>   | IX   |
| <b>1 Introduction</b>  | 1    |
| 1.1 Introduction to the Physical Principles of Coacervates Formation . . . | 3    |
| 1.1.1 Nucleation and Growth . . . . .                                      | 4    |
| 1.1.2 The Screening Effect of Salt . . . . .                               | 5    |
| 1.1.3 Gibbs Free Energy . . . . .  | 6    |
| 1.2 Theory of Partitioning . . . . .                                       | 7    |
| <b>2 Materials and Method</b>  | 9    |
| 2.1 RP3/polyU Phase Diagram . . . . .                                      | 9    |
| 2.2 WGR-4 Phase Diagram . . . . .  | 10   |
| 2.3 Partition Coefficient of Clients in Coacervates . . . . .              | 11   |
| 2.3.1 Image Segmentation and Analysis . . . . .                            | 13   |
| <b>3 Numerical Simulations</b>   | 15   |
| 3.1 Theoretical Background . . . . .                                       | 15   |
| 3.1.1 One Bead Per Amino Acids (1BPA) . . . . .                            | 15   |
| 3.1.2 Three Beads Per Nucleotide (3BPN) . . . . .                          | 16   |
| 3.1.3 1BPA-3BPN compatibility . . . . .                                    | 17   |
| 3.1.4 Simulation Protocol . . . . .  | 17   |
| 3.2 Phase Diagrams . . . . .   | 19   |
| 3.2.1 WGR-4 . . . . .  | 19   |
| 3.2.2 RP3/polyU . . . . .  | 19   |
| 3.3 Host-Client Interactions . . . . .                                     | 19   |
| <b>4 Results and Discussion</b>  | 21   |
| 4.1 Phase Diagrams . . . . .   | 21   |
| 4.1.1 Choice of Coacervates . . . . .                                      | 21   |

|          |  |           |
|----------|--|-----------|
| 4.1.2    | WGR-4 Phase Diagram . . . . .                        | 23        |
| 4.1.3    | RP3-pU Phase Diagram . . . . .                       | 27        |
| 4.2      | Client Partitioning . . . . .                        | 31        |
| 4.2.1    | Choice of Clients . . . . .                          | 31        |
| 4.2.2    | Effect of Fluorophore Dyes on Partitioning . . . . . | 32        |
| 4.2.3    | Partitioning of Small Peptides . . . . .             | 34        |
| 4.2.4    | Partitioning of Small RNAs . . . . .                 | 37        |
| <b>5</b> | <b>Conclusions and Outlook</b>                       | <b>41</b> |
| <b>A</b> | <b>Supplementary Plots and Figures</b>               | <b>45</b> |
| A.1      | WGR-4 . . . . .                                      | 46        |
| A.2      | RP3-pU . . . . .                                     | 47        |
| A.3      | Client partitioning . . . . .                        | 48        |
| <b>B</b> | <b>Chemical Potential for 4 Components System</b>    | <b>51</b> |
|          | <b>Bibliography</b>                                  | <b>53</b> |

# List of Tables

|     |   |    |
|-----|---|----|
| 3.1 | List of molecules, concentrations and related number of beads used in client partitioning simulations . . . . . | 20 |
| A.1 | Experimental Partitioning coefficients of small molecules in WGR-4 coacervates . . . . .                        | 50 |
| A.2 | Experimental Partitioning coefficients of small molecules in RP3/polyU coacervates . . . . .                    | 50 |

# List of Figures

|      |  |    |
|------|--|----|
| 1.1  | Classification of simple and complex coacervates and their phase diagrams . . . . .                  | 3  |
| 2.1  | Schematic representation of 3D images analysis . . . . .   | 12 |
| 3.1  | Snapshots of the final frame of different steps involved in MD simulations . . . . .                 | 18 |
| 3.2  | Snapshots of the final frame of MD simulations for RP3/polyU coacervates . . . . .                   | 20 |
| 4.1  | Molecular structure of WGR-4 (WGRGRGRGWY) . . . . .  | 22 |
| 4.2  | Molecular structures of RP3 (RRASL) <sub>3</sub> and Uridine . . . . .                               | 23 |
| 4.3  | Experimental and numerical phase diagram of WGR-4 coacervates . . . . .                              | 24 |
| 4.4  | Average droplet radius vs time . . . . .   | 26 |
| 4.5  | Experimental phase diagram of RP3/polyU coacervates . . . . .  | 28 |
| 4.6  | Positive:Negative charge ratio in RP3/polyU coacervates as a function of ion concentration . . . . . | 29 |
| 4.7  | Partitioning of small peptides in WGR-4 coacervates. . . . .   | 32 |
| 4.8  | Effect of FAM on partitioning of positively charged peptides in WGR-4 coacervates . . . . .          | 34 |
| 4.9  | Results of partitioning experiments of short peptides in WGR-4 and RP3/polyU coacervates . . . . .   | 35 |
| 4.10 | Results of partitioning of short RNAs in WGR-4 and RP3/polyU coacervates . . . . .                   | 38 |
| A.1  | Molecular structures of peptides involved in partitioning experiments                                | 45 |
| A.2  | Cluster size distribution in WGR-4 coacervates simulations . . . . .                                 | 46 |
| A.3  | Experimental phase diagrams for RP3/polyU coacervates . . . . .                                      | 47 |
| A.4  | Cluster size distribution in RP3/polyU coacervates simulations . . . . .                             | 47 |
| A.5  | Size dependence of partitioning in WGR-4 coacervates . . . . .                                       | 48 |
| A.6  | Effect of salt on WGR-4 coacervates . . . . .  | 48 |
| A.7  | Partition of positively charged peptides in MD simulations . . . . .                                 | 49 |

|  |    |
|--|----|
| A.8 Partition of negatively charged peptides in MD simulations . . . . . | 49 |
|--|----|

# Chapter 1

## Introduction

Liquid-liquid phase separation (LLPS) has been shown to be involved in vital processes in cells in modern day life organisms in which it constitutes the phenomenon underlying assembly of membraneless organelles (MLOs)[1]. Coacervates are liquid droplets that emerge by LLPS from a solution when one (simple coacervates) or multiple (complex coacervates) macro-molecular species have enough interactions to overcome mixing entropy and demix into a molecule dense and dilute phase coexisting at equilibrium [2]. They typically form in solutions of long polymers (IDPs, RNA or DNA stretches) rich in oppositely charged residues (associative phase separation) or hydrophobic residues (segregative phase separation) or both. Thanks to the possibility to be controlled by environmental conditions such as pH [3] and salt concentration and their ability of sequestering and concentrate a wide variety of solute molecules, coacervates are involved in a large amount of vital processes, such as the formation of nucleolus during mitosis or Cajal bodies, as well as in diseases [4]. Additionally, the high content of solvent of the dense phase makes coacervates liquid, allowing for coalescence and division phenomena resembling primitive cell-like functions [5], together with the possibility of create coacervates composed by short and simple organic polymers, formed starting by constituent likely present in a prebiotic environment and their ability of store and protect organic material without necessitate a membrane, inspired an increasing interest in coacervates as protocells model [6].

The transition from chemistry to biology remains an unexplored territory still full of unresolved questions. Despite our incomplete comprehension of the processes that give rise to life from a set of molecules, insights can be obtained from modern day organisms and their fossilized ancestors. A fundamental feature observed across all scales within living systems is the ability to segregate and compartmentalize molecules vital to life within protected spatial regions, ultimately shaping the organism [7]. Following the idea of coacervates as a model to describe such a compartmentalized system, many attempt of embedding life-like functions in

coacervates have been done [8][9][10]. A key role in the evolution and fate of a possible protocells is played by the interactions and exchange of molecules (clients) with the environment. The formation and stability of coacervates are governed by many different intra- and inter-molecular interactions, involving a huge number of molecules. Adding client molecules increases the variety of molecular species in the system and thus the number of inter-species interactions ruling their partitioning. Since the high number of molecules involved, the emerging complex network of interactions makes predicting the localization of molecules after uptake and the impact of sequestering on the coacervate properties elusive.

Here we explore the formation, properties, and behavior of coacervates, focusing on two distinct systems: an homotypic coacervate formed by a single short peptide, of which phase separation is driven by both segregative and associative effects, and an heterotypical coacervate composed by a positively charged peptide and RNA, in which the main driving forces for LLPS are associative electrostatic interactions between the oppositely charged polymers. The central objectives of this study were to quantitatively study the coacervate phase diagrams and determine the sequestration of an array of different guest molecules, combined with a numerical approach which could unravel the most important molecular interactions and possibly predict outcomes of further studies.

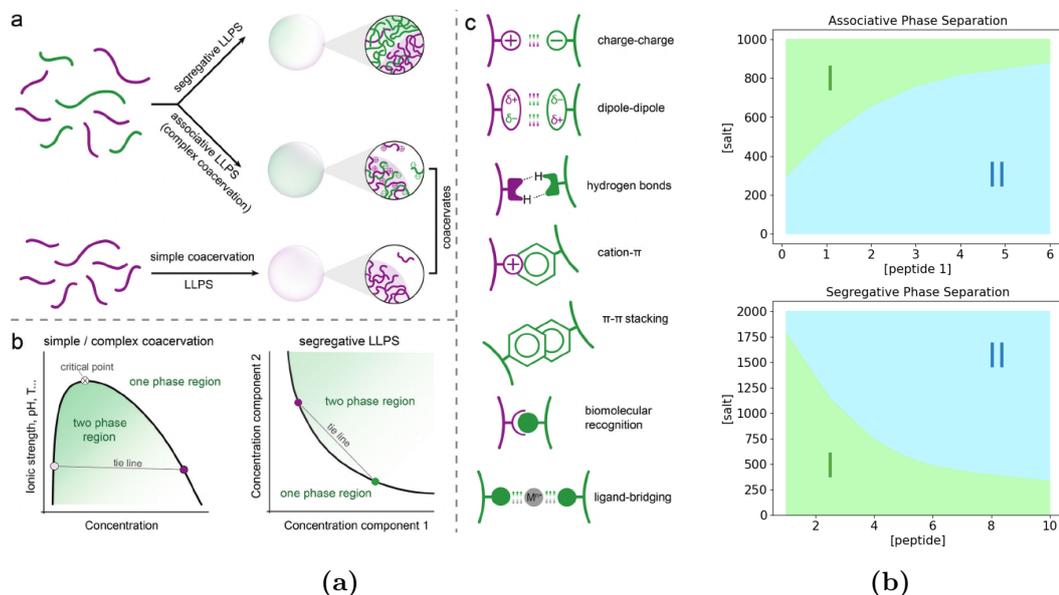
Chapter 1 offers an overview of the background literature, delving into the foundational principles underlying coacervate formation and their role as potential protocell models. This chapter contextualizes the research within the broader scientific landscape and outlines the motivations for studying these systems in greater detail.

Chapter 2 outlines the experimental methodology employed in this study. It details the preparation and characterization of the macro-molecular components crucial for the formation of coacervates as well as the analysis methods used to investigate phase diagrams and host-guest interactions. This chapter explains the phase diagram construction process, which served as a pivotal tool for understanding the conditions conducive to coacervate formation and stability. In addition, the experimental conditions used to study partitioning of client molecules into the two different coacervates are presented in this section.

Chapter 3 describes the coarse grained model used in molecular dynamics simulations. It offers an overview of the theoretical background behind the model used to simulate coacervate system and explore their interactions at a molecular level. A description of simulation steps and systems compositions is offered along this chapter.

In chapter 4 result from experiments and simulations are presented and compared. The discussion aims to investigate the main molecular forces that drive coacervation for the two systems analysed and their relevance in interaction with prebiotic relevant molecules such as short peptides or short RNA stretches.

In essence, this thesis encapsulates a careful examination of two distinct coacervate systems, unraveling their phase behaviors, molecular interactions, and experimental-numerical profiles.



**Figure 1.1:** (a) Image adapted from [11]. (a) Classification of coacervates and (b) their phase diagrams. (c) Molecular interactions driving coacervates. (b) (Top) Experimental phase diagrams in the peptide-salt concentration plane for associative phase separation of RP3/pU at constant [pU]. Bottom panel shows it for segregative phase separation of WGR-4. Blue corresponds to two-phase region, green to one-phase region.

## 1.1 Introduction to the Physical Principles of Coacervates Formation

Coacervates systems are commonly sub-classified as *simple* or *complex*. Simple coacervates are formed by only one macro-molecular species (plus solvent and eventually ions). Phase separation can occur both for segregative effects of hydrophobic residues or associative interactions between molecules of the same species. Complex coacervates are instead formed by two or more macro-molecular species, often oppositely charged, in which the main driving force of phase separation is the net electrostatic attraction between chains of different species.

In addition to net electrostatic interaction between charged molecules and hydrophobic effects, a lot of intriguing phenomena emerge from the complexity

of the system and several still poorly understood interactions play a role. An important group of LLPS-driving molecules are aromatic molecules, which are characterized by aromatic groups in which carbon atoms are  $sp^2$  hybridized with one electron delocalized in the molecular orbitals above and below the plane of the ring, as predicted by theory of linear combination of atomic orbitals (LCAO). The resulting distribution of negative charges can then attractively interact with positively charged molecules in the so called cation- $\pi$  interaction [12]. Moreover, it has been observed that a  $\pi$  orbital can also attract another  $\pi$  orbital in few well defined orientations in the so called  $\pi$ - $\pi$  stacking [13]. However, the phenomenon is still poorly understood and a definitive description of this kind of interactions necessitate further studies.

### 1.1.1 Nucleation and Growth

From a physical perspective, LLPS is guided by free energy minimization. For a two component system with free energy  $G$ , the condition delimiting the binodal line, the line at which phase separation becomes thermodynamically stable, is given by equal chemical potential for any species in the two phases and equal osmotic pressure

$$\mu_i^I = \mu_i^{II}, \quad \forall i \quad (1.1)$$

$$P^I = P^{II}, \quad P = \sum_i \mu_i \phi_i \quad (1.2)$$

with  $\mu_i$  chemical potential of the species  $i$  and  $P$  osmotic pressure.  $I$  and  $II$  refer to dilute and dense phase respectively,  $\phi_i$  is the volume fraction of component  $i$ . The spinodal line corresponds to the line at which the mixed states move from a local minimum to a local maximum in the free energy, becoming globally unstable. The spinodal line is defined as

$$\frac{\partial^2 G}{\partial \phi_i^2} = 0 \quad (1.3)$$

Binodal and spinodal lines coincide at the critical point. When crossing the phase boundary at the critical point the mixed state becomes a maximum in the free energy and thus globally unstable leading to phase separation driven by long-range critical fluctuations. This phenomenon is called spinodal decomposition. Away from the critical point spinodal and binodal lines form a region in the phase diagram in which phase separation is still thermodynamically favoured, but the mixed state may survive as a meta-stable state. In this region phase separation is driven by nucleation, typically of droplets, of the minority phase in solution. Then, droplets that overcome the critical size, are stable against environmental fluctuations and can undergo a growth process governed by several diverse phenomena such as diffusion limited growth, coarsening and coalescence.

It has been shown that the growing process is governed by dynamic scaling law and thus, spatial domains of the system exhibit self-similarity in time. This follows from the emergence of a single length scale  $L$ , governing domain kinetics, so morphology of a region of size  $L$  at time  $t_1$  is statistically equivalent to that of a region of size  $L(t_2/t_1)^n$  at time  $t_2$ , where  $n$  is the dynamic scaling exponent. This argument has been fully discussed by Berry et. al. with focus on the consequence of dynamic scaling law on properties related to droplets size distribution [14]. It has been shown, in particular, that the average radius distribution follows a power law of time

$$\langle R \rangle = (Kt)^n \quad (1.4)$$

where the exponent  $n$ , as well as the prefactor  $K$ , depend on the phenomena involved in the growth process. In a real system, as well as in MD simulations involved in our studies, growing processes are typically driven by contribution of more phenomena giving rise to possibly hybrid situation with exponents different from those found in [14].

### 1.1.2 The Screening Effect of Salt

A crucial role in the control of phase separation is played by salts. Thanks to their ability to dissociate in free ions they can freely move in solution and screen charged molecules. The presence of salt in solution can thus be crucial to formation or dissolution of coacervates that will depend on the nature of macro-molecules involved. In associative phase separation where the electrostatic force between oppositely charged polymers drives phase separation an increase of ions leads to dissociation of molecules in solution (1.1 (b) top panel). In segregative phase separation, the effect of salt depends on the molecules involved, considering a solution of a single macro-molecular species rich in positive charges and aromatic residues, used in experiments and simulations in this study, increasing salt concentration screens repulsion between charges of the same sign, allowing molecules to come closer and interact in weak and short-range interactions such as cation- $\pi$  and  $\pi - \pi$  stacking. It is then expected that, for such a system, segregative phase separation is enhanced by addition of salt (1.1 (b) bottom panel).

In addition to electrostatic interactions between charged poly anions and/or poly cations, there are some results that seem to show that salt can modify weak interactions involving aromatic residues enhancing cation- $\pi$  attractions. Kim et al. proposed that salt in sea water trigger coacervates formation from mussel's protein mfp-1, increasing cation- $\pi$  attraction between positively charged lysine and aromatic rings in tyrosine [15]. However, a theoretical description of how salt affects weak electrostatic interactions than Coulomb attraction still need to be investigated.

### 1.1.3 Gibbs Free Energy

In the following, a Flory-Huggins mean-field model is applied to segregative phase separation of simple coacervates [16]. To describe complex coacervation, Voorn-Overbeek theory was used that represent the natural extension of Flory-Huggins theory for charged polymers [17]. It is worth to mention that Voorn-Overbeek theory fails in reproduce phase behavior of short polymers, and large molecules and large charge densities are required to make the system phase separate. Further development have been done in the study of phase separation [18]. Despite that, because of its simplicity and the possibility to reproduce qualitative results of phase separation, Flory-Huggins theory represents a pivotal model in the study of LLPS and allows to identify the main aspect driving formation of both simple and complex coacervates.

The entropic contribution to free energy has the same form for both segregative and associative phase separation and is given by the Flory-Huggins entropy for polymers in solution. For a system in a volume  $V$  containing  $n_i$  molecules of the species  $i$ , of length  $r_i$ , the entropy density of mixing is given by

$$\frac{TS_{mix}}{Nk_bT} = - \sum_i \frac{\phi_i}{r_i} \log(\phi_i) \quad (1.5)$$

where  $k_b$  is the Boltzmann constant and  $\phi_i = \frac{n_i r_i v}{V}$  volume fraction of the species  $i$ . Here  $v$  is the elementary volume so that  $r_i v$  is the molecular volume of species  $i$  and  $N = \sum_i n_i r_i$  total number of elementary volume units in the system.

The energetic term can be separated in a nearest-neighbour interaction, derived from a mean-field lattice model that takes into account the inter-species interactions solute-solute and solute-solvent as well as intra-species interactions [16], and an electrostatic term for charged polymers, the distinctive term of a Voorn-Overbeek model for complex coacervation [17]. The intra- and inter-species solute-solute and solute-solvent interactions can be grouped in the following energy term

$$\frac{U_{FH}}{Nk_bT} = \frac{1}{2} \sum_i \sum_j \chi_{ij} \phi_i \phi_j \quad (1.6)$$

where the  $\chi_{ij}$  parameters shape interactions between different molecular species and is defined as

$$\chi_{ij} = \frac{z}{k_bT} \left( u_{ij} - \frac{u_{ii}u_{jj}}{2} \right) \quad (1.7)$$

with  $z$  number of nearest neighbour in the lattice, and  $u_{kl}$  free energy of interactions between adjacent segments of species  $k$  and  $l$ . These free energies can be either negative giving a contribution to the mixing entropy, or positive favouring demixing.

An additional energy term need to be considered when the theory is applied to complex coacervation of charged polymers, to consider the electrostatic interaction

between oppositely charged species. Has been shown by Debye and Huckel that for an aqueous solution such as that described above, if molecules of the species  $i$  have charge  $z_i e$ , with  $e$  elementary charge, the electrical free energy can be written as

$$U_{el} = -\frac{e^2}{3\epsilon} k N_z \quad (1.8)$$

with  $N_z = \sum_i n_i |z_i|$  total number of charges and  $k$  Debye length defined as

$$k = \sqrt{\frac{4\pi e^2 N_z}{\epsilon k_b T V}} \quad (1.9)$$

where  $\epsilon$  is the relative permittivity. The free energy density for charged polymer solutions is then given by

$$\frac{U_{el}}{N k_b T} = -\alpha \left( \sum_i \sigma_i \phi_i \right)^{3/2} \quad (1.10)$$

with  $\alpha = \frac{1}{12\pi\sqrt{v}} \left( \frac{e^2}{\epsilon\epsilon_0 k_b T} \right)^{3/2}$  and  $\epsilon_0$  vacuum permittivity.

At high salt concentrations the observed behavior is qualitatively different from the one predicted by Debye-Huckel theory and is dependent on the nature of salt [19]. However for a monovalent salt, such as NaCl used through this project, Debye-Huckel theory is in well agreement with in-vitro results up to 1M concentration.

It should be also mentioned that the original Voorn-Overbeek formulation for poly electrolyte complexes, only considers the electrostatic and entropic terms in free energy [17]. However, further developments introduce the term in (1.6) to take into account non-conventional electrostatic interactions as well as hydrophobic effects and predict coacervates in a wider window of parameters (size, charge density, etc.) [20].

## 1.2 Theory of Partitioning

Let us consider a system of two molecular species plus solvent that phase separates through interactions expressed by (1.6). The Gibbs free energy density of mixing for that system can be written as

$$\frac{G_{mix}}{N k_b T} = \sum_{i=1,2,3} \frac{\phi_i}{r_i} \log \phi_i + \chi_{12} \phi_1 \phi_2 + \chi_{13} \phi_1 \phi_3 + \chi_{23} \phi_2 \phi_3 \quad (1.11)$$

The chemical potential for any species can be calculated as

$$(\mu_i - \mu_i^0) = N_A \left( \frac{\partial G_{mix}}{\partial n_i} \right) \quad (1.12)$$

where  $N_A$  is the Avogadro's number and  $\mu_i^0$  the chemical potential of species  $i$  when  $\phi = 1$ . At equilibrium the chemical potential of any species in the two phases must coincide, and since the chemical potential of pure species is always the same

$$\frac{\partial G_{mix}^I}{\partial \phi_i} = \frac{\partial G_{mix}^{II}}{\partial \phi_i} \quad \forall i \quad (1.13)$$

Let's now consider the case in which a new molecular species (4) is added to the system. Under the approximation of low client concentration ( $\phi_4 \ll \phi_i$  for  $i = 1, 2, 3$  and so  $N \approx n_1 + n_2 + n_3$ ) and equally soluble molecules, that is equal interaction with the solvent (molecular species 1)  $\chi_{12} = \chi_{13} = \chi_{14} = \chi_S$ , then 1.11 becomes

$$\begin{aligned} \frac{G_{mix}}{k_b T} = & \sum_{i=1,2,3,4} n_i \log \phi_i + N \left[ \chi_S (\phi_1 \phi_2 + \phi_1 \phi_3 + \phi_1 \phi_4) + \right. \\ & \left. + \chi_{23} \phi_2 \phi_3 + \chi_{24} \phi_2 \phi_4 + \chi_{34} \phi_3 \phi_4 \right] \end{aligned} \quad (1.14)$$

using 1.12 (full derivation in B)

$$\frac{(\mu_4 - \mu_4^0)}{N_A k_b T} = 1 + \log \phi_4 + r_4 \left[ \phi_1 (\chi_S - 1) + \phi_2 \left( \chi_{24} - \frac{1}{r_2} \right) + \phi_3 \left( \chi_{34} - \frac{1}{r_3} \right) \right] \quad (1.15)$$

It is thus possible to extract the partition coefficient of clients (species 4) that is defined as the ratio between molecules in the dense phase (II) and in the dilute phase (I)

$$Kp_4 = \frac{\phi_4^{II}}{\phi_4^I} = \quad (1.16)$$

$$\exp \left\{ r_4 \left[ (\phi_1^{II} - \phi_1^I) (1 - \chi_S) + (\phi_2^{II} - \phi_2^I) \left( \frac{1}{r_2} - \chi_{24} \right) + (\phi_3^{II} - \phi_3^I) \left( \frac{1}{r_3} - \chi_{34} \right) \right] \right\} \quad [21]$$

- The partitioning coefficient depends exponentially on client size
- $\phi_1^{II} - \phi_1^I < 0$  since solvent concentration is expected larger in the dilute phase than in the dense one. Favourable interactions with solvent ( $\chi_S < 0 \rightarrow 1 - \chi_S > 0$ ) are then expected to decrease Kp. However, the difference in solvent fraction between dense and dilute phase is negligible compared to differences in polymers fractions making this term negligible.
- $\phi_2^{II} - \phi_2^I$  and  $\phi_3^{II} - \phi_3^I$  are instead expected positive. Favourable interactions with coacervates molecules ( $\chi_{i4} < 0 \rightarrow \frac{1}{r_i} - \chi_{i4}$ ) are then expected to increase Kp.

# Chapter 2

## Materials and Method

### 2.1 RP3/polyU Phase Diagram

#### Preparation of Stock Solutions

Stock solutions of peptides and RNA have been made starting from crude materials, dissolving them in milliQ ( $18.2 \text{ M}\Omega \text{ m}^{-1}$ ) and adding NaOH to reach a neutral pH.

RRASLRRASLRRASL (RP3) was ordered from genscript and its molecular weight was determined as  $2567.11 \text{ g mol}^{-1}$ . 10 mg of RP3 was dissolved in 390  $\mu\text{L}$  milliQ, 360  $\mu\text{L}$  milliQ was taken and 17  $\mu\text{L}$  0.1 M NaOH was added to reach pH 7. 73  $\mu\text{L}$  milliQ was added to reach a total volume of 450  $\mu\text{L}$  and a final RP3 concentration of 8 mM.

10 mg polyU was dissolved in 1 mL milliQ and 150  $\mu\text{L}$  10 mg  $\text{mL}^{-1}$  polyU was taken and 1.5  $\mu\text{L}$  0.1 M NaOH was added for pH 7. 148.5  $\mu\text{L}$  MQ was added to reach a total volume of 300  $\mu\text{L}$  and a final polyU concentration of 5 mg  $\text{mL}^{-1}$ .

#### Turbidity Analysis

A Tecan Spark Plate Reader was employed in turbidity analysis to measure absorption at 600 nm. Two phase diagrams have been constructed at constant polyU concentration, samples were prepared adding RP3 to milliQ until target concentrations (0.1 mM, 0.25 mM, 0.5 mM, 1 mM, 2 mM, 3 mM, 4 mM, 5 mM) and NaCl was added for each RP3 concentration in the range 0 to 1 M (0 mM, 50 mM, 100 mM, 200 mM, 250 mM, 300 mM, 400 mM, 500 mM, 600 mM, 750 mM, 1 M). Then, all the samples were brought to a final volume of 20  $\mu\text{L}$  adding polyU at 0.25 mg/mL and 0.5 mg/mL.

Four more phase diagrams were acquired fixing the RP3 concentration, sample were made adding polyU to milliQ in concentrations in the range 0.05:1.00 mg/mL (0.05 mg/mL, 0.1 mg/mL, 0.25 mg/mL, 0.5 mg/mL, 1 mg/mL) and NaCl in the

range 0 to 750 mM (0 mM, 50 mM, 100 mM, 200 mM, 250 mM, 300 mM, 400 mM, 500 mM, 600 mM, 750 mM). All the samples were then brought to a final volume of 20  $\mu\text{L}$  adding RP3 at 0.1 mM, 0.5 mM, 1 mM and 4 mM just before analysis. An additional set of samples at 1 M NaCl has been tested for each polyU concentration and 4 mM RP3.

A milliQ sample was added to each individual experiment to acquire blank reference value and phase diagrams were constructed using  $1.5 \times$  blank absorption as a threshold to identify coacervates. Samples with turbidity above the threshold were identified as phase separated, below the threshold as dissolved.

A second control was performed on samples in the phase boundary region, by detecting the presence of coacervates using an inverted fluorescence microscope (Olympus IX83).

## 2.2 WGR-4 Phase Diagram

### Preparation of Stock Solutions

WGRGRGRGWY (WGR-4) was ordered from genscript and its molecular weight determined as  $1706.47 \text{ g mol}^{-1}$ . Two different stock solutions of WGR-4 were prepared dissolving 20 mg of crude material in milliQ at a concentration of 50 mg/mL for each one. The solutions were adjusted adding NaOH (1 and 0.1 M) to pH 6 in the first and 7.3 in the second, and both solutions were brought to a total volume of 500  $\mu\text{L}$ . The concentration of any stock solution was established by UV-Vis absorption analysis employing UV-NanoDrop with an extinction coefficient of  $12 \text{ 660 cm}^{-1}\text{M}^{-1}$  at 280 nm. Samples used in the same set of experiments have always been prepared starting from the same stock solution to limit the effect of possible deviation in concentrations.

### Turbidity Analysis

Turbidity analysis has been used to construct phase diagrams employing a Tecan Spark Plate Reader to measure absorption at 600 nm. Samples in the range 1:10 mM WGR-4 and 0.1:2 M NaCl have been prepared adding WGR-4 stock solutions to the milliQ to reach target concentration (1 mM, 2.5 mM, 4 mM, 5 mM, 6 mM, 7.5 mM, 10 mM) and NaCl was added (100 mM, 250 mM, 400 mM, 500 mM, 600 mM, 750 mM, 1 M, 1.5 M, 2 M). A milliQ sample was measured for each experiment to acquire blank and  $1.5 \times$  blank has been used as a threshold to identify coacervates.

A second control was performed on samples in the phase boundary region by detecting the presence of coacervates using inverted light microscope.

## UV-Vis Measurements

Absorption analysis was used to validate concentration of any WGR-4 stock solutions. A dilution series was prepared by diluting WGR-4 (20 mM, pH 6). Three samples were instead made from pH 7.3 solution (10 $\times$ , 20 $\times$  and 40 $\times$ ). All the samples were tested at least twice employing UV-NanoDrop for the absorbance analysis.

An extinction coefficient of 12 660 cm<sup>-1</sup>M<sup>-1</sup> at 280 nm and a path length of 1 cm for the light beam were used in the translation of absorbance in concentration. Absorbance was measured at the wavelength corresponding to the actual peak in the spectrum with usually no more than 2 nm shift with respect to the theoretical one of 280 nm.

## 2.3 Partition Coefficient of Clients in Coacervates

### Preparation of samples

Experiments of WGR-4 with client molecules were carried out starting from stock solution of WGR-4 with a concentration in the range 17-21.5 mM validated by UV-Vis analysis. WGR-4 up to a concentration of 7.5 mM and NaOH up to a concentration of 40 mM were added to milliQ and vortexed. After the formation of coacervates client solution, previously prepared with molecules labelled with fluorescent dyes and unlabelled molecules of the same species, was added at three different concentrations (Table A.1).

Samples of RP3-polyU were prepared starting mixing RP3 at concentration of 2 mM and polyU at 1 mg/mL in milliQ and adding client solution at three different concentrations (Table A.2).

WCLRRA (M.W. 1146.05 g mol<sup>-1</sup>) and (W<sup>CN</sup>-CLRRA (M.W. 1171.05 g mol<sup>-1</sup>) were ordered from CASLO and used to prepare stock solutions at 10 mM in milliQ. Concentrations in stock solutions were tested through UV-Vis analysis employing UV-NanoDrop. An extinction coefficient of 5810 cm<sup>-1</sup>M<sup>-1</sup> at 280 nm and a light path length of 1 cm was used. A concentration window of 0.17 mM to 4.13 mM was determined from CyanoTrp emission coefficient. Three samples for any stock were tested at dilution of 5, 10 and 20 times the original concentration. Final solutions of WLCRRA and W<sup>CN</sup>-CLRRA were mixed at 90:10, 50:50, 0:100.

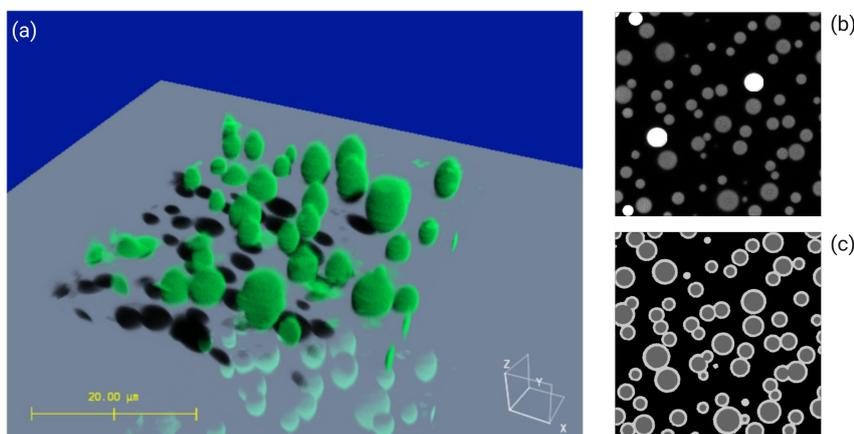
### Confocal imaging

Microscope slides from Ibidi with 18 square wells of 200  $\mu$ L were modified with 80  $\mu$ L pLL-g-PEG 0.01 (0.01 mg mL<sup>-1</sup>) in each well for 24 hours, after cleaning the slides in a O<sub>2</sub> plasma cleaner.

Leica SP8 confocal microscope was employed in analysis of clients partitioning. Any microscope setting has been fixed in a first round of experiments and taken constant along the subsequent experiments. 16 bit Z-stack of  $1024 \times 1024 \times Z$  voxels per plane were acquired using a  $40 \times$  Zeiss oil immersion lens with  $NA = 1.3$  and an additional digital  $7 \times$  zoom ( $40.6 \times 40.6$  nm per voxel), a laser intensity of 0.1%, speed 400 and a distance of 130 nm between planes (fig. 2.1 left panel). Variable gain has been used to avoid out of range voxels and reduce the amount of zero intensity voxels as much as possible. Variable number of planes have been used in each z-stack depending on the size of the droplets, approximately  $5 \mu\text{m}$  was acquired, leading to z-stacks of  $\sim 40$  planes.

Excitation wavelength of 405 nm and 410-500 nm acquisition window was used for NADH, 488 nm excitation wavelength and 500-540 nm acquisition was used for clients labelled with fluorescein amidite (FAM), 552 nm excitation and 560-600 nm acquisition was used for cyanine 3 (Cy3) and a wavelength of 638 nm and acquisition window of 650-680 nm was used for cyanine 5 (Cy5). For molecules with 5-Cyano-Tryptophan an excitation wavelength of 405 nm and a window of 410-600 nm was used.

No more than one z-stack for each acquisition area was acquired to avoid bleaching of the region and the samples were acquired in region of the well far apart the one from the other to be representative of the whole sample. At least three images for sample were acquired and almost all systems were analysed twice.



**Figure 2.1:** (a) 3D reconstruction of a Z-stack processed with Huygens Deconvolution SVI. Scale bar is  $20 \mu\text{m}$ . (b), (c) Schematic segmentation process of droplets in the maximum intensity plane of Z-stack shown on the left.

### 2.3.1 Image Segmentation and Analysis

The analysis of images from confocal microscopy was implemented in two steps. In a first analysis, raw data are passed to LabKit [22], a Fiji machine learning based plugins that allow to create a mask, using a pixel classifier, to divide voxels belonging to dense and dilute phase and blurs (2.1 bottom-right panel). Different images from the same sample were always analysed with the same pre-trained model.

A julia code was then implemented for the second part of the analysis<sup>1</sup>. The code allow for the reconstruction of droplets starting from the mask and computation of single droplets partition coefficients (2.1 top-right panel). The analysis proceed plane by plane, first all the Z stack is passed to the program to compute intensity of each plane and extract the maximum intensity plane. Then the maximum intensity plane is scanned pixel by pixel to reconstruct single droplets using the mask for pixels positions and the raw data for intensities. The image is analysed looking for each pixel at the left and top pixels and proceeding row by row, leading to a full analysis in  $\mathcal{O}(N^2)$  for an image  $N \times N$ . The code ends with a list of droplets in the plane and pixels belonging to each droplet. The Kp is then computed for each droplet as  $\frac{[Int.Rel.]_{droplet}}{[Int.Rel.]_{dilute}}$  with  $[Int.Rel.]_A = \frac{1}{\#A} \sum_{v \in A} Int.(v)$ .

The average Kp of a system is then extracted as the average of single droplets Kp's from 4 to 8 different images using as only manual parameters the threshold on droplets size. All the images were analysed with a cutoff of 500 pixels on size, approximately 1  $\mu\text{m}$  diameter.

---

<sup>1</sup>Code available at <https://github.com/RSchiroli/Kp.git>



# Chapter 3

## Numerical Simulations

### 3.1 Theoretical Background

A coarse grained model has been implemented in molecular dynamics simulations attempting to reproduce the phase behavior of coacervates systems employed in experimental analysis and to gain insight the molecular interactions involved in the formation and evolution of coacervates.

A coarse grained 1BPA-3BPN (one bead per amino acids - three bead per nucleotide) model has been chosen to describe coacervates and client molecules. Ions are implicitly considered through screening of electrostatic force and solvent is implicitly implemented in hydrophobic-hydrophilic interactions. GROMACS [23][24][25] has been used to run molecular dynamics simulations on the high performance cluster Snellius.

#### 3.1.1 One Bead Per Amino Acids (1BPA)

In the 1BPA model employed each amino acid (AA) is represented by a single bead of 120 Da average mass at the position of the alpha-carbon. Two amino acids in a peptide chain are then located at an equilibrium distance of  $b = 0.38$  nm and connected through an harmonic bonding potential

$$\phi_{bond} = K_b(r - b)^2 \quad (3.1)$$

with  $K_b = 8038$  kJ·nm<sup>-2</sup>·mol<sup>-1</sup>. The whole bonded potential is then computed as the sum of bonding, angle and torsion potentials. Angle and torsion potentials have been derived in [26]. The non bonded potential is instead given by

$$\phi_{nb} = \phi_{hp} + \phi_{cp} + \phi_{el} \quad (3.2)$$

sum of hydrophobic, cation- $\pi$  and electrostatic terms.

$$\phi_{hp} = \begin{cases} \epsilon_{rep} \left(\frac{\sigma}{r}\right)^8 - \epsilon_{ij} \left[\frac{4}{3} \left(\frac{\sigma}{r}\right)^6 - \frac{1}{3}\right] & \text{if } r \leq \sigma \\ (\epsilon_{rep} - \epsilon_{ij}) \left(\frac{\sigma}{r}\right)^8 & \text{if } r \geq \sigma \end{cases} \quad (3.3)$$

with  $\epsilon_{ij} = \epsilon_{hp} \sqrt{(\epsilon_i \epsilon_j)^\alpha}$ ,  $\sigma = 0.60$  nm,  $\epsilon_{rep} = 5$  kJ·mol<sup>-1</sup>,  $\alpha = 0.15$ ,  $\epsilon_{hp} = 6.5$  kJ·mol<sup>-1</sup> and  $\epsilon_i$ ,  $\epsilon_j$  single bead hydrophobicity constants described in [27].

$$\phi_{cp,ij} = \epsilon_{cp,ij} \left[ 3 \left(\frac{\sigma_{cp}}{r}\right)^8 - 4 \left(\frac{\sigma_{cp}}{r}\right)^6 \right] \quad (3.4)$$

with  $\sigma_{cp} = 0.60$  nm and  $\epsilon_{cp,ij}$  energy of cation- $\pi$  interactions computed in [28].

$$\phi_{el} = \frac{q_i q_j}{4\pi\epsilon_0\epsilon_r(r)} e^{-kr} \quad (3.5)$$

where  $k$  is the Debye length and depends on temperature and ion concentrations and  $\epsilon_r(r) = S_s \left[ 1 - \frac{r^2}{z^2} \frac{e^{r/z}}{(e^{r/z} - 1)^2} \right]$ ,  $S_s = 80$  and  $z = 0.25$  nm [29]. The exponents of 8 and 6 in the Lennard Jones potential of (3.3) and (3.4) were chosen in place of the usual 12 and 6 to make the interactions smoother [30].

### 3.1.2 Three Beads Per Nucleotide (3BPN)

A 3BPN model was used to simulate RNA molecules. Three beads are used to represent a nucleotide, one for the phosphate group in the P atom position with charge -1, one for the ribose located at C4' position and one for the nucleobase with position dependent on the specific nucleobase [31].

Bonded potential is the sum of bonds and bending harmonic potentials

$$\phi_b = \phi_{bond} + \phi_{bend} \quad (3.6)$$

with parameters described in [27]. Non bonded interactions are given by the sum of hydrophobic, base stacking and electrostatic interactions

$$\phi_{nb} = \phi_{hp} + \phi_{bs} + \phi_{el} \quad (3.7)$$

where  $\phi_{hp}$  and  $\phi_{el}$  have same form and parameter values as (3.3) and (3.5) respectively with hydrophobicities derived in [32].  $\phi_{bs}$  is introduced to take into account the potential energy of base stacking between nitrogen bases through a modified version of (3.3)

$$\phi_{bs} = \begin{cases} \epsilon_{bs1} \left(\frac{\sigma}{r}\right)^8 - \epsilon_{bs2} \left[\frac{4}{3} \left(\frac{\sigma}{r}\right)^6 - \frac{1}{3}\right] & \text{if } r \leq \sigma \\ (\epsilon_{bs1} - \epsilon_{bs2}) \left(\frac{\sigma}{r}\right)^8 & \text{if } r \geq \sigma \end{cases} \quad (3.8)$$

where  $\epsilon_{rep}$  becomes  $\epsilon_{bs1} = 2.71$  kJ·mol<sup>-1</sup> and  $\epsilon_{ij}$  becomes  $\epsilon_{bs2} = 1.25$  kJ·mol<sup>-1</sup>.

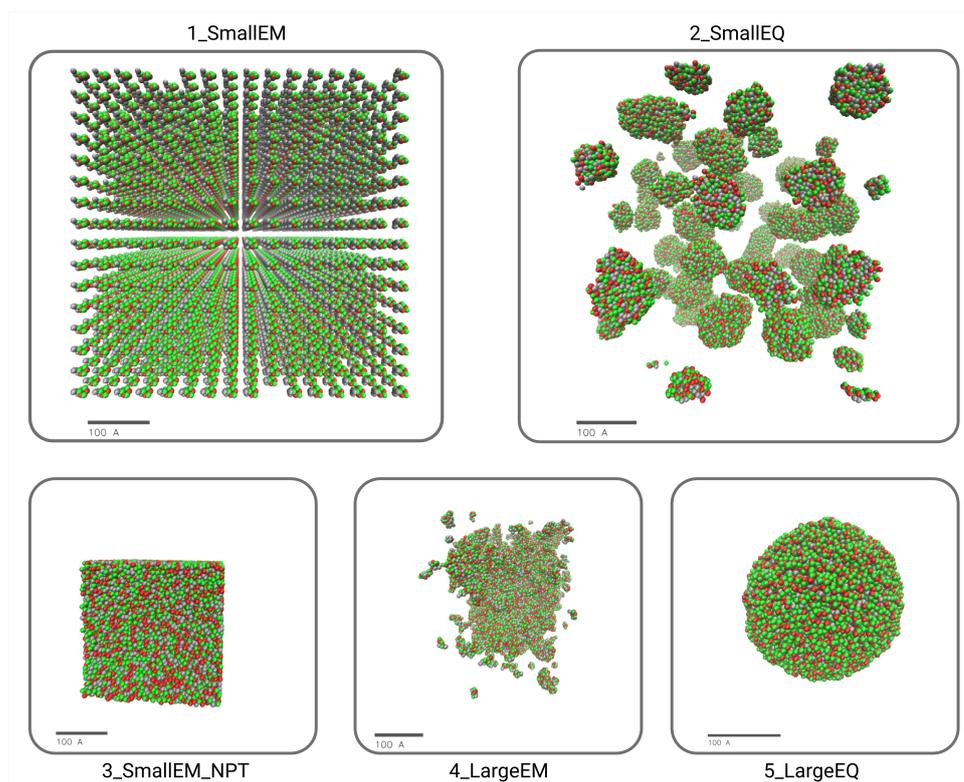
### 3.1.3 1BPA-3BPN compatibility

Since molecules in 1BPA and 3BPN models interact via potentials written in the same form, it is possible to implement an hybrid 1BPA-3BPN model in which proteins and nucleotides interact with intra-species interactions described in 3.1.1 and 3.1.2 respectively. The inter-species amino acids - nucleotides interactions are described via hydrophobic-hydrophilic interactions as in (3.3). In addition, charged residues can interact via electrostatic interactions (3.5) with P beads in the RNA backbone. Inter-species cation- $\pi$  and  $\pi - \pi$  stacking are instead not taken into account in the model. However, it has been shown that unfolded regions in IDP's and RNA can interact via non conventional electrostatic interactions [33]. A stacking effect between aromatic residues (Phe, Trp, Tyr) and nucleobases is in principle captured by hydrophobic interactions. Since they are both highly hydrophobic, so  $\epsilon_i$  and  $\epsilon_j \approx 1$  and  $\epsilon_{ij} > 5$  in (3.3) makes it attractive. However, the interaction energy is weaker than typical  $\pi - \pi$  energy involved in 3BPN model. Interactions between positively charged amino acids (Arg, Lys) and nucleobases are instead neglected by the model, Arg and Lys have  $\epsilon_i = 0$  since they are hydrophilic that means (3.3) is always repulsive for these beads.

### 3.1.4 Simulation Protocol

The simulation protocol implemented is divided in five steps that can be summarized as follow:

1. *Small Energy Minimization*: First the topology files for any molecules that need to be added to the simulation box is created according to the 1BPA-3BPN model implemented. Then the simulation box is created with an initial size dependent on the amount of molecules and divided into a lattice structure with a number of sites equal to the number of molecules that need to be added. Molecules are then allocated in any lattice site in their original structure (e.g. crystal structure). The energy minimization step applied involve a steepest descent method with an energy tolerance of  $1 \text{ kJ mol}^{-1}\text{nm}^{-1}$ , to allow molecules reach their minimum energy structure in the system.
2. *Small Equilibration*: A first NVT Langevin dynamics is ran to bring the system at the chosen temperature of 300 K. This step consists of 25 000 time steps with  $\Delta t = 20 \text{ fs}$  for a total of 0.5 ns.
3. *NPT Equilibration*: A second equilibration step is ran in the NPT ensemble bringing the system at a pressure of 1 bar to have molecules closer and start forming cluster faster. This step consists of 25 000 000 time steps with  $\Delta t = 20 \text{ fs}$  for a total simulation time of 500 ns.



**Figure 3.1:** Snapshots of the final frame of different steps involved in MD simulations. The process shows the formation of coacervates in a box with 4 000 WGR-4 molecules at 7.5 mM with 1 M ions. Green residues are Gly, red residues are Arg, silver residues are Tyr and Trp. Scale bars are 100Å.

4. *Large Energy Minimization:* The system is then inserted in a new box to reproduce actual concentration at room pressure and a second energy minimization step is then ran to relax the molecules after the expansion with same energy tolerance and time steps of *Small Energy Minimization*.
5. *Large Equilibration:* The last part of the simulation is the only one used to evaluate them and extract data and statistics. It consists of an equilibration step again in the NVT ensemble long 150 000 000 time steps with  $\Delta t = 20$  fs for a total of  $3 \mu s$  and allow the system to reach the equilibrium. The time needed to actually reach equilibrium highly depend on the system involved, for systems analysed the simulated interval allow to reach equilibrium and ran a considerable fraction of the simulation in an equilibrium condition in which data allowing in principle the study of both formation and stability of studied systems.

## 3.2 Phase Diagrams

### 3.2.1 WGR-4

We started simulating WGR-4 phase diagram in the plane of peptides-ions concentrations. Systems of 4 000 molecules, consisting of 10 amino acids (fig. 4.1) (10 beads per molecules), were simulated leading to a total of 40 000 beads per simulation. Boxes with different peptide concentrations (0.1 mM, 1 mM, 2.5 mM, 5 mM, 7.5 mM, 10 mM, 15 mM) and ion concentrations (10 mM, 50 mM, 100 mM, 400 mM, 1 M) were simulated and trajectories were used for evaluation.

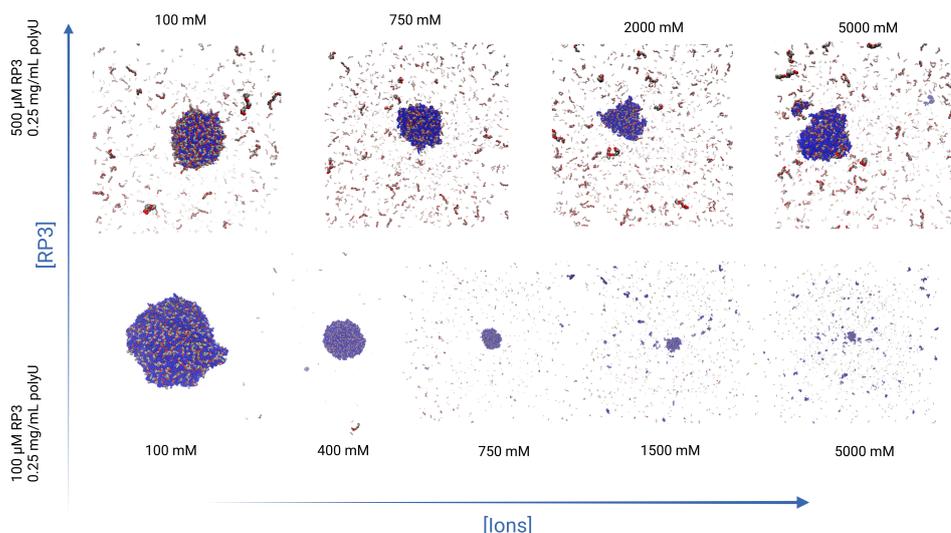
Only the large equilibration step is used for evaluation. One frame every 5 ns is extracted and used to compute contact map, setting a threshold of 0.7 nm for molecules in contact. The contact map at any time frame is then represented as a tree. Clusters corresponding to connected components of the graph are extracted and used for statistics. Specifically the evolution of average and largest cluster size was considered in the analysis. An additional set of information about intra- and inter-molecular interactions is extracted from contact map and discussed in section 4. In the analysis volume and radius of droplets were considered proportional to the number of molecules in the droplet itself.

### 3.2.2 RP3/polyU

Phase behavior of RP3/polyU coacervates was simulated at constant polyU concentration of 0.25 mg/mL for RP3 concentrations of 0.1 mM and 0.5 mM. Simulations boxes were created using a fixed number of 2 000 RP3 molecules of 15 beads each one (fig. 4.2) for a total of 30 000 peptide beads and varying RNA number accordingly. In boxes with 0.1 mM : 0.25 mg/mL 150 RNA molecules were used whereas 30 were added to 0.5 mM : 0.25 mg/mL boxes. Each polyU molecule contains 100 uridine nucleotides leading to a total number of nucleotide beads of 45 000 and 9 000 respectively. The total number of beads in a simulation box 75 000 (12 000 positively charged beads, 15 000 negatively charged beads) or 39 000 (12 000 positively charged beads, 3 000 negatively charged beads) depending on RP3 concentration. Any RP3/polyU combination was simulated at 10 mM, 100 mM, 400 mM, 750 mM, 1 M, 1.5 M, 2 M, 2.5 M, 5 M ions concentration and results used to extract contact maps and graphs as in 3.2.1.

## 3.3 Host-Client Interactions

For the partitioning experiment only WGR-4 coacervates were simulated as peptide-RNA interactions were not quantified correctly in the model. Simulation boxes were composed of two macro-molecular species, coacervates and client molecules,



**Figure 3.2:** Snapshots of the final frame of MD simulations for RP3/polyU coacervates. PolyU concentration is fixed at 0.25 mg/mL. Moving from left to right increase ion concentration from 100 mM up to 5 M. Bottom line corresponds to 100  $\mu$ M RP3, showing dissolution of coacervates at high ion strength. Top line shows simulations for 500  $\mu$ M RP3 where increase ion concentration leads to a release of RP3 in solution but never dissolves coacervates. Red beads are Arg, silver are Ala, Ser, Leu and blue are U.

both added to the box at the beginning of the simulation. 3 500 WGR-4 molecules were used at 7.5 mM and 400 mM ion concentration. Client molecules analysed and concentrations used are displayed in table 3.1

**Table 3.1:** List of molecules, concentrations and related number of beads used in client partitioning simulations. Numbers in molecule names represent repetition of the elementary unit. Concentrations corresponds to a number of molecules of 50, 100, 250 for nucleotides and 50, 125, 250 for peptides.

| Client                                   | Concentrations (mM) | # Beads            |
|--|---------------------|--------------------|
| (polyA/C/G/U) <sub>15</sub>              | 0.1, 0.2, 0.5       | 2250, 4500, 11250  |
| Trp-Cys-Leu-Arg-Arg-Ala                  | 0.1, 0.25, 0.5      | 300, 750, 1500     |
| Arg <sub>10</sub> , Glu <sub>10</sub>    | 0.1, 0.25, 0.5      | 500, 1250, 2500    |
| Arg <sub>30</sub> , Glu <sub>30</sub>    | 0.1, 0.25, 0.5      | 1500, 3750, 7500   |
| Arg <sub>100</sub> , Glu <sub>100</sub>  | 0.1, 0.25, 0.5      | 5000, 12500, 25000 |
| RP3 = (Arg-Arg-Ala-Ser-Leu) <sub>3</sub> | 0.1, 0.25, 0.5      | 750, 1875, 3750    |
| RRP = (Arg-Gly-Arg-Gly-Gly) <sub>5</sub> | 0.1, 0.25, 0.5      | 1250, 3125, 6250   |

# Chapter 4

## Results and Discussion

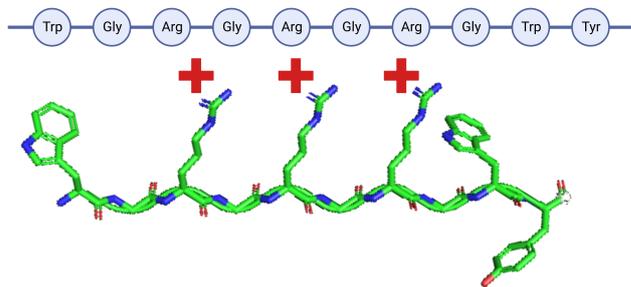
### 4.1 Phase Diagrams

To gain insight on the phase behavior of coacervate systems used throughout this project, the first step involved an experimental investigation of their phase diagrams in the peptide-salt concentrations plane, attempting to replicate theoretical patterns and capture their essential features. The resulting data have been used to select an optimal concentration range in subsequent analyses, ensuring the stability of coacervates used in imaging while minimizing susceptibility to temperature fluctuations and agitation. Furthermore, these experimental phase diagrams will serve as a benchmark for comparing and evaluating simulations before their application in host-client interaction studies.

#### 4.1.1 Choice of Coacervates

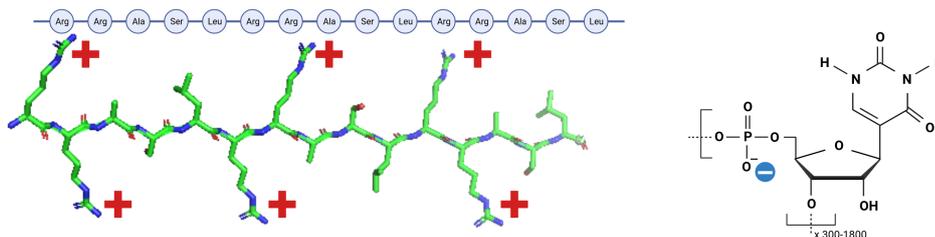
Coacervates found in modern day living organisms are typically formed by long and complex polymers constituted by unfolded regions of IDP's or RNA and DNA stretches, all unlikely present in a prebiotic environment. However, thanks to an increasing interest in LLPS in biological systems, a wide variety of short molecules forming coacervates have been found [11]. Since our interest in prebiotically relevant molecules we focused our study on short and simple peptides inspired from modern day proteins and RNA stretches made up by repetition of the same nucleotide.

We started studying phase behavior in the plane of peptides-salt concentrations showing the salt dependence of phase separation and qualitatively reproducing expected theoretical behaviors. Then we used the experimental results to evaluate accuracy of the coarse grained model employed, attempting to reproduce dependence on concentrations of phase separation. The first coacervate analysed was formed by WGR-4, a peptide 10 monomers in length (*WGRGRGRGWY*) [34] formed by repetition of *RG* domains with addition of hydrophobic head and tail (fig. 4.1).



**Figure 4.1:** Molecular structure of WGR-4 (WGRGRGRGWY).

It has been found in a study in which researchers tried to obtain coacervates from peptides inspired by *RG* low complexity domains driving coacervates in ribonucleoproteins and it constituted the shortest peptide found to phase separate along the study. Its structure is rich in hydrophobic aromatic residues constituted by tryptophan (*W*) and tyrosine (*Y*) and allows phase separation driven by segregative effects. In addition to that, the presence of positively charged arginine play a double role: for low ion strength the repulsive electrostatic force is expected to prevent WGR-4 molecules to stay close and the system to phase separates fig. 4.3 (e), whereas for high ion strength the screening effect explained in 1.1.2 decrease repulsion between molecules. In that case cations can interact with aromatic residues via cation- $\pi$  interactions, two aromatic rings can interact themselves via  $\pi - \pi$  stacking and the attractive contribution of segregative and short-range associative forces dominates leading to coacervate formation. It is then expected that an increase in salt concentration enhances coacervation fig.4.3 (d). The effect of pH on phase separation depends on the nature of molecules involved. In basic solutions the excess of negatively charged  $\text{OH}^-$  ions deprotonates positively charged molecules and renders them neutral. The effect of pH is then expected to be different in phase separation driven by hydrophobic or electrostatic effects. In WGR-4 coacervates, an increase in pH is expected to make positively charged arginine residues neutral and enhance segregative phase separation. Should be noticed that the effect of pH is not taken into account in simulations in which C and N termini are both neutral and charges of residues are those expected at neutral pH. We then decided to work in two nearly neutral conditions at pH 6 and 7.3. The second coacervate studied is heterotypical and is formed by two macro-molecular species: a 15 amino acids length peptide RP3 (*RRASLRRASLRRASL*) and poly uridine (*polyU*) of polydisperse sizes but much more longer than RP3 (fig. 4.2) [35]. We carried out experiments with polyU stock solutions in the range 300-1800 monomer units and, to have a sufficient number of polyU molecules in the simulated



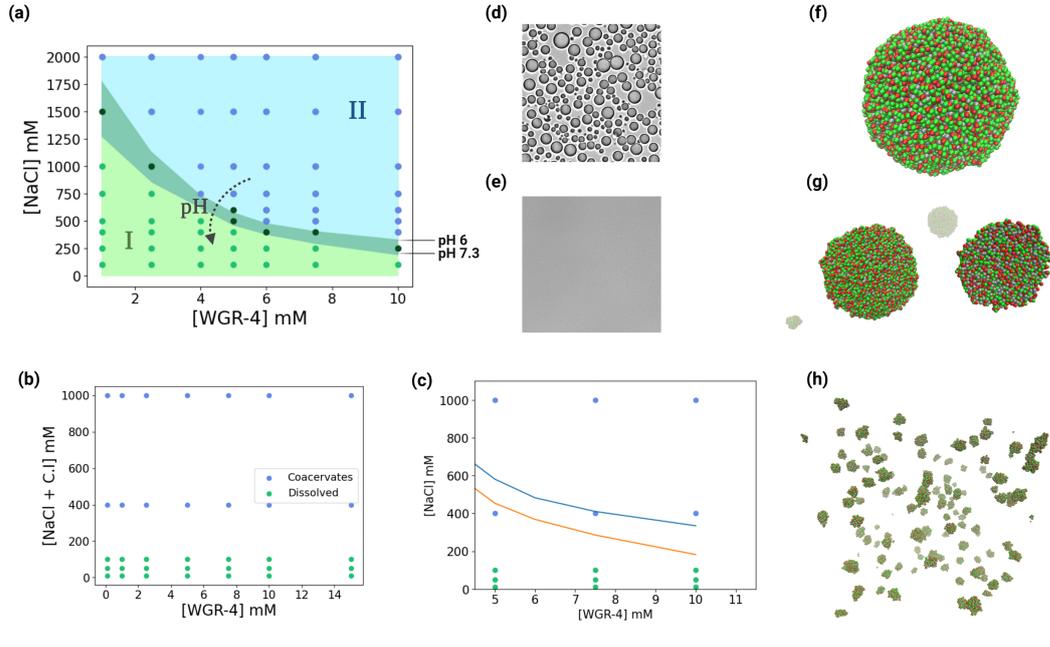
**Figure 4.2:** Molecular structures of RP3 (RRASL)<sub>3</sub> and Uridine.

box (30 to 150), we used 100 monomers length RNA stretches in simulations. The associative phase separation is driven by the net electrostatic attraction between positively charged arginine in RP3 molecules and negatively charged phosphate groups in the RNA backbone. However, similarly to the previous system, also cation- $\pi$  interactions between arginine and uracil as well as  $\pi - \pi$  stacking uracil-uracil may be present. Since coacervate is formed by associative phase separation, ions are now expected to dissolve coacervates since they screen the main attractive force between the two species. As in the first case pH is not involved as a simulation parameters and we carried out experiments at pH 7. Moreover in this case, since the combination of 1BPA-3BPN models in simulations, also inter-species cation- $\pi$  interactions are not involved and interactive forces between arginine and uracil are always repulsive.

#### 4.1.2 WGR-4 Phase Diagram

We started analysing a first set of samples through light microscopy to determine extreme values for phase diagram construction and to define a concentration window in which to analyse them. To reproduce the effect of both salt and pH, and create a phase boundary window that should enclose the simulated pH, we acquired phase diagrams at two different pH values, namely 6.0 and 7.3, in the peptide-salt concentration plane. As shown in fig.4.3 (a), the phase separation is favoured for high peptide and high salt concentrations (blue region) in which system forms coacervates as shown in fig. 4.3 (d). The dark green region in the graph represents the increase of the two phase region as a consequence of increasing the pH from 6.0 to 7.3. It shows how pH contributes to phase separation, shifting the binodal line to the bottom, enhancing coacervation at lower salt concentration. For the numerical fit of experimental results we used a 3<sup>rd</sup> order polynomial fit to qualitatively reproduce the shape of the theoretical phase diagram found in [36].

To explore the possibility of application of the 1BPA-3BPN coarse grained model to simulate segregative LLPS and subsequently apply it to partitioning



**Figure 4.3:** (a) Experimental phase diagram of WGR-4 coacervates at pH 6 and 7.3. Blue represents coacervates while green dissolved solutions. Dark green shows the shift of the binodal line increasing pH from 6 (top line) to 7.3 (bottom line). Dots show concentrations tested in-vitro. (b) Simulated phase diagrams of WGR-4 and (c) high density region in which experimental and simulated results agree. Blue and orange plots are experimental fits for pH 6 and 7.3 respectively. (d) WGR-4 coacervates and (e) dissolved solution. Simulated images shown the formation process of coacervates triggered by ions at 400 mM (f), 100 mM (g) and 10 mM (h)

experiments, we tried to reproduce the salt-peptide concentration dependence of phase separation. We started simulating ion-peptide concentrations in the region analysed experimentally and analysed the trajectories qualitatively observing that three main behaviors can be distinguished after the system is brought at room temperature and released at room pressure (Step 5 in 3.1.4) depending on ion concentration. For low ion strength (10 mM, 50 mM ions), that is low electrostatic screening, the system forms small clusters with size of order of tens of molecules that increase when WGR-4 concentration is increased, as shown in figure 4.3 (h). By analysing the cluster size distributions (fig. A.2 (a),(b)) it is possible to see that both size of clusters and number of molecules that form clusters are stable after the initial equilibration. This suggests that even for longer simulations these clusters will not merge, leading to the conclusion that in this situation coacervates are not stable, which means the system was labelled as dissolved.

In the large ion concentration region ( $\geq 400$  mM), the screening of electrostatic

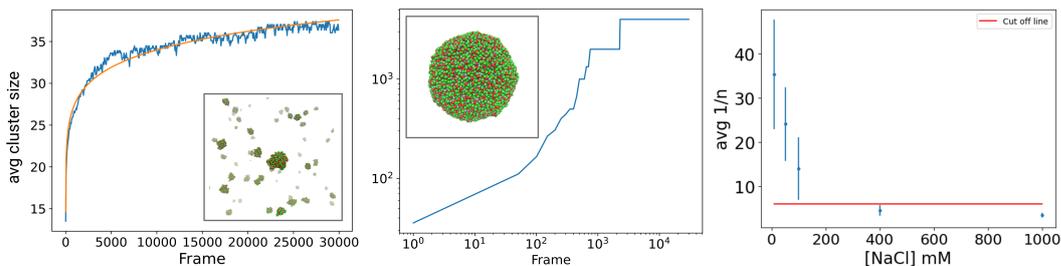
repulsion between arginine (+) is high enough that when molecules start to be closer in the NPT step (Step 3 in 3.1.4) then they are in a favoured condition that will rapidly bring the system to a fully phase separated state after the releasing step. This results in the formation of a single cluster containing all the molecules in the box. Figures A.2 (e) and (f) in the supplemental plots show the behavior of cluster distributions for 10 mM WGR-4 and 400 mM ions, corresponding to image 4.3 (f). In particular, A.2 (f) show that the phase separation and subsequent growth process is driven by the formation of an initial large cluster of 3700-3800 molecules and some small cluster that is rapidly absorbed by the larger one until all the molecules are in a single stable droplet that never dissolves or loses molecules. The growth process involve the first 2-3 thousand frames where a frame corresponds to 100 ns. For the remaining time no molecules are dissociated from the cluster suggesting that the system reaches equilibrium and in such condition the phase separation is stable. So we classified this case as coacervates.

The last case corresponds to an intermediate situation between dissolved and phase separated and emerges at 100 mM ions. At this ion concentration the system forms droplets composed by hundreds to thousands of molecules (fig. 4.3 (g)) that emerge from the NPT step and seem to be in equilibrium. Looking at the evolution of the cluster size distribution (fig. A.2 (c) and (d)), no growing processes involve large droplets except for some small clusters that merge along equilibration. Comparing the droplets on the left (surface) and on the right (core) in fig. 4.3 (g) (green residues are Gly, red Arg, grey Tyr and Trp), it is possible to see how droplets tend to distribute aromatic and hydrophobic residues in the core, leaving on the surface a shell of hydrophilic and positively charged residues. The emergence of a structure implies that aromatic residues don't participate in the interaction between droplets and the electrostatic repulsion between shells prevent coalescence.

To investigate the effect of the NPT step on the formation of such clusters, we decided to run a set of simulations without the step 3 in 3.1.4, using an initial box of size comparable with the final one at room pressure and running the NVT equilibration in point 2 for a longer time. We found that leaving the system free to evolve at room pressure and temperature still forms aggregates but made up by 3-400 molecules (10% of the total system). Interestingly, also in this case the clusters in the last equilibration are already formed at the firsts time steps and the only coalescence phenomena involves small clusters. This suggests a behavior more similar to simulations represented in 4.3 (h) rather than (f). We then classified this group of results as dissolved and got a simulated phase diagram shown in 4.3 (h), where dots represents simulated concentrations.

To enforce our classification, inspired by the scaling theory described in 1.1.1, we decided to investigate the growing behavior of droplets trying to extrapolate the characteristic exponent  $n$  of the average radius of clusters in the system (eq.1.4). Different exponents are related to different mechanisms governing growth. From

Berry et al., since neither flow nor gravity plays a role in simulations we expect to find a scaling law governed by an exponent  $n \in [\frac{1}{3}, 1]$  [14]. For the average cluster size we used the assumption that the size of the cluster is proportional to the number of molecules in the cluster itself, since molecules have the same volumes. So the volume, and thus the radius, is directly related to the amount of molecules in the cluster. By analysing average volumes plots, it is always possible to observe a region in which the growth process satisfies a power law over time. For small ions distribution, and in general simulations in which the final configuration contains more than one cluster, the average size continues to growth for the full simulated times, highlighting growing processes not evident from cluster distributions previously analysed. In the case in which the system reaches a fully separated state it follows that the growing process is arrested when all molecules are in the cluster, that is shown by the plateau in the central panel of 4.4. We used the power fit function in *CurveFit* package of Julia to extract the scaling exponents in the growing region using the full simulation time for the first class described and the time interval before reaching the plateau for the second. The first 15 frames were always removed to remove the effect of initial higher local concentration due to the NPT step. We found that only for simulations in which the full phase separation is achieved the extracted exponents are close to theoretical ones found in [14]. In 4.4 the panel on the left shows the average of the reciprocal exponents found averaging over different peptides concentrations. Simulations that fully phase separate show a maximum reciprocal exponent of 5.5 – 6 with an average of 4.5 for 400 mM ions (with a maximum of 5.3 for 1 mM WGR-4) and 3.5 for 1000 mM ions (with a maximum of 4.2 for 15 mM WGR-4). For 100 mM ions the average  $1/n$  is 14 with a minimum of 8.4 found for 10 mM WGR-4.



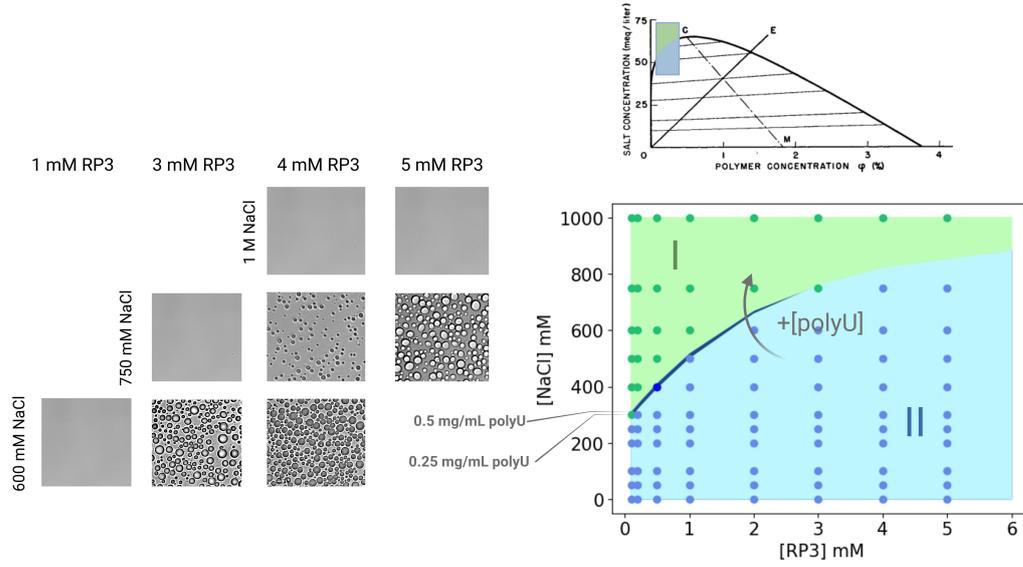
**Figure 4.4:** (Left) Power fit of average droplets volume for low ion concentration. (Center) average volume for high ion concentration. The last frames of simulations show the plateau reached when the system is fully separated. Simulation images in the boxes show the final frame of the related simulation. (Right) Plot of the reciprocal of the exponents of the average droplet radius. Cut off line (red) is 6. Averages and error bars are computed grouping simulations with same ion concentration

Even lacking in a good numerical match with theoretical predictions, we suggested to interpret the results as follow: when the system forms more than one droplets the average size continue to increase, however the increase is much slower than theoretically expected for any driving phenomena considered in [14]. This is because the growing process only involves small clusters that can eventually merge and increase the average size along simulation. However, the size of these clusters is small compared to the maximum cluster size (or sizes if more large clusters are involved as in 100 mM simulations) that is instead constant along time and makes the average cluster distributions almost constant. The resulting phase diagram obtained is then the same represented in 4.3 (b), the only effect of peptides concentration on phase separation seems to be slightly affect the average size of clusters for low salt concentration. The full phase separation and emergence of a single cluster, as well as the dynamic properties of growth seems to be more dictated by ions concentration than peptides. We then decided to carry out experiments and simulations for the second part of the project, involving host-guest systems, in the region of phase diagrams shown in figure 4.3 (c), in which numerical and experimental results are numerically comparable (In particular 7.5 mM WGR-4 and 400 mM ions).

### 4.1.3 RP3-pU Phase Diagram

The next system studied is the heterotypical RP3/polyU coacervate. Since it is formed by two macro-molecular species, analysing the phase behavior as a function of salt involves in principle the analysis of a 3D phase diagram in the salt-peptide-RNA space plus eventually the pH dependence. To simplify the analysis we decided to work at neutral pH, since it is the expected pH reproduced by charge distribution in MD simulations. We acquired section of the 3D phase diagram working first at fixed RNA concentration and variable peptides, then fixing peptides concentration and varying RNA. We found that phase separation is much more sensible to change in peptide concentration than RNA. As shown in fig. A.3 (right panel), the phase boundaries in the plane [salt]-[RNA] are almost constant in RNA concentration, the only region influenced by changing in amount of polyU corresponds to small RNA:peptide ratio ( $[\text{polyU}] < 0.25 \text{ mg/mL}$  and  $[\text{RP3}] = 4 \text{ mM}$ ).

In the [ions]-[peptides] plane shown in fig.4.5, the dependence on peptides concentration is more appreciable and it qualitatively reproduces the low concentration region of theoretical phase diagrams found with Voorn-Overbeek theory for associative phase separation in polyelectrolites solutions [17]. Accordingly with what is discussed in 1.1.2, we found that increasing the amount of ions in the system dissolves coacervates, as shown in phase diagram of fig.4.5. The effect of an increase in peptides concentration also increases the two phase region. The points on the boundary line we found, theoretically corresponds to the dilute phase density of

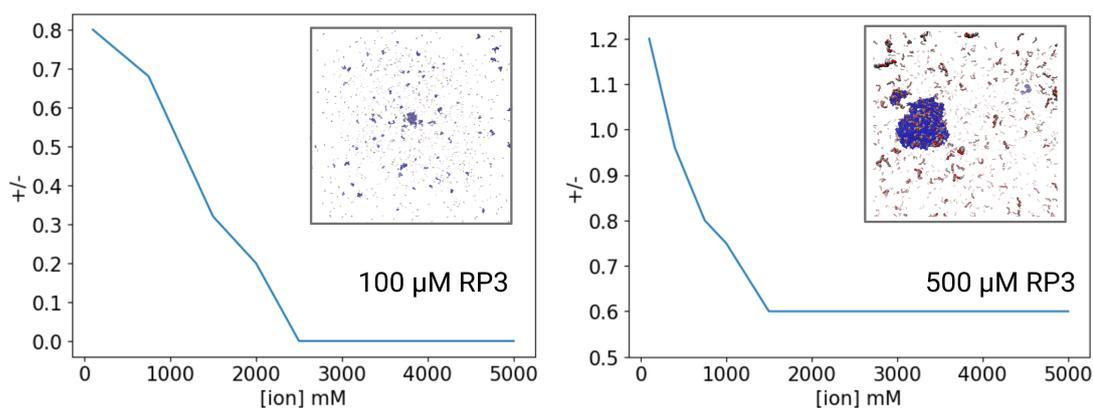


**Figure 4.5:** Top figure is adapted from [17]. It show phase diagram predicted by Voorn-Overbeek theory, in the polymer-salt concentration plane, for a symmetric mixture of water, positive and negative ions and positively and negatively charged polymers. Green-blue panel is a pictorial representation of region analysed in experiments. Plot on the bottom shown experimental phase diagram in the peptide-salt plane varying RNA concentration. Blue and green dots represent experiments conducted, the green region corresponds to mixed state, blue region to phase separation. Microscope images in the left shown solutions at different salt-peptides concentration. In the central line is possible to appreciate the effect on coacervates amount increasing concentration.

the system. This is true as long as the peptide concentration is smaller than that at critical point  $c_p^C$ , corresponding to the point in phase diagram above which phase separation never occurs. At the critical point dilute and dense phase have same concentration. Moving to the bottom, and so reducing salt concentration, the two phase region is enclosed between a low concentration ( $c_p, c_s$ ) and a high concentration ( $C_p, C_s$ ) point on the phase boundary. All the points on the line connecting ( $c_p, c_s$ ) and ( $C_p, C_s$ ) (*tie line*) correspond to concentrations for which the system splits in two phases, characterized by dilute and dense concentration  $c_p$  and  $C_p$  respectively. Tie lines are typically nearly parallel to the x axis [37] (and so  $c_s \approx C_s$ ), which means that an increase peptide concentration inside the two phase region, at almost constant salt concentration, is expected to produce systems in which dilute and dense phases always have concentrations  $c_p$  and  $C_p$  respectively. This translates into an increase in the volume of the dense phase that is well

shown by the 750 mM NaCl line of microscope images in 4.5.

Because of the small effect of polyU concentration on phase separation we decided to focus the numerical analysis on the [salt]-[peptides] plane working at 0.25 mg/mL polyU. We started simulating RP3 concentrations of 100  $\mu\text{M}$  and 500  $\mu\text{M}$ , running a first set of simulations from 10 mM to 1 M ions, expecting to dissolve both simulations. However, the 500  $\mu\text{M}$  case wasn't dissolved and also in the 100  $\mu\text{M}$  case the coacervates were not fully dissolved. We then decided to run a second set of simulations with same peptides-RNA amounts but ions in the range 1.5 M : 5 M. Results are shown in fig. 3.2. While 100  $\mu\text{M}$  seems to be dissolved by ions, 500  $\mu\text{M}$  simulations always present coacervates. Analysing the amount of molecules in the cluster shown in fig.A.4 for 100  $\mu\text{M}$  and in fig. 4.6 for 500  $\mu\text{M}$  RP3 and comparing them with trajectories analysis 3.2, it is possible to see how the main effect of ions is a release of RP3 molecules in solutions. The release is caused by screening of the attractive interactions between polyU and RP3, which are purely electrostatic in the model. This release, in the case of low RP3 concentration, leads to a complete dissolution of coacervates at 1.5-2 M ions. The amount of ions needed is larger than what expected from experimental results in fig.4.5 but qualitatively in agreement with observed behavior.



**Figure 4.6:** Positive:Negative charge ratio in RP3/polyU coacervates as a function of ion concentration. Left panel shows behavior for low RP3 concentration in which ions dissolves coacervates. Right panel shows behavior for high RP3 concentration in which, after a first release, the amount of RP3 stays stable in polyU droplets. In simulation images, representing the final snapshot of simulations, blue is polyU while small red (Arg) and silver (Ala, Ser, Leu) molecules are RP3.

In the case of high RP3 concentration, we found that ions never dissolve coacervates, also at 5 M concentration, that is close to the saturation concentration for NaCl in water (6.14 M). Since at the [peptide]-[RNA] involved in simulations we

expected to dissolve them around 0.5 M, we tried to use a different approach and look at the composition of the droplet formed. We found that as in the previous case an increase ions in solution releases RP3. However, the release seems to be balanced by the high amount of RP3 in solution. Considering that the interaction in 1BPA-3BPN model is almost purely electrostatic, neglecting hydrophobic effects, the amount of charges in the  $100\mu\text{M}$  simulation is  $150 \times 100 = 15\,000$  negative charges and  $6 \times 2\,000 = 12\,000$  positive charges that justify the full absorption of RP3 in the previous case. In addition the excess of polyU doesn't lead to dissociation of RNA molecules that are favoured in coacervates since  $\pi - \pi$  stacking. On the other hand, in the  $500\mu\text{M}$  simulation the number of RNA molecules is 30, leading to a total of only  $30 \times 100 = 3\,000$  negative charges. 4.6 shows the ratio between positive and negative charges in coacervates increasing salt concentration. Looking in particular at the right panel, corresponding to  $500\mu\text{M}$  RP3 we found that the amount of RP3 in coacervates initially decrease reaching 1/2 of the initial concentration forming a coacervates rich in polyU. The tendency of polyU to form cluster can be explained by hydrophobic and  $\pi - \pi$  interactions between nucleobases that are not screened by ions. Additionally increase ions in solution decrease repulsion between P groups on the RNA backbones favouring demix.

The effect of salt on RP3/polyU coacervates has been studied by Onuchic et. al. that found that using divalent salt ( $\text{MgCl}_2$ ) the droplets composition change increasing concentration [35]. For low salt concentrations an heterotypic RP3/polyU coacervates, that dissolves increasing salt, was found. Adding more salt to the solution they shown that homotypic coacervates composed by only polyU are formed in the area above the mixing region. We proposed this phenomenon as an explanation for coacervates found in our simulations. Looking again at fig. 4.6, the amount of RP3 in coacervates, after the first release up to 1.5 M ions, stays stable also triplicating ions. This suggest that the composition found will likely be stable and the system we are looking on could interpreted as an homotypic polyU coacervates sequestering RP3 molecules in solution, with the same principle of client partitioning discussed in the following.

Since the different behavior in coacervates found for small change in RP3 concentration it is difficult to get a phase diagrams for RP3/polyU such as done for WGR-4 in 4.3 (b). Looking at 4.5 the simulated systems corresponds to a small portion in the low RP3 concentration region. It is then expected that increasing RP3 in accord to experimental values will likely produce boxes with always more RP3 in solution. In addition, as discussed before, inter-species cation- $\pi$  interactions play no role in simulations. However, the high concentration of arginine and uracil in RP3/polyU coacervates, suggests that cation- $\pi$  attraction is important in phase separation for such polymers. A complete quantification of RP3/polyU ratio in coacervates necessitate further refinement to takes into account the contribution of non-Coulomb electrostatic interactions.

Since we didn't observed the heterotypic-homotypic switch in our experiments (conducted with monovalent salt NaCl), we decided not to proceed with other simulations of this system. In addition, since simulations involve no inter-species  $\pi$  interactions, and since these interactions are likely relevant for both formation of coacervates and their interactions with client molecules, we decided to proceed the numerical study using only WGR-4 coacervates.

## 4.2 Client Partitioning

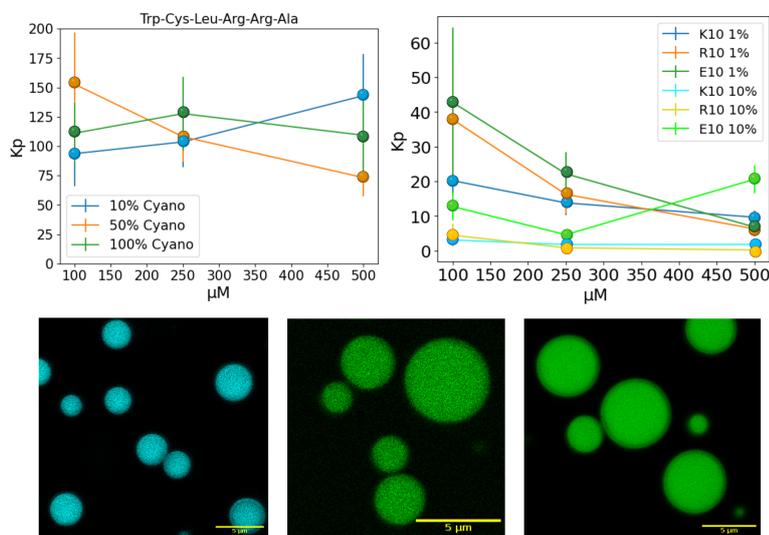
The last part of the project involved the study of client molecules in coacervates environments. To do that, we used a confocal technique to acquire 3D images of systems composed by the two coacervates described so far and small client molecules. Then we developed analysis tools to segment and extract data from the images acquired. We carried out experiments in both segregative and associative phase separation highlighting differences and similarity in partition properties of clients and we tried to reproduce features in MD simulations. Only clients in WGR-4 coacervates were analysed in simulations driven by motivations discussed in section 4.1.3.

### 4.2.1 Choice of Clients

Since the interest in prebiotically relevant molecules, we choose molecules for the partitioning experiments following the same principle of choice of coacervates, looking at short peptides and short RNA stretches constituted by a single nucleobase. The peptides involved in experiments are short positively charged peptides: Arg10 (charge +10, linear charge density = 1), Lys10 ( $q = +10$ ,  $\sigma = 1$ ) with same charge and density but constituted by different residues, (RRASL)<sub>3</sub> ( $q = +6$ ,  $\sigma = 0.4$ ), (RGRGG)<sub>5</sub> ( $q = +10$ ,  $\sigma = 0.4$ ), positively charged but with different lengths and a negatively charge peptides Glu10 ( $q = -10$ ,  $\sigma = 1$ ). In MD simulations we also analysed longer positively charged peptides, Arg30 and Arg100, and two more negatively charged peptides, Glu30 and Glu100. Molecular structures of all the peptides involved in experiments are shown in appendix A.1. In addition, the partitioning of NADH and polyA<sub>15</sub>, polyC<sub>15</sub>, polyU<sub>15</sub> has been studied.

Since the experimental analysis involved confocal microscopy, we needed a solution of client molecules containing a small percentage of molecules labelled with fluorophore dyes, as indicated in chapter 2. The practice of labelling biomolecules is commonly applied in biological studies involving the labelling of large molecules such as proteins. Since the size of proteins compared to those of dyes is typically large it is in general assumed that labelled and unlabelled molecules behave as the same. However, since molecules involved in our experiments are much smaller than proteins, and also dyes are formed by large structures involving many aromatic

rings that in principle can have favourable interactions with the environment in coacervates, we expected that in our case the presence of dyes could affect the properties of the molecules enough to have an influence on partitioning. We then decided to investigate behavior of clients involving a six monomers length peptide WCLRRA ( $q = +2$ ,  $\sigma = 0.33$ ) containing tryptophan and using as a labelled version of that molecule the same peptide with a modified version of tryptophan ( $\alpha$ Cyano-Tryptophan) that exhibits an excitation and emission maximum in the part of the spectrum that can be observed with a confocal fluorescence microscope. We used this molecules since Trp and CyanoTrp differs in a single atom we expected that the behavior of labelled and unlabelled peptides would be the same and so the ratio between them shouldn't affect the partitioning.



**Figure 4.7:** Partitioning of small peptides in WGR-4 coacervates. Left plot shows  $K_p$  of WCLRRA peptides varying concentration of CyanoTrp at 10%, 50%, 100%. Right plot shows partition coefficient of K10, R10 and E10 labelled at 1% and 10% with FAM showing how label concentration affect partitioning. Confocal images are from left to right WCLRRA 100% Cyano, K10 10% FAM and E10 10% FAM in WGR-4 coacervates. Scale bars are  $5 \mu\text{m}$ .

## 4.2.2 Effect of Fluorophore Dyes on Partitioning

The first result discussed is the influence of fluorophore dyes on partitioning properties of small peptides. We analysed our control system constituted by solution of WCLRRA and CyanoTrp-CLRRA in WGR-4 and RP3/polyU coacervates at three different percentages of CyanoTrp (10%, 50% and 100%). However, the 100% solution in RP3/polyU coacervates produced samples with zero intensity at full

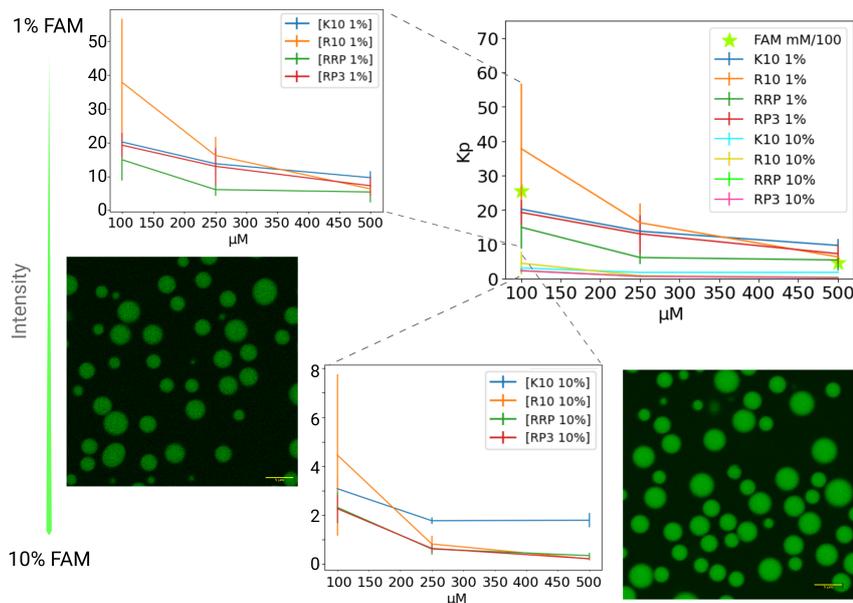
laser intensity. The acidic pH involved, or the high amount of arginine and polyU in solution seem to bleach CyanoTrp lading to sample without any signal. Lacking a control system on RP3/polyU, we decided to analyse effect of dyes restricted to WGR-4 coacervates.

As expected  $K_p$  of WCLRRA does not seem to be influenced by the amount of labelled molecules in solution (fig. 4.7 left plot). It is worth to notice that the weak fluorescence of CyanoTrp together with the larger acquisition window used to increase the acquired signal (see section 2.3) produces images with small intensities and an higher environmental noise. The large variances that arise from that make an exact investigation of numerical values of  $K_p$  difficult. However, a comparison with results obtained for FAM (discussed in the next paragraph, right plot of 4.7), suggests that differences in CyanoTrp partitioning are more likely due to variability of experimental systems and the techniques involved than its interaction properties. Then, changing the amount of dyes of a factor 10, we expect to see differences in  $K_p$  if dyes affect partitioning.

We analysed samples involving small peptides labelled at 1% and 10% with FAM and compared the results with control system. Results are shown in fig. 4.7 and 4.8. The behavior for positively charged peptides shown in fig. 4.8 seems to be largely affected by the amount of dye. Increasing the percentage of labelled molecules in solution from 1% to 10% results in a reduction of  $K_p$  of one order of magnitude, suggesting that dyes decrease the sequestering of these molecules. Stars in the top right panel of 4.8 show the partitioning coefficient of free dye when the concentration is the same of dye in a solution at 100  $\mu\text{M}$  and 500  $\mu\text{M}$  labelled at 1% and the results are comparable with those of labelled molecules at that concentration.

The behavior with FAM is different when is applied to the negatively charged peptide E10 (fig. 4.7 left plot). At low client concentration, dyes have a weaker effect on partitioning of E10 but it is still reduced. At high client concentration (500  $\mu\text{M}$ ), we observed an unexpected increase of  $K_p$ . The results have been tested on two different samples acquiring four different z-stacks for any sample and the small variance obtained averaging  $K_p$  over the images seems to confirm this result.

The other labels used in experiments are Cy3 and Cy5, involved in RNA's labelling. However, the high intensity of these dyes together with the high partitioning observed for RNAs in WGR-4 required to use a smaller percentage of labelled molecules with respect to FAM, no more than 0.1% Cy3 and Cy5. Since the dilution process involves errors, we decided to don't dilute them more than 0.1% and limit the analysis of dyes effects to FAM.

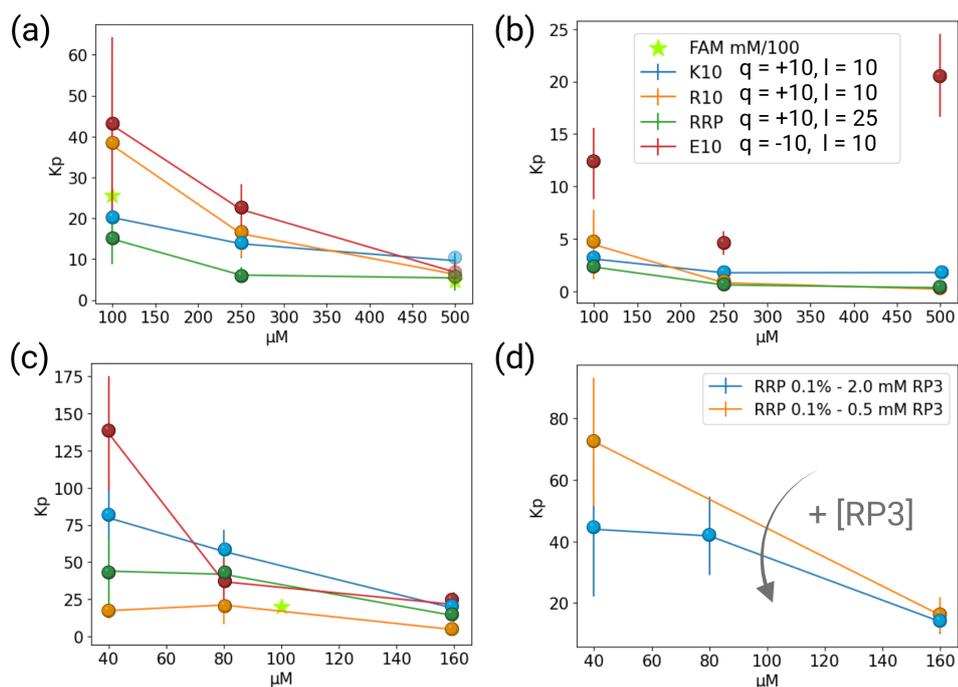


**Figure 4.8:** Effect of FAM on partitioning of positively charged peptides in WGR-4 coacervates. From top to bottom, label concentration is increased from 1% to 10% leading to a decrease in  $K_p$  of one order of magnitude. Plot on the left shows  $K_p$  of labelled molecules compared to those of free FAM in solution. Concentration for FAM are computed considering the amount of FAM in a solution at that concentration, labelled at 1%. Confocal images are R10 1% FAM (left) and R10 10% FAM (right) in WGR-4 coacervates. Scale bars are 5  $\mu\text{m}$ .

### 4.2.3 Partitioning of Small Peptides

Even if the influence of dyes discussed above has a non negligible effect on partitioning properties, since all peptides used were labelled with FAM, we decided to proceed the investigation of host-guest systems analysing differences in partitioning between molecules labelled with the same dyes. The first set of experiments involved a group of small positively charged peptides: Arg10 and Lys10, RP3 and RRP (see A.1). Results are shown in fig. 4.8 and 4.9(a) for WGR-4 and (c) for RP3/polyU.

To analyse results is important to keep in mind the conditions in which we performed experiments. With WGR-4, the amount of ions necessary to make the system phase separate at neutral pH, is such that droplets have the tendency to undergo rapid coalescence. This leads to the formation of large regions of dense phase in the system instead of droplets, making the acquisition process difficult A.6. We then carried out experiments with WGR-4 systems working at basic pH (11-12), at which the system phase separates without adding salt. At this pH both arginine and lysine will be mainly deprotonated and so neutral. However, the  $pK_a$  (the



**Figure 4.9:** Results of partitioning experiments of short peptides in WGR-4 (a), (b) and RP3/polyU (c), (d) coacervates. (a) K10, R10, RRP, E10 labelled at 1% with FAM in WGR-4. (b) same peptides in WGR-4 coacervates nut labelled at 10% with FAM. (c) Peptides labelled at 1% FAM in RP3/polyU coacervates. (d) Effect of change in coacervate composition on partitioning of RRP.

values of pH at which 50% of a molecular species is deprotonated) of side chain of arginine (12.48) is higher than lysine (10.43). Then for peptides with equal amount of arginine and lysine we expect a higher charge for the first. It should be also noted that WGK-4 peptides, that has the same structure of WGR-4, but with lysine in place of arginine, have been shown to not phase separate, suggesting arginine play a key role in interactions with aromatic rings in the structures of tyrosine and tryptophan [34]. Fig. 4.9 (a) shows for small client concentrations (100  $\mu\text{M}$  and 250  $\mu\text{M}$ ) Arg10 has a higher partition coefficient in WGR-4 coacervates than Lys10, enforcing what discussed so far. The second difference between arginine and lysine behaviors is in the effect of client concentration. Increasing amount of client in solution has a small effect on lysine partitioning compared with the large decrease in  $K_p$  for arginine, for which client concentration and  $K_p$  seems to be inversely proportional in this region. This could suggest a saturation effect for Arg10 due to the large contribution of clients molecules to the total amount of positive charges in the system. The behavior of lysine partitioning is instead nearly constant, also

lysine is less hydrophilic than arginine, suggesting the partitioning of lysine at this pH is mainly driven by hydrophobic-hydrophilic interactions.

The effect of size on  $K_p$  in WGR-4 seems to be the opposite of what expected from (1.16) for which size should increase partitioning. Results for system composed by arginine of length 10 (Arg10), 15 (RP3) and 25 (RRP) were analysed and compared in fig. A.5. Except for high concentrations, increasing size seems to decrease partitioning. However, molecules used have also different charges and charge densities and a full characterization of size and charge dependence need further investigation. Comparison between top and bottom plots of 4.8 also shows that even lacking in an absolute meaning of numerical  $K_p$  computed, the main features of partitioning are preserved at different label concentrations, suggesting they are proper of the molecules involved.

RP3/polyU experiments were instead done working with nearly neutral pH (6-7). In these conditions arginine and lysine will be mainly protonated and so positively charged, leading to strong electrostatic interactions with both positive (Arg in RP3 peptides) and negative (P in uridine backbone) charges in coacervates. Fig.4.9 (c) shows results for molecules labelled with 1% FAM. In this case lysine seems to have an higher  $K_p$  than arginine.

To better understand the effect of electrostatic interactions we analysed systems with Glu10 (fig. 4.9 a, b, c). Negatively charged molecules seems to have the higher partition coefficient in both coacervates. Since WGR-4 coacervate is expected to be positively charged, the partition of Glu10 suggests a central role of Coulomb attraction in client sequestration. At low concentrations  $K_p$  is significantly larger in RP3/polyU coacervates where the neutral pH is such that arginine are charged and can interact more with glutammic acid. Interestingly, in both coacervates the sequestering of Glu10 shown the faster decreasing when client concentration increases.

Two additional observation can be taken into account trying to explain the decreasing of  $K_p$  observed. Fig. 3.2 shows how partitioning of peptides change in RP3/polyU coacervates at different peptides concentrations. Even if in this case peptides are involved in coacervates formation, the similarity in interactions and structure between client and coacervates molecules, suggests that partitioning behavior can be ruled by the interplay of host and guest in formation of coacervates. We then analysed partitioning of RRP molecules in RP3/polyU coacervates at two different RP3 concentrations and we found support for our hypothesis, observing that increasing RP3 in coacervates decreases the partitioning of RRP. In addition, we performed an experiment with WGR-4 molecules in a region of the phase diagram in which usually doesn't phase separate. We added Arg10 molecules to a solution of WGR-4 at high pH but concentration lower than the critical one at which phase separation occurs. Then we added Arg10 molecules to bring the total concentration of WGR-4 and Arg10 just above the critical line (considering Arg10

as WGR-4 in concentration) and we observed phase separation.

### Numerical Simulations

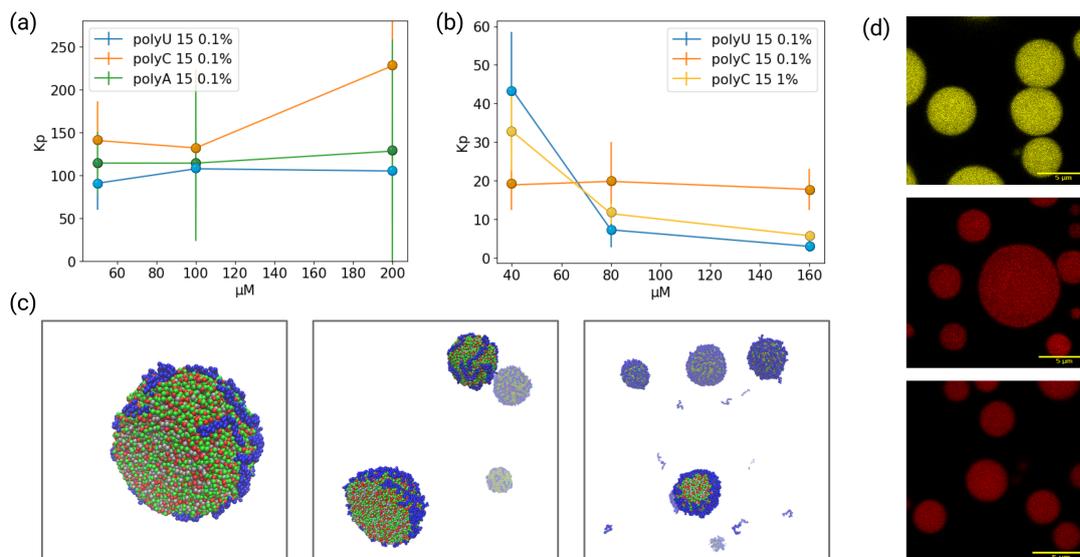
Since arginine and lysine peptides have same charge in 1BPA model involved in simulations and both have vanishing hydrophobicity, it is expected that poly-Arg and poly-Lys behave equally in the model. We thus analysed only arginine based peptides. Fig. A.7 left panel shows partition of R10 (500  $\mu\text{M}$ ) and results are the same for RP3 and RRP. Both are fully sequestered by coacervates and partitioned mainly in the core, well mixed with WGR-4 molecules. This leads to infinite  $K_p$  for almost all positive charged molecules tried in simulations (see R30 250  $\mu\text{M}$  in central panel of fig. A.7). Moreover, from left to right, fig. A.7 shows that increasing size (Arg10, Arg30 and Arg100) destabilizes coacervates leading to irregularly shaped droplets and/or aggregates.

We then simulated systems with opposite charged molecules (Glu10, Glu,30, Glu100) and results are shown in A.8. It is possible to see that negatively charged molecules are always partitioned on the surface, that is also confirmed by simulations with RNA's discussed in next chapter. Because the surface partitioning, first the droplet is divided in more smaller droplets to increase the available surface to allocate client molecules. Then, further increasing in client concentration leads to surfaces saturation leaving some molecules in solution and thus producing a finite  $K_p$ . However, as it is possible to see in the right image of fig. 4.8, showing partition of E10 in WGR-4, the effect is not observed in-vitro.

#### 4.2.4 Partitioning of Small RNAs

To conclude the analysis of client partitioning in our coacervates systems we conducted experiments with short stretches of RNAs composed by 15 repetition of the same nucleotide (U, C, A). Results of partitioning are shown in fig. 4.10 for WGR-4 (a) and RP3/polyU (b). The negative charge carried by phosphate groups on the backbone can be attracted by electrostatic interactions with positive charges in coacervates. In addition to that, the nitrogen bases are aromatic and can interact via cation- $\pi$  and  $\pi - \pi$  interactions with positively charged or aromatic residues respectively. Since the presence of both charged and aromatic molecules in our coacervates we expected RNA molecules to be well partitioned with high  $K_p$ .

In WGR-4 coacervates, comparing fig. 4.10 (a) with fig. 4.8, it seems that RNAs is in general more partitioned in the dense phase than any peptide involved in experiments. Comparing in particular results with Glu10, it shows that RNAs are more sequestered than negatively charged peptides, suggesting that the presence of aromatic rings in nucleobases plays an important role in partition properties. In addition, electrostatic interactions are expected to be unfavoured by deprotonation



**Figure 4.10:** Results of partitioning of short RNAs in WGR-4 and RP3/polyU coacervates. (a)  $K_p$  of polyU(Cy3), polyC(Cy5) and polyA(Cy5) labelled at 0.1% in WGR-4. (b)  $K_p$  of polyC 0.1% Cy5 and polyU labelled with Cy3 at 0.1% and 1%, showing a weak effect of Cy3 on polyU's  $K_p$ . (c) Final snapshots of MD simulations for WGR-4 coacervates and polyA at 100  $\mu\text{M}$ , 200  $\mu\text{M}$  and 500  $\mu\text{M}$  showing surface partitioning of RNAs, droplets separation to maximize the surface and subsequent surface saturation. (d) Confocal images of RNAs in WGR-4 coacervates. From top to bottom polyU(Cy3), polyC(Cy5) and polyA(Cy5).

of positive charged molecules because of the basic environment. The partitioning of RNAs in WGR-4 seems to contradict the decreasing behavior in concentration shown by other molecules. Increasing the amount of RNA in solution leads to constant or higher  $K_p$ , showing an higher tendency of molecules containing aromatic rings to partition than purely electrostatically interacting molecules. Following the idea mentioned in the previous section of an interplay of host and guest in phase separation for not too small concentrations, we analysed the possibility to enhance coacervates formation with addition of clients as done for Arg10 in WGR-4 coacervates. We performed an experiment at neutral pH, with zero salt added and 1 mM WGR-4 that as shown in fig. 4.3 (a) is expected to stay mixed. Then, inspired by RP3/polyU phase separation and by the high partitioning observed for RNAs, we added to the solution polyU in concentrations comparable to those at which RP3 forms coacervates and we observed phase separation. In addition, droplets observed were smaller and undergo much slower coalescence than coacervates formed by only WGR-4.

In RP3/polyU coacervates, the results of peptides and RNAs partitioning are

instead comparable (fig. 4.9 (c) and fig. 4.10 (b) respectively). In both cases is observed a decreasing in  $K_p$  increasing client concentration. We carried out experiments using only polyU and polyC because small concentrations of polyA make the system aggregates. The similarity in partitioning of positive charge peptides (such as RP3) and RNA (such as polyU), together with the effect of RP3/polyU ratio on partitioning (fig. 4.9 (d)), could suggest that partitioning in such coacervates is ruled by the balance between the two components forming coacervates and sequestered material could participate in this balance affecting the very properties of sequestering.

One observation could be done looking at numerical values obtained in experiments, is the influence of dyes, that in case of RNAs are only the 0.1%. Considering the influence discussed in 4.2.2 the higher  $K_p$  could also be looked as an effect of low dyes. We didn't verified the effect of such dyes in WGR-4 but only in RP3/polyU where fig. 4.10 (b) shows that results are comparable for dyes concentration different of a factor 10. In addition, Cy3 and Cy5 are formed by aromatic rings disposed in a less rigid structure than FAM and also, aromatic rings are already present in RNAs. All this leads us to the assumption that RNAs are much less affected by Cy3 and Cy5 dyes than how hydrophilic peptides are affected by FAM.

### Numerical Simulations

We simulated WGR-4 coacervates with RNA stretches of 15 nucleotides (U, C, A, G) at 100  $\mu\text{M}$ , 200  $\mu\text{M}$  and 500  $\mu\text{M}$ . Similarly to Glu10, Glu30 and Glu100 simulations, molecules are partitioned on the surface of coacervates at any concentration and for any nucleobase. In figure 4.10 (c), snapshots of the last simulation frame are presented, showing a behavior comparable with those found for negatively charged peptides (fig. A.8). First the increase in client concentration (from left to right 4.10 (c)) reduce droplets size to increase the total surface available to allocate molecules. Then, further increase leads to a saturation of the surface and molecules are found in solution, making  $K_p$  finite. Comparing left and central panel of fig. A.8 (Glu10 500  $\mu\text{M}$ ) with right panel of fig. 4.10 (c) the repulsion on the surface between RNA molecules seems to be smaller leading to a more compact RNA shell around WGR-4 coacervates and, as a consequence, an higher partitioning of RNA than GluN (with  $N = 10, 30, 100$ ). However, the surface partition of RNAs, as well as for Glu10, is not observed experimentally, as shown in fig. 4.10 (d) that from top to bottom shows partition of polyU, polyC and polyA in WGR-4 coacervates.



## Chapter 5

# Conclusions and Outlook

The phase diagram of two coacervate systems and their interactions with small client molecules were investigated by numerical simulations and experimental work, with focus on prebiotically relevant molecules. We first investigated the phase behavior of homotypic WGR-4 and heterotypic RP3/polyU coacervates in-vitro by looking at the influence of salt and pH on phase separation.

We qualitatively reproduced the theoretical phase behavior of WGR-4 coacervates in the peptide-salt concentration plane by applying a 1BPA-3BPN coarse grained model in MD simulations. We numerically reproduced the effect of ions on WGR-4 coacervates showing that salt enhances phase separation. A salt-peptide concentration window in which numerical simulations agree with experiments has been found, which was used to investigate partitioning of small molecules.

We experimentally investigated RP3/polyU phase diagrams in the peptide-salt concentration plane, highlighting the effect of salt on the dissolution of such coacervates by screening electrostatic charges. We reproduced the theoretical phase diagrams for associative phase separation found with Voorn-Overbeek theory and tried to simulate behavior with MD simulations. Two different behaviors have been found for high and low peptide concentrations working at constant polyU, showing ions dissolving coacervates for low RP3. In addition, for high RP3 concentrations, results could be interpreted as an heterotypic-homotypic switch, where complex peptide/RNA coacervates turn slowly into homotypic RNA coacervates, similar to experimental observations by Onuchic et al [35].

In the second part of the project, we analysed interactions of coacervates and small client molecules composed by short charged peptides or short RNA stretches. We developed an automated analysis tool for the segmentation and analysis of z-stacks, allowing computation of  $K_p$  for single droplets without any manual intensity thresholding. We performed experiments with two coacervates systems and we found that coacervates made up by small and simple peptides or RNAs can sequester a wide variety of prebiologically relevant molecules and, with some exception, phase

separation is not destabilized by clients. We investigated the interplay of host and guest in coacervates formation and client sequestration and we found that client can enhance phase separation and that for not too small concentrations host and guest roles are not always distinct. Both peptides and RNAs has been shown to enhance phase separation of WGR-4 and RRP partitioning was found dependent on RP3/polyU coacervates composition.

We then simulated systems with WGR-4 and client molecules in the region of phase diagrams found to be in agreement with experimental behavior, and we found that the charge sign rules the partition of molecules, leading to surface partition for negatively charged molecules and core partition for positive charges. The surface partition lead to a finite  $K_p$  for negative molecules due to surface saturation, whereas  $K_p$  of positively charged peptides is always infinite due to a full sequestration. In addition, increase size of positively charge molecules seems to destabilize coacervates producing aggregates.

However, phenomena such as surface partition or droplets size reduction increasing concentration, have been only observed in simulations and have no experimental support. Comparing the size of a ten monomer peptide in simulation with the size of a droplet, it's clear that the relation is very different from what expected in a real system, as the droplets in reality are orders of magnitude larger than single molecules. A step could be to experimentally investigate the partitioning of longer negatively charged peptides to verify a possible surface partition and have insight about other approaches to understand simulations results.

In addition, the absence of any dependence of WGR-4 phase diagrams on peptide concentration and the poor numerical and qualitative agreement with experiments for both coacervates systems analysed, suggest that a refinement of parameters, and eventually inter-species 1BPA-3BPN interactions, should be done to fully reproduce features of LLPS. Given the central role of interactions involving aromatic beads in coacervate formation, the incorporation of inter-species cation- $\pi$  and  $\pi - \pi$  interactions into the model, followed by a subsequent fine-tuning of parameters through experimental fitting, establishes coarse-grained MD simulations as a pivotal tool for investigating coacervate systems. Furthermore, the ability of 1BPA-3BPN MD simulations, including implicit ions and solvent descriptions, to predict the salt dependency of homotypic and heterotypic phase separation holds promise for future studies. Andreev et al. have previously discussed the agreement between experimental and numerical phase diagrams for a range of oppositely charged polymers, utilizing coarse-grained simulations with explicit ions and solvent [38]. However, the computational demands of including beads to simulate the solvent become restrictive, limiting simulations to small systems. By employing an implicit ion and solvent model, we focus exclusively on simulating coacervate molecules, enhancing our resolution of molecular-level interactions.

The possibility of enhancing phase separation with client molecules of different

species suggests that in a prebiotic environment, the emergence of coacervate protocells could result from inter-species cooperation, thus reducing the constraints on the emergence of such compartmentalized systems. This concept carries implications for both biological and prebiotic life studies, offering insights into these interactions and the phase behavior of coacervates in the presence of clients. Building upon the results discussed, an investigation of the phase diagram in the peptide-RNA-salt space for WGR-4/polyU coacervates would provide valuable insights into the phase behavior of WGR-4 within a more complex environment. This could lead to emergent phenomena arising from the dual ability of both WGR-4 and polyU to self-separate into complex or simple coacervates depending on ion.

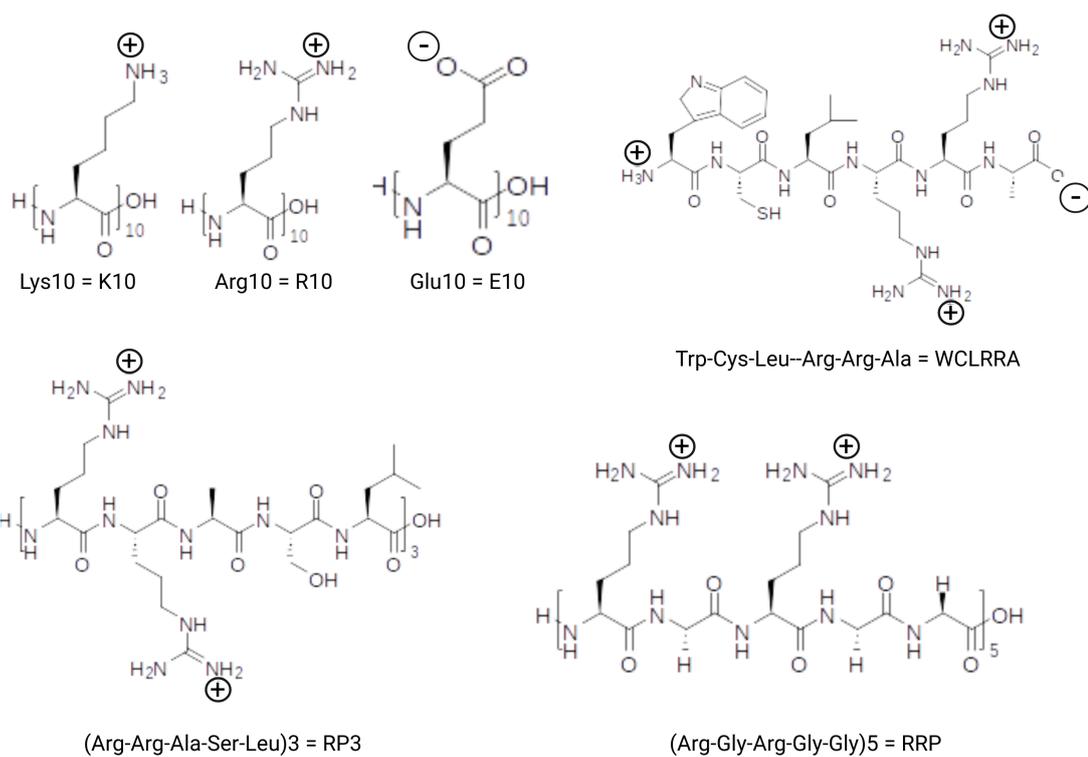
To conclude, also the experimental technique involved, even offering valuable insights on coacervates systems analysed, has some limitations. The most important limitation discussed along the study was the labelling of molecules with fluorophore dyes and the consequence on partition properties. It is important to acknowledge that more precise results can be achieved using state of the art techniques such as Nuclear Magnetic Resonance (NMR). However, it requires a considerably larger amount of material for any sample analysed, in the order of tens to hundreds times the amount needed with confocal microscopy, resulting in a much more expensive research process. It should be a good control system however, having some pair coacervates-clients also analysed with this technique, to verify results and better quantify both partition of molecules and effect of dyes.

All in all, the numerical and experimental results presented here improve our understanding of client partitioning and could be valuable for future studies.



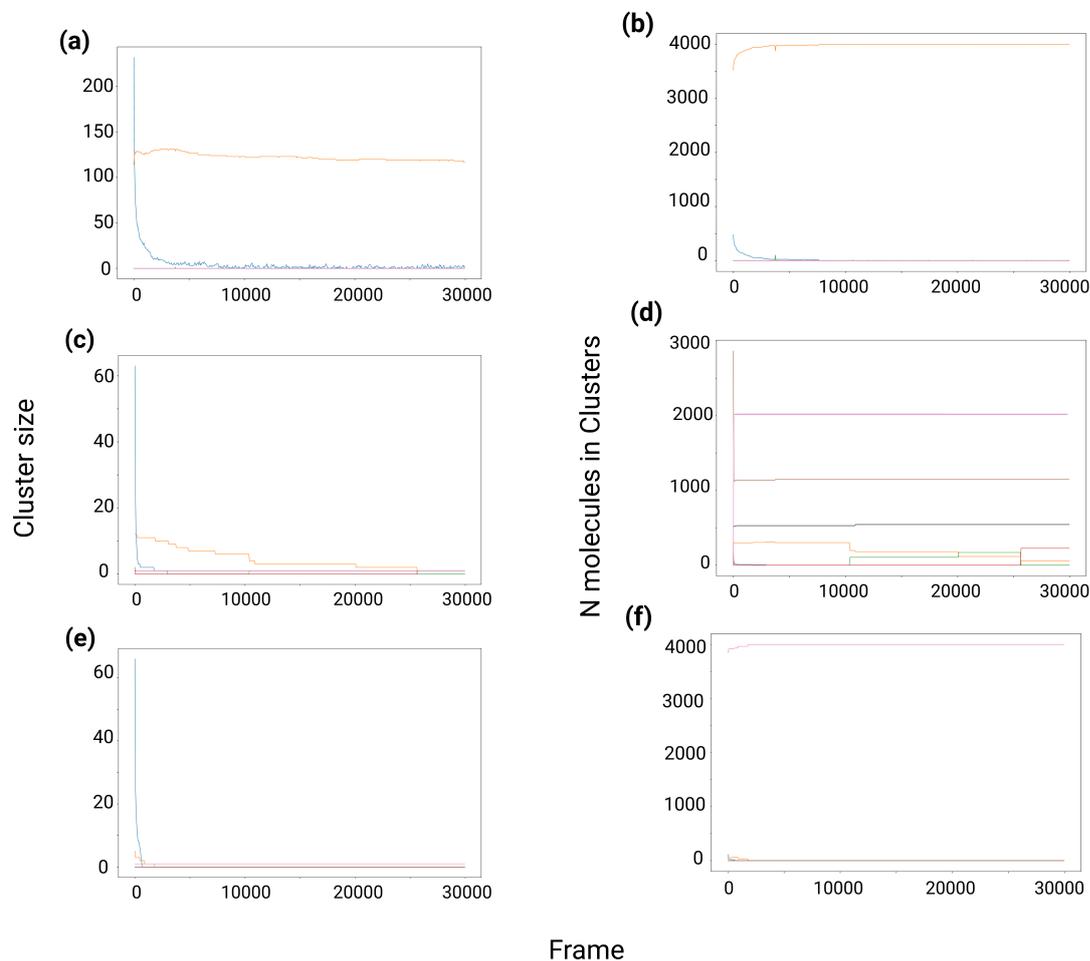
# Appendix A

## Supplementary Plots and Figures



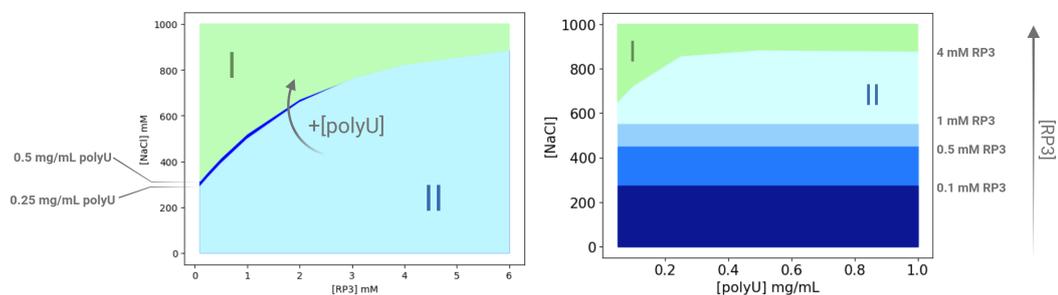
**Figure A.1:** Molecular structures of peptides involved in partitioning experiments. Numbers refer to repetition of the elementary units represented.

## A.1 WGR-4

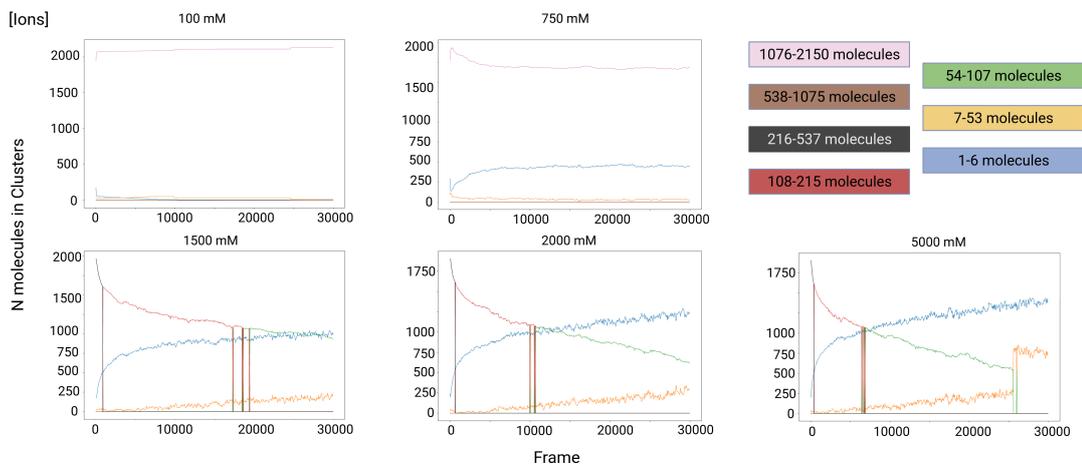


**Figure A.2:** Cluster sizes in WGR-4 simulations. Left column are cluster size distributions over time, right row are number of molecules in clusters, grouped by cluster size, over time. All simulations are 10 mM WGR-4, ion concentration increases from top to bottom. Top row is 50 mM ions, central row is 100 mM ions, bottom row is 400 mM ions. Legend: 1-6 (Blue), 7-100 (Orange), 101-200 (Green), 201-400 (Red), 401-1000 (Black), 1001-2000 (Brown), 2001-4000 (Pink).

## A.2 RP3-pU

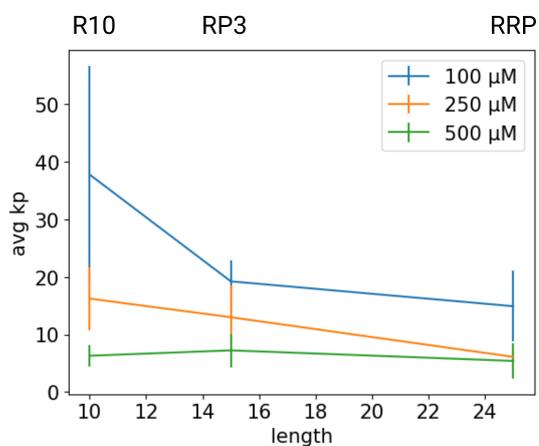


**Figure A.3:** Experimental phase diagrams for RP3/polyU coacervates. Coloured regions corresponds to polynomial fit of experimental data. Plot on the left shows how phase diagram in the [peptide]-[salt] plane change changing polyU concentration (0.25 mg/mL and 0.5 mg/mL). Plot on the right shows the phase diagram in the [RNA]-[salt] plane for four different concentrations of RP3 (0.1 mM, 0.5 mM, 1 mM and 4 mM) evidencing a constant behavior except for the 4 mM case.

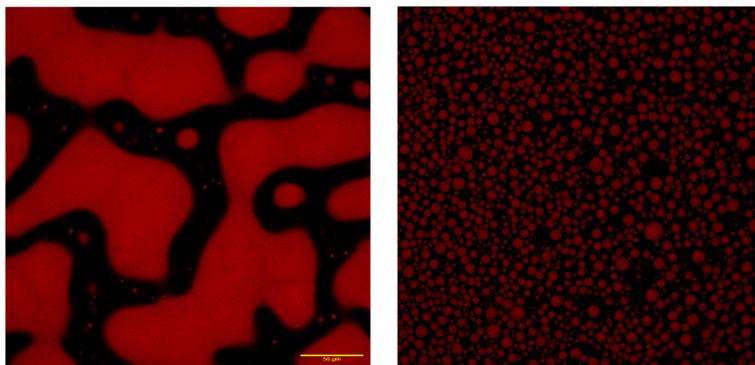


**Figure A.4:** Number of molecules in clusters, grouped by cluster size, over time. One frame corresponds to 100 ns.

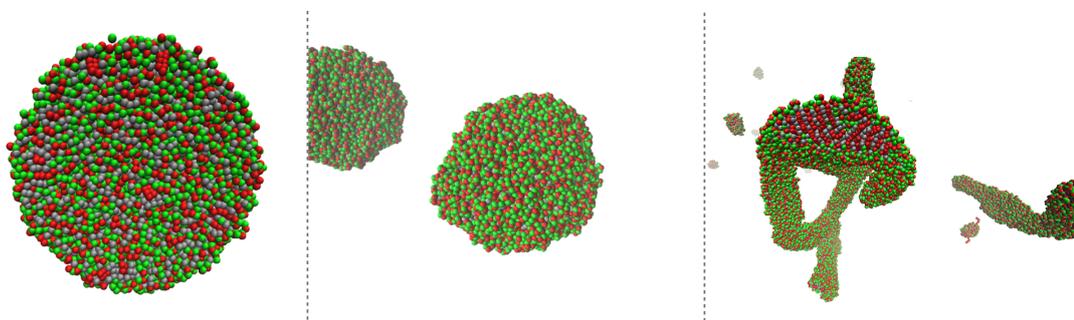
### A.3 Client partitioning



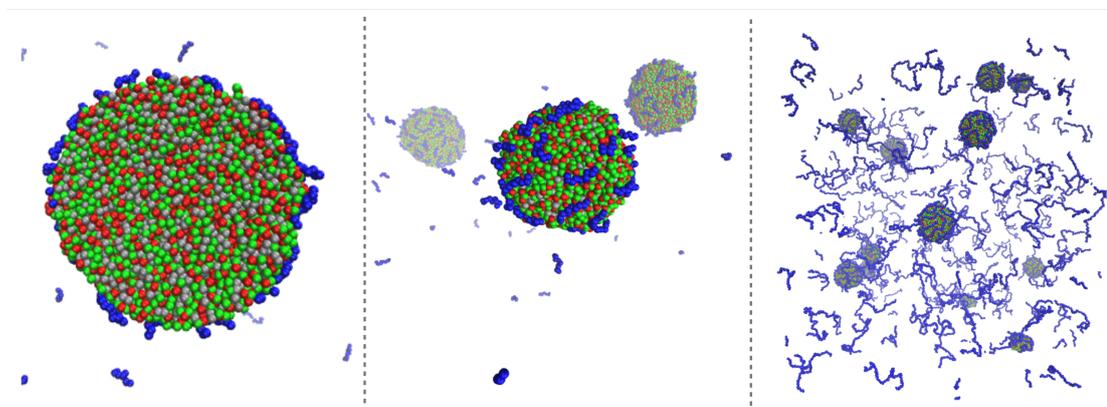
**Figure A.5:** Size dependence of partitioning in WGR-4 coacervates showing a reduction of  $K_p$  increasing client size. Experiments were performed at 7.5 mM WGR-4 and 40 mM NaOH.



**Figure A.6:** Left figure shown WGR-4 coacervates at 7.5 mM WGR-4 and 1 M NaCl, right figure at 7.5 mM WGR-4 and 40 mM NaOH. Both systems contains 100  $\mu\text{M}$  of polyC 0.1% Cy5. Scale bar is 50  $\mu\text{m}$ . Images were taken after a comparable amount of time.



**Figure A.7:** Partition of positively charged peptides in MD simulations. From left to right, R10 500  $\mu\text{M}$ , R30 250  $\mu\text{M}$ , R100 250  $\mu\text{M}$ . Arg (red), Tyr and Trp (silver) are preferred inside the coacervates in the solvent "poor" phase, forming a compact core that destabilize spherical shape increasing size. The shell is always mainly constituted by Gly (green) and Arg (red) energetically favoured by exposition to the solvent.



**Figure A.8:** Partition of negatively charged peptides in MD simulations. From left to right, E10 500  $\mu\text{M}$  section, E10 500  $\mu\text{M}$ , E100 500  $\mu\text{M}$ . Arg (red), Tyr and Trp (silver), Gly (green), Glu (blue). Negatively charged molecules are always partitioned on the surface.

**Table A.1:** Experimental Partitioning coefficients of small molecules in WGR-4 coacervates. Experiments were taken at 7.5 mM WGR-4 and 40 mM NaOH.

| Client         | mM   | K <sub>p</sub> | mM   | K <sub>p</sub> | mM   | K <sub>p</sub> |
|----------------|------|----------------|------|----------------|------|----------------|
| polyU 0.1% Cy3 | 0.05 | 90.6 ± 31.0    | 0.10 | 107.4 ± 46.9   | 0.20 | 104.9 ± 57.3   |
| polyC 0.1% Cy5 | 0.05 | 140.4 ± 45.6   | 0.10 | 131.8 ± 92.2   | 0.20 | 227.7 ± 213.1  |
| polyA 0.1% Cy5 | 0.05 | 114.1 ± 36.1   | 0.10 | 114.0 ± 90.5   | 0.20 | 128.1 ± 130.2  |
| RP3 1% FAM     | 0.10 | 19.2 ± 3.7     | 0.25 | 13.0 ± 5.5     | 0.50 | 7.2 ± 2.9      |
| RP3 10% FAM    | 0.10 | 2.6 ± 0.6      | 0.25 | 13.0 ± 5.5     | 0.50 | 7.2 ± 2.9      |
| RRP 1% FAM     | 0.10 | 14.9 ± 6.2     | 0.25 | 6.1 ± 1.9      | 0.50 | 5.4 ± 3.1      |
| RRP 10% FAM    | 0.10 | 2.3 ± 0.7      | 0.25 | 0.6 ± 0.2      | 0.50 | 0.3 ± 0.1      |
| R10 1% FAM     | 0.10 | 37.8 ± 18.9    | 0.25 | 16.2 ± 5.5     | 0.50 | 6.3 ± 1.9      |
| R10 10% FAM    | 0.10 | 4.5 ± 3.3      | 0.25 | 0.8 ± 0.4      | 0.50 | 0.2 ± 0.1      |
| K10 1% FAM     | 0.10 | 20.2 ± 4.1     | 0.25 | 13.8 ± 3.6     | 0.50 | 9.6 ± 1.9      |
| K10 10% FAM    | 0.10 | 3.1 ± 0.5      | 0.25 | 1.8 ± 0.1      | 0.50 | 1.8 ± 0.3      |
| E10 1% FAM     | 0.10 | 42.7 ± 21.6    | 0.25 | 22.2 ± 6.2     | 0.50 | 6.8 ± 2.7      |
| E10 10% FAM    | 0.10 | 12.7 ± 4.0     | 0.25 | 4.6 ± 1.1      | 0.50 | 20.6 ± 4.0     |
| NADH           | 0.10 | 155.0 ± 44.3   | 0.25 | 107.5 ± 40.9   | 0.50 | 77.9 ± 18.9    |
| WCLRRA 10% CN  | 0.10 | 93.6 ± 27.6    | 0.25 | 103.6 ± 21.6   | 0.50 | 143.0 ± 35.2   |
| WCLRRA 50% CN  | 0.10 | 152.9 ± 44.0   | 0.25 | 108.1 ± 21.7   | 0.50 | 73.2 ± 15.9    |
| WCLRRA 100% CN | 0.10 | 111.0 ± 25.7   | 0.25 | 127.7 ± 31.2   | 0.50 | 109.2 ± 35.1   |

**Table A.2:** Experimental Partitioning coefficients of small molecules in RP3/polyU coacervates. Experiments were taken at 2 mM RP3 and 1 mg/mL polyU. \* Low RP3 concentration (0.5 mM RP3 and 1 mg/mL polyU).

| Client         | mM   | K <sub>p</sub> | mM   | K <sub>p</sub> | mM   | K <sub>p</sub> |
|----------------|------|----------------|------|----------------|------|----------------|
| polyU 0.1% Cy3 | 0.04 | 43.73 ± 14.85  | 0.08 | 7.26 ± 4.50    | 0.16 | 2.97 ± 1.42    |
| polyC 0.1% Cy5 | 0.04 | 18.91 ± 14.85  | 0.08 | 19.82 ± 10.30  | 0.16 | 17.74 ± 5.37   |
| RRP 1% FAM     | 0.04 | 17.33 ± 5.68   | 0.08 | 21.10 ± 12.65  | 0.16 | 4.53 ± 1.38    |
| * RRP 1% FAM   | 0.04 | 72.46 ± 20.76  | 0.08 | N.A.           | 0.16 | 16.06 ± 5.71   |
| R10 1% FAM     | 0.04 | 17.33 ± 5.68   | 0.08 | 21.10 ± 12.65  | 0.16 | 4.53 ± 1.38    |
| K10 1% FAM     | 0.04 | 79.78 ± 27.96  | 0.08 | 57.16 ± 14.77  | 0.16 | 18.83 ± 5.19   |
| E10 1% FAM     | 0.04 | 136.77 ± 38.35 | 0.08 | 36.80 ± 15.63  | 0.16 | 21.48 ± 8.34   |
| NADH           | 0.04 | 42.80 ± 11.75  | 0.08 | 37.39 ± 12.47  | 0.16 | 93.41 ± 8.94   |

## Appendix B

# Chemical Potential for 4 Components System

Let us consider the generic volume fraction rewritten as  $\phi_i = \frac{n_i r_i}{n_1 + n_2 r_2 + n_3 r_3 + n_4 r_4}$  with  $n_4 \ll n_1, n_2, n_3$ . Then at 1<sup>st</sup> order in  $n_4$  the terms of the free energy can be rewritten as

$$\begin{aligned} n_i \log \phi_i &= n_i \log \phi_i \Big|_{n_4=0} + \frac{\partial(n_i \log \phi_i)}{\partial n_4} \Big|_{n_4=0} n_4 + \mathcal{O}(n_4^2) = \\ n_i \log \frac{n_i r_i}{n_1 + n_2 r_2 + n_3 r_3} - n_i \frac{1}{n_i r_i} \frac{n_i r_i r_4}{n_1 + n_2 r_2 + n_3 r_3} n_4 + \mathcal{O}(n_4^2) &\approx \\ &\approx n_i \log \phi_i - \frac{r_4}{r_i} n_4 \phi_i \end{aligned} \quad (\text{B.1})$$

$$\begin{aligned} \phi_i \phi_4 &= \phi_i \phi_4 \Big|_{n_4=0} + \frac{\partial(\phi_i \phi_4)}{\partial n_4} \Big|_{n_4=0} n_4 + \mathcal{O}(n_4^2) = \\ \frac{n_i r_i n_4 r_4}{(n_1 + n_2 r_2 + n_3 r_3)^2} + \frac{n_i r_i r_4 (n_1 + n_2 r_2 + n_3 r_3)^2}{(n_1 + n_2 r_2 + n_3 r_3)^4} n_4 + \mathcal{O}(n_4^2) &\approx \phi_i \phi_4 \end{aligned} \quad (\text{B.2})$$

$$\begin{aligned} \phi_i \phi_j &= \phi_i \phi_j \Big|_{n_4=0} + \frac{\partial(\phi_i \phi_j)}{\partial n_4} \Big|_{n_4=0} n_4 + \mathcal{O}(n_4^2) = \\ \frac{n_i r_i n_j r_j}{(n_1 + n_2 r_2 + n_3 r_3)^2} - \frac{n_i r_i n_j r_j * 2r_4}{(n_1 + n_2 r_2 + n_3 r_3)^3} n_4 + \mathcal{O}(n_4^2) &\approx \phi_i \phi_j - 2\phi_i \phi_j \phi_4 \end{aligned} \quad (\text{B.3})$$

Neglecting third order terms the chemical potential is thus given by

$$\frac{(\mu_4 - \mu_4^0)}{N_A k_b T} = \frac{\partial G_{mix}}{\partial n_4} =$$

$$\begin{aligned}
 & \frac{\partial}{\partial n_4} \left[ \sum_{i=1,2,3,4} n_i \log \phi_i - \sum_{i=1,2,3} \frac{r_4}{r_i} \phi_i n_4 + \right. \\
 & \left. + N (\chi_S (\phi_1 \phi_2 + \phi_1 \phi_3 + \phi_1 \phi_4) + \chi_{23} \phi_2 \phi_3 + \chi_{24} \phi_2 \phi_4 + \chi_{34} \phi_3 \phi_4) \right] = \\
 & 1 + \log \phi_4 + r_4 \left[ \phi_1 (\chi_S - 1) + \phi_2 \left( \chi_{24} - \frac{1}{r_2} \right) + \phi_3 \left( \chi_{34} - \frac{1}{r_3} \right) \right] \quad (\text{B.4})
 \end{aligned}$$

# Bibliography

- [1] Vladimir N. Uversky. «Protein intrinsic disorder-based liquid–liquid phase transitions in biological systems: Complex coacervates and membrane-less organelles». In: *Advances in Colloid and Interface Science* 239 (2017), pp. 97–114 (cit. on p. 1).
- [2] Yewdall N.A., Alain A.A.M., Tiemei L., and Spruijt E. «Coacervates as models of membraneless organelles». In: *Current Opinion in Colloid & Interface Science* 52 (2021) (cit. on p. 1).
- [3] M. G. F. Last et al. «pH-Controlled Coacervate–Membrane Interactions within Liposomes». In: *ACS Nano* 14.4 (2020), pp. 4487–4498 (cit. on p. 1).
- [4] Wang B., Zhang L., Dai T., et al. «Liquid–liquid phase separation in human health and diseases». In: *Signal Transduction and Targeted Therapy* 6.290 (2021) (cit. on p. 1).
- [5] Zwicker D., Seyboldt R., Weber C., et al. «Growth and division of active droplets provides a model for protocells». In: *Nature Phys* 13 (2017), pp. 408–413 (cit. on p. 1).
- [6] I.B.A. Smokers, M.H.I van Haren, L. Tiemei, and E. Spruijt. «Complex Coacervation and Compartmentalized Conversion of Prebiotically Relevant Metabolites». In: *Chem. Systems Chem.* 4 (2022) (cit. on p. 1).
- [7] Stephen Mann. «The Origins of Life: Old Problems, New Chemistries». In: *Angew. Chem. Int. Ed.* 52(1) (2012), pp. 155–162 (cit. on p. 1).
- [8] Slootbeek A.D., van Haren M.H.I, Smokers I.B.A., and Spruijt E. «Growth, replication and division enable evolution of coacervate protocells». In: *Chem. Commun.* 58 (2022), pp. 11183–11200 (cit. on p. 2).
- [9] Merlijn H.I. van Haren, Karina K. Nakashima, and Evan Spruijt. «Coacervate-Based Protocells: Integration of Life-Like Properties in a Droplet». In: *Journal of Systems Chemistry* 8 (2020), pp. 107–120 (cit. on p. 2).
- [10] Hua Wu and Yan Qiao. «Engineering coacervate droplets towards the building of multiplex biomimetic protocells». In: *Supramolecular Materials* 1 (2022), pp. 2667–2405 (cit. on p. 2).

- [11] Abbas M., Lipiński W. P., Wang J., and Spruijt E. «Peptide-based coacervates as biomimetic protocells». In: *Chem. Soc. Rev.* 50 (2021), pp. 3690–3705 (cit. on pp. 3, 21).
- [12] Dennis A. Dougherty. «The Cation- $\pi$  Interaction». In: *Acc. Chem. Res.* 46 (4) (2013), pp. 885–893 (cit. on p. 4).
- [13] Brian Jacobus Jozefus Timmer and Tiddo Jonathan Mooibroek. «Intermolecular  $\pi$ - $\pi$  Stacking Interactions Made Visible». In: *J. Chem. Edu.* 98 (2) (2021), pp. 540–545 (cit. on p. 4).
- [14] Joel Berry et al. «Physical principles of intracellular organization via active and passive phase transitions». In: *Rep. Prog. Phys.* 82 (2018) (cit. on pp. 5, 26, 27).
- [15] Sangsik Kim, Hee Young Yoo, Jun Huang, et al. «Salt Triggers the Simple Coacervation of an Underwater Adhesive When Cations Meet Aromatic  $\pi$  Electrons in Seawater». In: *ACS Nano* 11.77 (2017), pp. 6764–6772 (cit. on p. 5).
- [16] Karin Bergfeldt, Lennart Piculell, and Per Linse. «Segregation and Association in Mixed Polymer Solutions from Flory-Huggins Model Calculations». In: *J. Phys. Chem.* 100 (1996), pp. 3680–3687 (cit. on p. 6).
- [17] J. T. G. Overbeek and M. J. Voorn. «Phase separation in polyelectrolyte solutions. Theory of complex coacervation». In: *J. Cell. Comp. Biol.* 49 (1957), pp. 7–26 (cit. on pp. 6, 7, 27, 28).
- [18] Charles E. Sing. «Development of the modern theory of polymeric complex coacervation». In: *Adv. Coll. Int. Sci.* 239 (2017), pp. 2–16 (cit. on p. 6).
- [19] Smith A. M., Lee A. A., and Perkin S. «The Electrostatic Screening Length in Concentrated Electrolytes Increases with Concentration». In: *The Journal of Physical Chemistry Letters* 7.12 (2016), pp. 2157–2163 (cit. on p. 7).
- [20] Evan Spruijt et al. «Binodal Compositions of Polyelectrolyte Complexes». In: *Macromolecules* 43(15) (2010), pp. 6476–6484 (cit. on p. 7).
- [21] Donald E. Brooks et al. «Partitioning in Aqueous Two-Phase Systems: Theory, Methods, Uses and Applications in Biotechnology». In: Orlando, Florida 32887: Academic. Press., Inc., 2012. Chap. 2 (cit. on p. 8).
- [22] Arzt M. Deschamps et al. «LABKIT: Labeling and Segmentation Toolkit for Big Image Data». In: *Front. In Comp. Sci.* 4 (2022) (cit. on p. 13).
- [23] H. Bekker, H.J.C. Berendsen, E.J. Dijkstra, et al. «Gromacs: A parallel computer for molecular dynamics simulations». In: *Phys. Comp.* 92 (1993), pp. 252–256 (cit. on p. 15).

- [24] H.J.C. Berendsen, D. van der Spoel, and R. van Drunen. «GROMACS: A message-passing parallel molecular dynamics implementation». In: *Comp. Phys. Comm.* 91 (1995), pp. 43–56 (cit. on p. 15).
- [25] D. van der Spoel et al. «GROMACS: Fast, Flexible and Free». In: *J. Comp. Chem.* 26 (2005), pp. 1701–1718 (cit. on p. 15).
- [26] Ali Ghavami, Erik van der Giessen, and Patrick R. Onck. «Coarse-Grained Potentials for Local Interactions in Unfolded Proteins». In: *J. Chem. Theory Comput.* 9 (2012), pp. 432–440 (cit. on p. 15).
- [27] L. Kind, M. Driver, A. Raasakka, P. R. Onck, P. R. Njølstad, T. Arnesen, and P. Kursula. «Structural properties of the HNF-1A transactivation domain». In: *bioRxiv* (2023) (cit. on p. 16).
- [28] Hamidreza Jafarinaia, Erik Van der Giessen, and Patrick R. Onck. «Molecular basis of C9orf72 poly-PR interference with the  $\beta$ -karyopherin family of nuclear transport receptors». In: *bioRxiv* (2022) (cit. on p. 16).
- [29] Ghavami A., Veenhoff L.M., van der Giessen E., and Onck P.R. «Probing the Disordered Domain of the Nuclear Pore Complex through Coarse- Grained Molecular Dynamics Simulations». In: *Nat Commun* 107.6 (2014), pp. 1393–1402 (cit. on p. 16).
- [30] Michael Levitt. «A simplified representation of protein conformations for rapid simulation of protein folding». In: *J. Mol. Biol.* 104 (1976), pp. 59–107 (cit. on p. 16).
- [31] E.D. Holmstrom, Z. Liu, Nettels D., et al. «Disordered RNA chaperones can enhance nucleic acid folding via local charge screening». In: *Nat Commun* 10.2453 (2019) (cit. on p. 16).
- [32] Lauren H. Kapcha and Peter J. Rossky. «A Simple Atomic-Level Hydrophobicity Scale Reveals Protein Interfacial Structure». In: *J. Mol. Biol.* 426 (2014), pp. 484–498 (cit. on p. 16).
- [33] Zeke A. et al. «Deep structural insights into rna-binding disordered protein regions. WIREs RNA». In: *WIREs RNA* 13(5) (2022) (cit. on p. 17).
- [34] Baruch Leshem A., Sloan-Dennison S., Massarano T., et al. «Biomolecular condensates formed by designer minimalistic peptides». In: *Nat. Commun.* 14.421 (2023) (cit. on pp. 21, 35).
- [35] Onuchic PL, Milin AN, Alshareedah I, Deniz AA, and Banerjee PR. «Divalent cations can control a switch-like behavior in heterotypic and homotypic RNA coacervates». In: *Sci. Rep.* 9.12161 (2019) (cit. on pp. 22, 30, 41).
- [36] Allen P. Minton. «Simple Calculation of Phase Diagrams for Liquid-Liquid Phase Separation in Solutions of Two Macromolecular Solute Species». In: *J. Phys. Chem. B* 124 (2020), pp. 2363–2370 (cit. on p. 23).

- [37] P. Zhang et al. «Salting-Out and Salting-In of Polyelectrolyte Solutions: A Liquid- State Theory Study». In: *Macromolecules* 49(24) (2016), pp. 9720–9730 (cit. on p. 28).
- [38] Marat Andreev, Vivek M. Prabhu, Jack F. Douglas, Matthew Tirrell, and Juan J. de Pablo. «Complex Coacervation in Polyelectrolytes from a Coarse-Grained Model». In: *Macromolecules* 51.17 (2018), pp. 6717–6723 (cit. on p. 42).