Chemical control of liquid phase separation in the cell

Dissertation

zur Erlangung des akademischen Grades

Doctor rerum naturalium (Dr. rer. nat.)

vorgelegt

der Fakultät Physik

Bereich Mathemathik und Naturwissenschaften

der Technischen Universität Dresden

von

Omar Adame Arana

geborem am 11.11.1987 in Ensenada, Mexiko

Max Planck Institut für Physik komplexer Systeme

Dresden, 2019

Wissenschaftlicher Betreuer:

Prof. Dr. Frank Jülicher

Eingereicht am 14.11.2019

Gutachter:
1. Prof. Dr. Frank Jülicher
2. Prof. Dr. Stephan Wolfgang Grill
3. Prof. Dr. Hossein Fazli

Verteidigt am 29.01.2020

A mis padres, por todo lo que me han enseñado en la vida.

Acknowledgments

First of all, I would like to thank Frank Jülicher for giving me the chance of working at the MPI-PKS and for supervising my work. I am grateful for his patience, his clear and concise feedback and for teaching me so much on how to do research. Second, I would like to thank Christoph A. Weber, who was a great supervisor as well and with whom I had plenty of enjoyable scientific discussions. Christoph acted like a supervisor but importantly also as a friend, I am indebted with him for all of his support during these past years.

I also would like to thank Tim Minzoni, whose lectures were always impressive and whom encouraged me to study abroad, he was my teacher, my mentor and a role model that sadly is no longer with us. I thank Víctor Romero Rochín, who is the best at teaching Thermodynamics as well as a great person and scientist, incredibly committed to preparing the future generation of scientists in Mexico and a really fun person to talk to. They were both crucial in the development of my own scientific taste.

During my time here, I had the chance to discuss and collaborate with great scientists, one of them was Jacques Prost who was always extremely nice to me and ready to talk about science, therefore, I would like to express my gratitude to him for his contribution to my scientific development. I also thank Vasily Zaburdaev, always willing to discuss with me about our project and with his good sense of humor provided a nice atmosphere to have a chat. I would also like to thank A.A. Hyman and his group for sharing part of their knowledge with me. Other collaborators whom I would particularly like to acknowledge are: Shamba, Chrissy, Andrés and Anatol. They were always very kind and joyful and it has been a pleasure to learn from them. After our discussions, I always felt motivated to learn more and more about science. I also want to thank all the PhD students that have been part of the Biological Physics group during my stay here, the environment in the weekly seminars was always friendly and motivating, I learned a great deal by listening to others but also by trying to explain physics in a precise language. There are also the members of the Droplet meeting, I thank them for providing me with a great space to discuss results in an informal way. Regarding this Thesis, I must thank Bahareh, Giacomo, Johanna, Lars, Sherry, Stefano, Sudarshana and Talía for reviewing parts of it, I've got plenty of great suggestions from them. Additionally, I am grateful to Johanna and Jonathan who were so kind to translate my abstract in such a short notice. I thank the IT department and the staff of the MPI-PKS, most of them were very friendly and helpful during my stay. In particular I thank Ulrike, for doing her job in such a wonderful way.

Now comes the hard part, where I hope that my friends and some family members can forgive me for not giving a detailed expression of gratitude to each one of them, but if they really know me, I think they will understand why I decided not to do so. It feels utterly unfair to me, to only thank those who are present in my mind at this moment, so what I

will do, is to try to remember activities and events where in some way or another, many of my friends were present. For example, I have to thank all of those who traveled with me around Europe in the past years, I will never forget the nice moments that I experienced in all these trips. There are also the friends with whom I spent Christmas during my stay in Dresden, the food we cooked, the dancing and the drinking were all part of great fun days, and I am extremely grateful to them for that. I thank the friends that I met when I was a bachelor student, my former flatmates and my former teammates, they have given me so many joyful moments, which I treasure with me wherever I go. Some of these friends have come to visit me and filled me in with plenty more good memories for the future, I cannot but be grateful to them for being such wonderful friends. The friends who studied German with me know that we had a lot of fun there, maybe sometimes it was too much fun and that might be the reason why my German remains so poor to this moment, I thank them for the nice time that we spent together. This brings me to my German teacher, who became a dear friend and was the one that introduced me to one activity that became a definite part of my week, which is to play volleyball every Thursday with a very nice group of people, some of which, became a very important part of my life too, so much that I could even say that they, with their warm hearts, finally made me feel at home in Dresden. Then come the friends and colleagues with whom I discussed politics and the current affairs of the world where, in order to make a point, I many times had to invoke the words and knowledge of "him". I thank them for pushing me to read more and for teaching me that in order to get somewhere in a discussion, one must first find some common ground with the others. I thank my family as well, although I have not seen many of them lately, they have always supported me and they mean a lot to me.

In these last words, I first want to thank my brother Damian, for looking after me when we were young, also for teaching me about life in very unexpected ways, and lately, for providing me with names of authors that have worked on topics that suddenly became essential to me, I hope we can once again work for a good cause together as we were doing some time ago. Now, I need to switch the language since this part is intended for my parents: A mis papás les agradezco su amor, reflejado en sus cuidados, en sus atenciones y en sus esfuerzos por darme todas las herramientas necesarias para que yo sea el protagonista de mi propia vida. Además de agradecer todo el amor que me han dado, les agradezco porque mediante su ejemplo, me enseñaron a ser justo y respetuoso con los otros, como también me mostraron a apreciar las cosas simples de la vida y que los lujos no significan nada, puedo decir que gracias a mis padres disfruto casi todos los momentos que paso en este mundo. Por todo esto y más les doy las gracias. Finally, my deepest gratitude goes to Talía, who has the strange ability of getting inside my brain to know exactly what my interests are. I thank her for her love and her company, sharing my life with her has been the most enjoyable ride I've had in my life. She always provides me with the courage to become who I want to be. Thanks to her, life just seems to get better everyday.

Abstract

One of the main features of cells is their incredible ability to control biochemical processes in space and time. They do so by organizing their interior in sub-compartments called organelles, each of them with a different biochemical environment that allows them to perform specific tasks in the cell. It is sometimes believed that these compartments need a membrane in order to have a stable biochemical environment and regulate their compositions. However, there are some organelles which lack a membrane and seem to form and organize via liquid-liquid phase separation. Some of the components that form these membraneless organelles have the ability to bind to RNA and form complexes, while some others react to changes in the intracellular environment such as pH variations, which in turn affects their protonation state. In order to study these processes from a theoretical perspective, we develop a generic thermodynamic framework to study systems exhibiting liquid-liquid phase separation at chemical equilibrium. This framework, based on the use of conservation laws in chemical reactions, allow us to identify thermodynamic conjugate variables at chemical equilibrium, which are given by a set of conserved quantities and the corresponding conjugate chemical potentials.

Within the aforementioned framework, we introduce a minimal model to study the effect of pH on liquid-liquid phase separation. Our model explains macromolecular phase separation controlled by protonation and deprotonation reactions, which are tuned by the pH of the system. We study the phase behavior of the system as a function of pH. Our main findings are: Firstly, the broadest region of phase separation is typically found at the isoelectric point. Secondly, the system exhibits reentrant behavior. Thirdly, that the dominating interaction in the system determines the topology of the phase diagrams. Our model is in agreement with experimental observations of *in vitro* protein phase separation of pH-responsive intrinsically disordered proteins, as well as with observations of protein phase separation exhibited by many cytosolic proteins when the intracellular pH in yeast cells is brought close to the isoelectric point of such proteins.

Moreover, this work analyses the physical mechanism behind the positioning of liquid-like organelles in the *Caenorhabditis elegans* organism known as P granules. In order to study this phenomenon, we first present firm experimental evidence showing that PGL-3 protein, a key component of P granules, forms liquid-like drops whose assembly can be

II

modulated by RNA. We then present data showing that the RNA-binding affinity differs significantly between proteins relevant for the positioning of P granules, such as MEX-5 and the proteins forming the P granules, like the aforementioned PGL-3. This points to a possible mechanism of RNA-binding competition between P granule constituents and MEX-5 in order to spatially control the condensation and dissolution of P granules. Based on the experimental evidence, we propose a minimal model in which we couple phase separation of PGL-3 to a set of binding reactions involving the MEX-5 protein and RNA. We find that in order to explain the experimental data, the tendency for phase separation of the PGL-3 protein increases with the formation of complexes of PGL-3 bound to RNA. This therefore supports the idea that MEX-5 inhibits this protein phase separation by depleting the RNA available for PGL-3 to form such complexes. This simple mechanism is at the core of how P granules localize to the posterior side of the *Caenorhabditis elegans* embryo.

Zusammenfassung

Zellen sind in der Lage, gleichzeitig ganz unterschiedliche biochemische Prozesse zu bewältigen. Dies gelingt ihnen durch eine Einteilung ihres Inneren in Kompartiemente, sogennante Organellen, die die jeweils geeignete biochemische Umgebung für die unterschiedlichen Aufgaben schaffen. Bei membranumschlossenen Kompartimenten ist leicht vorstellbar, dass sie eine andere biochemische Zusammensetzung als ihre Umgebung haben können. Jedoch existieren auch Organelle ohne Membran die durch eine flüssig-flüssig Phasenseparation entstehen. Manche dieser Kompartiemente haben die Fähigkeit, RNA zu binden und Proteinkomplexe auszubilden, während andere auf die Veränderungen innerhalb der Zelle, wie z.B. die Veränderung des pH-Werts und der damit Verbunden Änderung ihres Protonierungszustands, reagieren können. Um diese Prozesse theoretisch analysieren zu können, entwickeln wir zunächst ein allgemeingültiges, thermodynamisches Gerüst, um Systeme zu untersuchen, die im chemischen Gleichgewicht flüssig-flüssig phasensepariert vorliegen können. Dies erlaubt, basierend auf den Erhaltungsgrößen, im chemischen Gleichgewicht thermodynamisch konjungierten Variablen zu identifizieren, welche aus den erhaltenen Komponenten und den zugehörigen chemischen Potentialen bestehen.

Mithilfe des obig erwähnten Gerüsts können wir den Einfluss des pH-Wertes auf die flüssig-flüssig Phasenseparation in einem minimalen Modell untersuchen. Dies beschreibt die makromolekulare Phasenseparation, kontrolliert durch Protonierungs- und Deprotonierungreaktionen, welche wiederum vom pH-Wert abhängig sind. Unsere Untersuchung der pH-Abhängigkeit der Phasenseparation kommt zu folgenden Ergebnissen: Erstens liegt die größte Region von Phasenseparation im Phasendiagramm typischerweise im Bereich des isoelektrischen Punkts. Zweitens zeigt das Modell eine Fähigkeit der erneuten Mischung auf. Drittens ist die Topologie des Phasendiagrams von der dominantesten Interaktion bestimmt. Unser Modell stimmt mit experimentellen Beobachtungen zur Phasenseparation von intrinsisch ungeordneten, Proteinen, deren Struktur sich pH abhängig verändern, überein. Das Modell ist außerdem konsistent mit Beobachtungen von Phasenseparation von Proteinen im Zytosol von Hefezellen, die entsteht, wenn der intrazellulare pH-Wert in die Nähe des isoelektrischen Punkt dieser Proteine gebracht wird.

Des Weiteren geht diese Arbeit auf den physikalischen Mechanismus ein, mit dem flüssigkeitsähnliche Organellen, sog. P granules, im Organismus Caenorhabditis elegans

positioniert werden. Um dieses Phänomen zu analysieren, stellen wir zunächst experimentelle Beobachtungen vor, die zeigen, dass PGL-3, eine Hauptkomponente der P granules, flüssigkeitsähnliche Tropfen bildet, deren Zusammensetzung von RNA moduliert werden kann. Darüber hinaus zeigen wir Daten, die großen Unterschiede zwischen der RNA-Bindungsaffinität von Proteinen wie Mex-5, die für die Positionierung der P granules relevant sind, und solchen, die P granules bilden, wie PGL-3, zeigen. Dies deutet darauf hin, dass eine Konkurrenz zwischen den Bestandteilen der P Granula und MEX-5 um die zur Bindung zur Verfügung stehende RNA besteht, die die Kondensation und Auflösung von P Granula räumlich kontrollieren könnte. Auf diesen experimentellen Befunden aufbauend führen wir ein minimalles Modell ein, in dem wir die Phasenseparation von PGL-3 an Bindungsreaktionen der MEX-5 Proteine und RNA koppeln. Um die experimentellen Beobachtungen beschreiben zu können, muss die Neigung des PGL-3 Proteins zur Phasenseparation zunehmen, wenn es Komplexe mit RNA bildet. Dies unterstützt die Idee, dass MEX-5 diese Phasenseparation unterdrückt, indem es die Anzahl an möglichen RNA-Bindungspartner für PGL-3 herabsetzt und damit die weitere Entstehung derartiger Protein-RNA-Komplexe erschwert. Dieser einfache Mechanismus scheint die Hauptursache dafür zu sein, dass P granules auf der posterioren Seite des Caenorhabditis elegans Embryos zu finden sind

Contents

Chapter 1

Introduction

The question of how a cell organizes its interior in time and space is a fascinating one. In order to properly function, a cell has to orderly and stringently control a variety of both chemical and physical process in its interior. The cell organizes these processes by sub-compartmentalizing its interior in a number of organelles, some of which are delimited by well-defined membranes that provide them with distinct biochemical environments to perform diverse tasks.. However, there is a growing body of evidence showing that there are liquid-like organelles which lack a membrane. These membraneless organelles can dynamically exchange material and just as in the case of membrane-bound organelles, provide distinct environments where specific biochemical reactions can take place. In order to know more about how liquid-like organelles form and how their compositions are controlled inside a living cell, we develop a generic thermodynamic framework to study how liquid-liquid phase separation can be controlled by different types of reactions. Based on experimental observations we construct minimal models which can account for some of the protein phase separation behavior observed in living cells.

We organize this introduction as follows: In order to motivate our study, we first provide a general picture of the role played by liquid-liquid phase separation in the cell. We continue by showing evidence that phase separation can be regarded as an organizing principle in the cell and discuss relevant experiments for the development of this work. After introducing all the relevant biological phenomena, we present a thorough discussion of the thermodynamics of phase separation in multicomponent mixtures, in which we also include the interplay between chemical reactions and phase separation. We conclude by giving an overview of this thesis.

1.1 Spatial organization of cells

The cell is the building block of all living organisms, it organizes its interior in sub-compartments known as organelles. These organelles are in charge of performing different
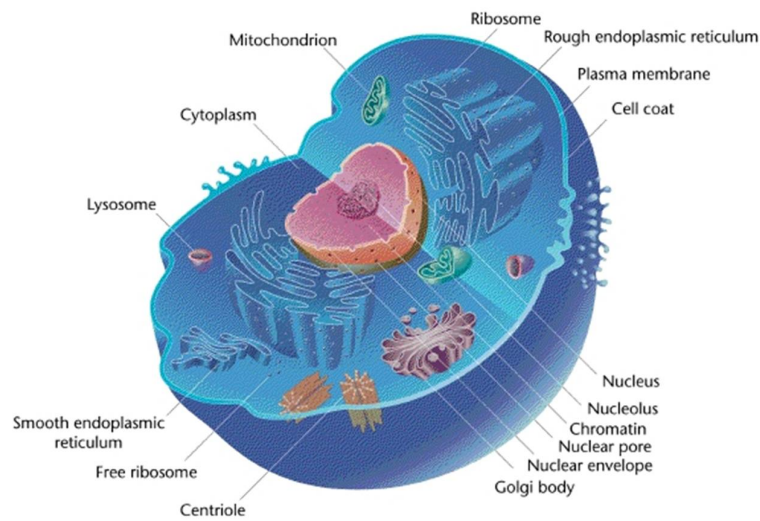
FIGURE 1.1: Organizational structure of an animal cell. The majority of the structures shown in the sketch are membrane-bound organelles. Image from [4]

tasks inside the cell, and in some cases, they are bound by membranes that separate them from the intracellular fluid known as cytosol. The basic structure of a cell, schematically shown in Fig. 1.1, has long been considered to be mostly defined by compartmentalization via membrane-bound organelles. Using their membrane, these organelles are capable of: Selectively choose which proteins can enter or leave via translocation processes [1], control their pH [2], and protect its interior from alterations in the biochemistry of the cytosol. It thus seems reasonable to think that membranes are needed in order to control the internal biochemistry of organelles, however, this idea has been challenged in the past decade. There is growing evidence showing that many organelles without a membrane (hereafter membraneless organelles) exist and that they also provide distinct biochemical environments that are able to perform specific tasks inside the cell [3]. As a result, liquid-liquid phase separation, the process in which a homogeneous mixture splits into different liquid phases, has gained an enormous amount of attention in the context of intracellular organization [3, 5–7]. Although the idea of liquid-liquid phase separation as an organizing principle for the interior of the cell had been hypothesized before [8] and maybe even observed [9–11], it was not until Brangwynne et al. [12] carefully showed that the germ granules of the *Caenorhabditis elegans*[1] (*C. elegans*) behave like liquid-like droplets. Germ granules in general are organelles composed of various RNA and protein molecules that are present in all germ cells [14], and they pass the genetic material to the next generation of the organism. In the particular case of *C. elegans*, the germ granules are known as P granules, which were the focus of the study of Brangwynne et al where they showed that

---

[1] *C. elegans* is a nematode (roundworm) which is an ideal model organism for experimentation due to its very regular process of development [1] and due to its transparency that makes the animal ideal for fluorescent microscopy [13].
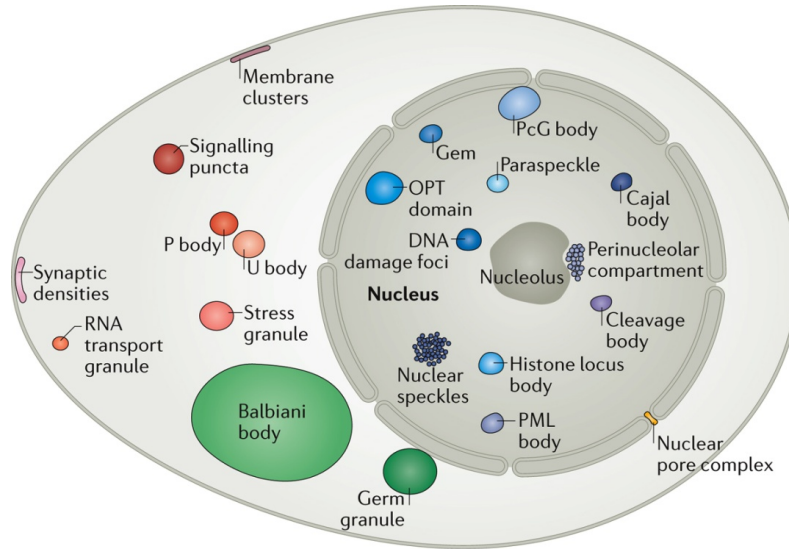
Figure 1.2: Biomolecular condensates in eukaryotic cells (cells with a nucleus). Examples of biomolecular condensates in the cytoplasm, nucleoplasm and in the cell membranes. Image adapted from [3]

P granules fuse, wet, exchange material with their surroundings, and drip the nucleus of the cell just as if they were liquid drops. They further showed that P granules condense and dissolve in a spatially regulated way, leading to the positioning of P granules to one side of the *C. elegans* embryo before the first cell division. This work thus provided strong evidence for liquid-liquid phase separation as an organizing principle in the cell.

After P granules were shown to behave like liquid drops, many new studies have characterized a rich variety of so-called membraneless organelles which are also termed biomolecular condensates [3]. In this work we use the terms indistinguishably. The name biomolecular condensates might be preferred sometimes since it reflects their true nature given that they are distinct biochemical environments that condense together specific proteins and RNA. One possible function of these condensates is that they may serve as distinct biochemical environments inside the cell. Such biomolecular condensates emerge in many contexts, with some of them forming as a response to changes in the intracellular environment such as temperature increase or a drop in the cytosolic pH [15, 16], whereas others exist in normal intracellular conditions [12, 17, 18]. The picture of intracellular organization shown in Fig. 1.1 should then be complemented with one that includes the growing body of membraneless organelles or biomolecular condensates (that are being currently studied). An important remark is that not all of the known biomolecular condensates behave like liquids but still many of them do seem to form via phase separation, and at least on early stages after their formation they behave like liquid drops. A sketch showing some biomolecular condensates in the cell is shown in Fig. 1.2.
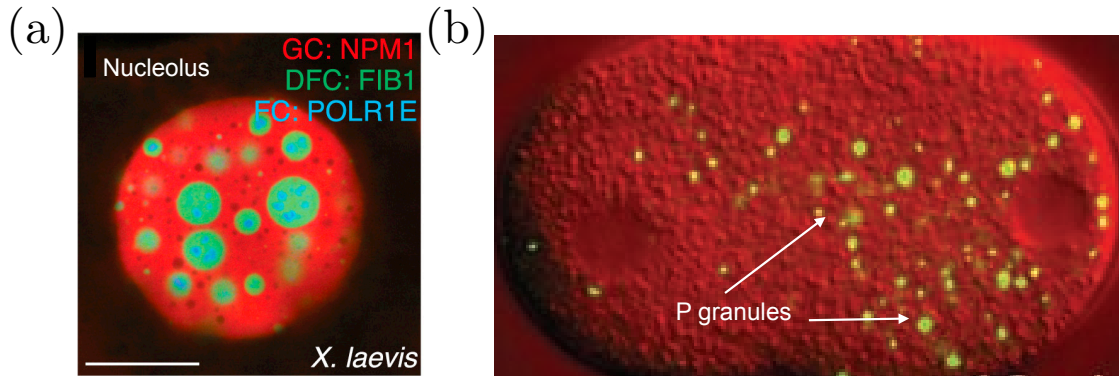
FIGURE 1.3: Liquid-like organelles in the cell. (a) Nucleoluls in *Xenopus laevis* exhibits a multilayer structure, the colors reflect the fluorescence of proteins that localize in different layers of the nucleolus. Scale bar, $20\,\mu$m. Image adapted from [18]. (b) P granules in the *C. elegans* embryo. Image adapted from [12].

1.2 Liquid phase separation as an organizing principle in the cell

Since the ground breaking work of Brangwynne et al. [12], there have been myriads of studies focused on studying the role played by liquid-liquid phase separation in the compartmentalization of the cell [15, 17–31]. Some of these studies showed that previously known membraneless compartments actually have liquid-like properties and that they are likely formed via liquid-liquid phase separation. In addition to P granules (Fig. 1.3(b)), some of these compartments are nucleoli [19] (Fig. 1.3(a)), stress granules [15, 32] and Cajal bodies [17].

Biomolecular condensates that form via liquid phase separation have a natural degree of selectivity without the need of a membrane. This is a natural consequence of phase separation, if the demixing process is driven by attractive inter-molecular interactions, then one expects that proteins which phase separate will attract some of their binding partners. Because biomolecular condensates are enriched in specific proteins, it has been proposed that they may also act as suitable environments for specific biochemical reactions with increased reaction rates as compared to those in the cytosol [3]. Another very interesting aspect of these liquid-like organelles is that they can assemble and disassemble in short time scales, providing a response mechanism to drastic changes in the environment [15, 33]. A prominent example in this regard are stress granules, that are rapidly formed after the cell is exposed to stress conditions such as heat shocks or acidification of the cytosol. Many different proteins are recruited into these granules as a protective measure [30, 34] and the granules persist until the environmental conditions are brought back to normal. Another interesting example of biomolecular condensates is the nucleolus, an organelle located inside the nucleus [18, 19, 22] which is in charge of producing the necessary elements for ribosome

assembly [35]. It has been shown [18], that the nucleolus is a liquid phase that has an inner multilayered structure. This structure is reminiscent of liquid multi-phase behavior in which two of the liquid phases have different surface tensions relative to the solvent, leading to one of them being engulfed by the other in order to reduce the energetic cost of forming an interface with the solvent. A picture of a nucleolus taken with fluorescent microscopy is shown in Fig. 1.3(a) [18], where one can appreciate different layers, that were shown to behave as distinct liquid phases. The colors in the image show fluorescently labeled proteins that are known to localize in different layers inside the nucleolus. The multilayered structure is thought to facilitate the processing of rRNA (ribosomal RNA) by different enzymes at different layers before the rRNA ultimately exits the nucleolus and is transported to the cytoplasm to assemble ribosomes [18]. Interestingly, one can see that a complex but flexible degree of organization can be achieved by liquid phase separation without the need of a membrane. It is worth saying that RNA plays a prominent role in the assembly of biomolecular condensates [18, 22, 23, 25, 27–29, 36–38], this is related to the fact that many constituent proteins of biomolecular condensates are RNA-binding proteins. We address the RNA effect on phase separation in one chapter of this thesis where we study an RNA-binding competition mechanism underlying P granule segregation.

In this section we did not intend to provide an exhaustive list of biomolecular condensates and their functions. Instead, by discussing some of their features and their ubiquity in the cell, we wanted to motivate a reason for studying the process of liquid-liquid phase separation in the context of a cell. In the following, we discuss two examples of biological phenomena that motivated the physical models presented in this thesis. The first one is the P granule segregation to the posterior side of the *C. elegans* embryo prior to the first cell division [12, 28, 39] and the second one is the widespread macromolecular assembly in yeast cells occurring as a response to stress conditions such as starvation [16].

1.3   P granule segregation in the *C. elegans* embryo

P granules were the first organelles which were thoroughly characterized as liquids in the seminal work of Brangwynne et al. [12]. Where they used a fluorescently labeled protein known to localize in P granules, namely PGL-1, to observe how P granules localize on one side of the worm (*C. elegans*). The process of P granule segregation to the posterior happens as follows (Fig. 1.4): First, P granules are homogeneously distributed across the embryo. Then, there is a symmetry breaking in the embryo and an anterior-posterior axis is established. After the symmetry breaking in the embryo, a MEX-5 gradient is established along the anterior-posterior axis (not shown in the figure); MEX-5 is an RNA-binding protein implicated in the degradation of P granule components [12]. The gradient is established as a consequence of a diffusivity gradient related to different phosphorylation rates of MEX-5 along the anterior-posterior axis [40], with the establishment of the gradient
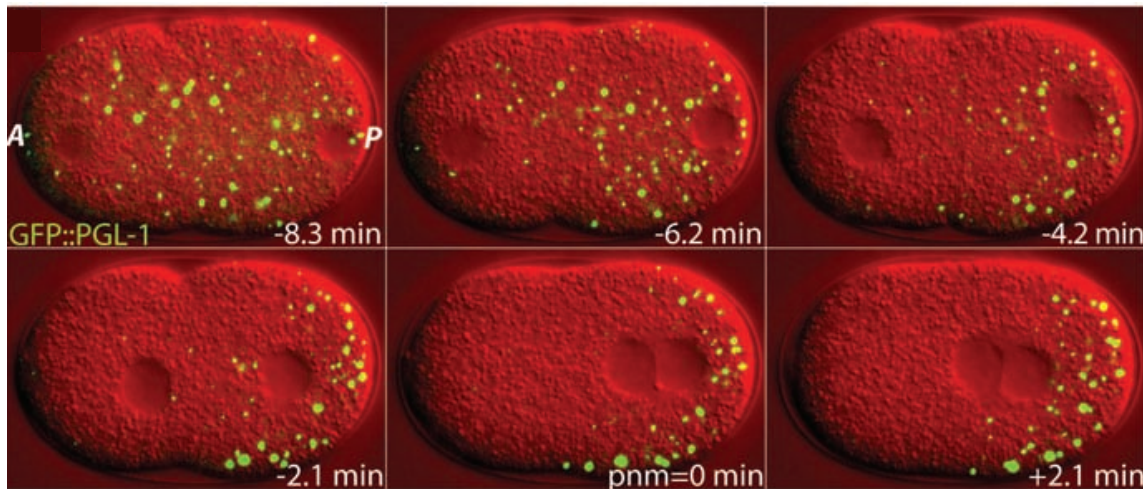
FIGURE 1.4: P granule segregation in the *C. elegans* embryo. Time lapse showing the localization of P granules from an originally homogeneous distribution across the embryo to the posterior side. The protein PGL-1, which localizes in P granules, is tagged with a green fluorescent protein. In the upper left panel, A represents the anterior side of the embryo and P the posterior. The time is measured relative to the time when the two pronuclei (hollow circles) meet, panel with legend pnm $= 0$ min. Panels are adapted from [12]

there is a concomitant positioning of P granules to the posterior side of the embryo. Finally, at the stage prior to cell division, P granules are fully localized on the posterior side. The cell which contains the P granules after the first cell division gives rise to the sexual progenitors of the organism. A hint at the underlying mechanism of P granule positioning is given by the fact that the MEX-5 gradient anti-correlates with the positioning of the P granules.

Let us briefly review the experimental observations that led to conclude that P granules are liquid-like drops whose dissolution and condensation is spatially controlled. We first discuss the evidence showing that P granules behave like liquids. P granules were observed to fuse with one another, the fluorescence of tagged proteins was shown to recover after photobleaching (disabling the fluorescent label) by exchanging material with their surroundings. They further seemed spherical just as one expects from a liquid due to surface tension, and finally, they were shown to flow when they were under pressure. All these facts together led the authors to recognize P granules as liquid-like drops. Regarding the spatial control of dissolution and condensation, it was reported that P granules do not segregate to the posterior side by a net current flow, but instead they dissolve at the anterior side where MEX-5 has a higher concentration and condense at the posterior side where the concentration of MEX-5 is lower. This was corroborated by studying a genetically modified roundworm (*C. elegans*) which did not express MEX-5, and as a result there was no concentration gradient established. This led formation of to worms with a homogeneous distribution of P granules that did not dissolve with time. Brangwynne et

al. concluded the work by hypothesizing that the localization of P granules is due to a demixing transition where the saturation concentration for phase separation depends on the properties set by the concentration gradient along the anterior-posterior axis.

The aforementioned work set the basis for understanding liquid phase separation as an organizational mechanism in the cell by showing how the spatially controlled dissolution and condensation of P granules lead to their localization. Although the mechanism of a controlled saturation concentration was proposed, the underlying physics of this control remained elusive. The question on how the condensation and dissolution of P granules are set by the gradient of MEX-5 was addressed in a collaboration [28]. In this work we found that an mRNA-binding competition between P granule components and MEX-5 might be the mechanism underlying this segregation. We present these results in Chapter 5.

1.4 Macromolecular assembly as a response to stress

Cells under stress conditions often transition from a normal growth state to a stalled condition where growth is inhibited [33, 41]. One of their responses to a wide range of stresses such as starvation and heat shocks, is to form so-called stress granules, which are assemblies of many different cytosolic proteins and RNA [34, 42]. This assemblage seems to be a protective measure. For instance, when stress granules form as a response to heat shocks, they sequester the mRNA[2] encoding for most of the proteins of the cell except for those who encode for heat shock proteins which are needed to diminish the damage caused by the heat stress [43].

Here, we particularly discuss the budding yeast (hereafter yeast)[3] response to changes in its intracellular acidity, quantified by a drop of the cytosolic pH after the cells are exposed to periods of starvation. This response was studied in the work of Munder et al. [16], in which they showed that a drop in the cytosolic pH leads to a liquid-solid transition of the cytosol. They assessed this behavior by tracking the mobility of tracer particles via fluorescent labeling. Munder et al. showed that the dormant stage attained when the cells are exposed to starvation periods, is associated to a decrease in the mobility of the particle tracers. They then quantified the cytosolic pH in starvation conditions. At normal growth conditions for yeast cells, the cytosolic pH is close to neutrality pH $\simeq$ 7.3. In contrast, when the cells experienced starvation, their cytosolic pH drops to pH $\simeq$ 5.8. Interestingly, since it is well known that pH has a strong effect on the solubility of different proteins [45–47], they analyzed the isoelectric point[4] distribution of the yeast proteome[5] and found

---

[2]mRNA is the sequence of base pairs translated by ribosomes in order to produce specific proteins encoded in the mRNA sequence.

[3]Budding yeast (*Saccharomyces Cerevisiae*) has been used as an experimental organism for more than 50 years, it is easy to treat and plenty of its genes are found in many other eukaryotic cells [44].

[4]The isoelectric point of a macromolecule is defined as the pH value for which the macromolecule does not carry a net charge on average.

[5]The proteome is the entire set of all different proteins of an organism
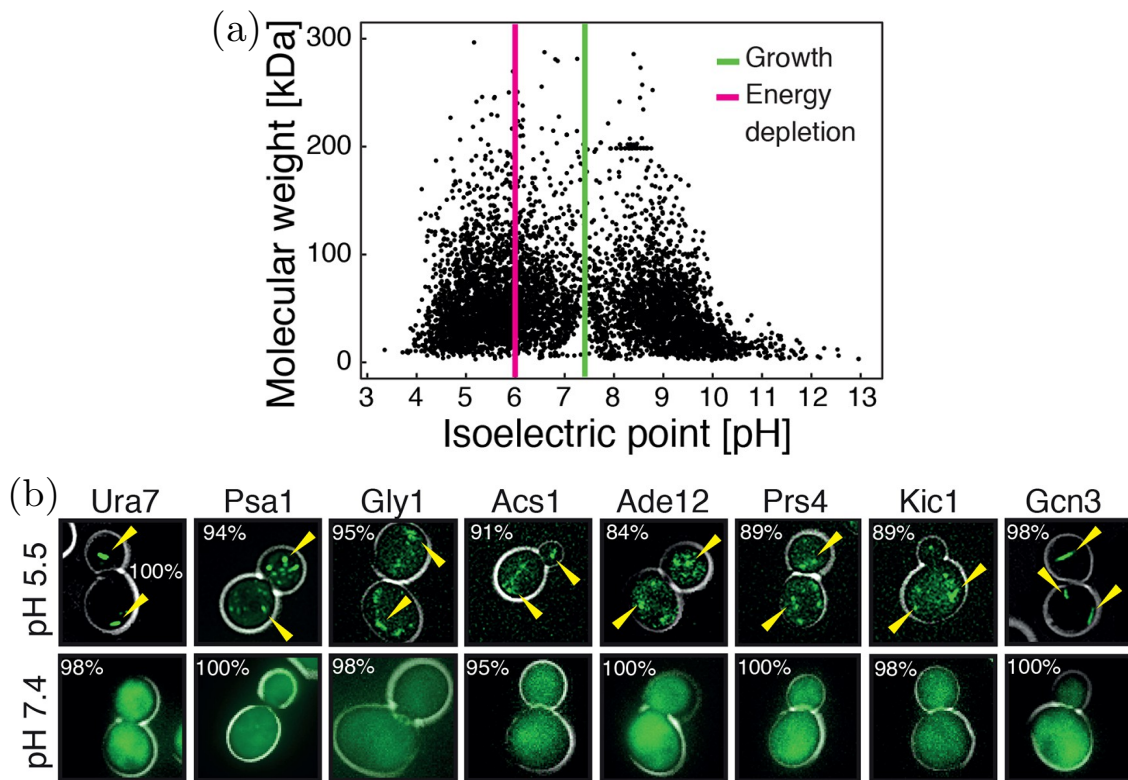
FIGURE 1.5: Macromolecular assembly dependence on pH. (a) Isoelectric point distribution for the yeast proteome shown as a function of the protein molecular weight. The green line represents the pH at optimal growth for yeast cells and the pink line shows the cytosolic pH upon energy depletion. (b) Macromolecular assembly upon acidification of the cytosol. Top row: fluorescently labeled proteins which form assemblies at pH = 5.5 . Bottom row: Proteins are well-mixed for normal conditions of pH. for different. Both panels are adapted from [16]

that the distribution has two peaks, one that is close to values of pH $\simeq 5.5$ and other close to pH $\simeq 9$, which is in agreement with previous work [48]. The bimodal distribution is shown in Fig. 1.5(a). Remarkably, what they found was that the pH of the cytosol under starvation conditions (pink line in Fig. 1.5(a)), was relatively close to the isoelectric point of many cytosolic proteins. With this in mind, using fluorescent microscopy they probed a large set of proteins and observed what happened if the cytosol was brought to a pH value of 5.5 by equilibration with a buffer. Results for some proteins tagged with a green fluorophore are shown in Fig. 1.5(b), where it can clearly be seen that at pH = 7.4 the proteins are well-mixed in the cytosol. In stark contrast, they found that these proteins formed different types of assemblies for acidic pH. Then, they probed the relevance of these pH-dependent protein assemblies in the cell under starvation conditions but keeping the cytosolic pH close to pH $\simeq 7$ to prevent macromolecular assembly. What was found is that the macromolecular assemblies were needed for survival of the cells and thus concluded that these assemblies play a protective role under stress conditions [16].

    Although the work that we have just discussed was concerned with the transition from

fluid to solid of the cytosol, it emphasized the importance of pH control in the cell, and not only that, it clearly showed a close connection between protein assembly in the cell and the isoelectric point distribution of the cell proteome. An interesting question that arises is the following: Can we explain the process of phase separation as a function of pH? We address this question in Chapter 4, where we propose a minimal model to study pH-dependent liquid-liquid phase separation.

In this section we have only discussed the pH-dependence of macromolecular assembly in living organisms, that is, we discussed what is known as *in vivo* experiments. Nevertheless, important insights into the mechanisms of protein liquid-liquid phase separation (hereafter protein phase separation) can also be obtained via experiments in a test tube (*in vitro*). For example, there is *in vitro* evidence showing that a constituent protein of stress granules exhibits reversible pH-dependent phase separation [30]. Given the usefulness of *in vitro* experiments to unravel the driving mechanisms of protein phase separation, in the next section, we briefly discuss experiments where it was found that many constituent proteins of different biomolecular condensates phase separate *in vitro*.

1.5   Protein phase separation *in vitro*

Interesting insights of the mechanisms driving protein phase separation can be drawn from experiments in test tubes. The great advantage of performing *in vitro* experiments is the possibility of having a high level of control as compared to experiments performed on living organisms which undergo all kind of processes that cannot be controlled by the experimenter.

Since the work by Brangwynne et al., experimental studies on protein phase separation *in vitro* have sharply increased [18, 22, 23, 26–28, 30, 31, 34, 38, 39, 49, 50]. One important observation that has been repeatedly made is that many constituent proteins of biomolecular condensates undergo liquid-liquid phase separation in buffer solutions[6] at remarkably low concentrations of the order of μM [23, 28, 30, 38, 39, 51]. It has also been shown that purified proteins which are found in the nucleolus show a similar multilayered structure *in vitro* as the one observed *in vivo* [18]. Suggesting that the properties of biomolecular condensates might be defined by only a few key components. Another interesting aspect of studying protein phase separation *in vitro*, is the possibility of assessing the interactions among components of the same biomolecular condensate. Given that RNA is found in many if not all of the biomolecular condensates, it is of particular interest to study the effect of RNA in protein solutions. This has been addressed experimentally in many different studies [3, 18, 22, 23, 25, 27–29, 37, 38, 52], where the studies were done with biomolecular condensates constituents that can bind to RNA. The results of the influence of RNA on

---

[6]buffers are mixtures of ions and water which mimic the physiological conditions of salt and pH, among other properties)

protein phase separation are various. It has been reported that small amounts can promote protein phase separation [28] but it was also shown that it may act as an inhibitor of the demixing of other proteins [34][7]. One study [38] showed that low ratios of RNA to protein promote phase separation, whereas high ratios may be used to inhibit it. The authors of this study suggested that this may explain why some proteins are soluble in the nucleus (which has a high content of RNA) in comparison to their irreversible association in the cytoplasm that might start by protein phase separation. Thus protein-RNA associations seem to play an important role in the regulation of biomolecular condensates assembly. Lastly, the fact that some proteins undergo pH-dependent phase separation indicates that protonation and deprotonation reactions might also play an important role in order to determine the phase behavior of some proteins. We devote Chapter 4 to theoretically study this scenario.

Thus far we have reviewed the ubiquity of protein condensates in living organisms, some prominent examples of liquid phase separation in the cell, and experimental evidence of protein phase separation *in vitro*. In addition, we discussed that biomolecular condensates are enriched in proteins and RNA, however, we have not yet described the nature of the proteins that take part in these processes. These proteins have the special attribute of belonging to a class of proteins which are called intrinsically disordered. In the next section we discuss what an intrinsically disordered protein is and introduce their relevant properties.

1.6   Intrinsically disordered proteins

Proteins are complex macromolecules defined by a unique sequence of amino acids. For a long time it was believed that each sequence of amino acids corresponded to a protein with a well defined three dimensional structure that was directly linked with their biological function [53]. However, in the past two decades there has been growing evidence relating what is called structural disorder, i.e., the lack of a fixed three dimensional structure, with protein function[53, 54]. Remarkably it has been consistently shown that constituent proteins of biomolecular condensates often belong to the class of intrinsically disordered proteins. We now discuss their properties in comparison to those of globular proteins, trying to understand why are they found in protein condensates.

There are important differences between proteins which have a well defined three dimensional structure (globular proteins) and proteins with intrinsically disordered regions (intrinsically disordered proteins) that do not fold in a three dimensional structure but have an ensemble of configurations. For example, intrinsically disordered proteins are flexible [53], resembling polymer chains. In contrast, globular proteins tend to be rigid

---

[7]This work concerns the Pab1 protein, which seems to form irreversible aggregates after phase separation, thus RNA may be acting here as an aggregation inhibitor

structures, which might be better described as colloids [55]. Another difference is the ability of globular proteins to form crystals, whereas intrinsically disordered proteins cannot [56]. Interestingly both types of proteins undergo liquid-liquid phase separation but with different features; globular proteins undergo liquid phase separation with a condensed liquid phase that is considered to be a metastable state for protein solutions undergoing slow crystallization [57, 58]. It is metastable because the condensed phase will slowly turn into a crystal as the final stable structure. Some intrinsically disordered proteins undergo a stable liquid-liquid phase separation, meaning that having a phase with low protein concentration coexisting with a condensed protein phase will remain so for very long periods of time, hence, in practice we consider this be stable[8]. A last point we would like to mention is that intrinsically disordered proteins have the ability of interacting with many simultaneous partners due to their flexibility and they are known to bind to RNA via specific motifs that are found along their sequences [55]. Curiously, even though there were signals of the ubiquity of structural disorder in proteins for a long time, they were all but ignored. It is now estimated that between 15-45% of eukaryotic proteins are intrinsically disordered [53]. This takes us back to protein phase separation in living organisms, which is also ubiquitous and it was long overlooked. Most of the constituent proteins of biomolecular condensates possess intrinsically disordered regions, characterized by their ability of having multiple simultaneous interacting partners as a consequence of their structural disorder. Interestingly, intrinsically disordered proteins and protein phase separation in living organisms seem to go hand in hand, thus, the fact that intrinsically disordered proteins are polymer-like flexible structures, encourage us to use the models and techniques developed for the study of liquid phase separation in polymer systems in our endeavor to understand the underlying principles driving protein phase separation inside the cell.

In all of the previous sections, we have tried to motivate the interest in studying protein phase separation in the context of living organisms as well as in *in vitro* experiments. We have seen that protein phase separation is a versatile tool for organizing the interior of a cell, and thus, we believe that developing a physical understanding of the processes that give rise to this organization is in order. In the rest of the chapter, we define the basic physical concepts needed to understand protein phase separation, and as a starting point we introduce the Flory-Huggins model [59, 60], a well-known model used to describe the thermodynamic behavior of polymer solutions.

---

[8]Think of a glass of water left on a table, for practical purposes the water is at equilibrium with its vapor. Nevertheless, if we leave the glass of water for long enough sitting on the table, all the liquid water would become vapor, which would mean that the liquid water was actually a metastable state. Thus metastability is in practice a matter of time scales.

1.7   The Flory-Huggins free energy

As we discussed in the previous section, intrinsically disordered proteins may be, to a certain extent considered flexible polymers. Following this line of reasoning, in this work, we describe protein solutions (mixtures) by means of the Flory-Huggins model, which was developed by Huggins [59] and Flory [60] to explain the thermodynamic behavior of polymer mixtures. This model successfully captures the effects of large differences in the molecular volumes of the components of the system. This is a desirable quality in a model given that proteins are much larger than their solvent counterparts. In addition, it takes into account interactions per monomer, which seems reasonable if we want to consider weak protein interactions via flexible segments of intrinsically disordered proteins. We now introduce the Flory-Huggins model in its simplest version and provide a brief explanation of the terms describing the thermodynamics of a binary polymer-solvent mixture.

Let us consider a system composed of two components: a polymer (a large macro-molecule) and solvent (a small molecule like water). The polymers and solvent molecules are placed on a regular lattice which has a fixed volume. Each cell of this lattice has a volume v_0 and the lattice has N cells, thus the total volume of the lattice is given by $V = v_0 N$. The volume of a solvent molecule is equal to the volume of a lattice cell v_0, and polymers are composed of monomers with each of them having a volume v_0. Hence a polymer composed of ℓ monomers occupies ℓ lattice cells and has a volume $v_P = v_0 \ell$. In this model we also consider that the total volume is the sum of the volumes occupied by polymers and solvent molecules, given by $V = v_P N_P + v_0 N_0$, where N_P and N_0 are the particle numbers of polymers and solvent, respectively. If additionally to the previous considerations, we also consider short-range interactions among the components, the system can be described by the free energy of mixing F_{mix}, given by

$$F_{\mathrm{mix}} = k_{\mathrm{B}} T \left[N_P \ln \phi_P + N_0 \ln \phi_0 + \chi N_0 \phi_P \right] \quad , \tag{1.1}$$

where k_{B} is the Boltzmann constant, T is the temperature, $\phi_P = v_P N_P / V$ is the polymer volume fraction, ϕ_0 is volume fraction of the solvent and χ is the Flory-Huggins parameter defined by $\chi = z(2\epsilon_{P0} - \epsilon_{PP} - \epsilon_{00})/2k_{\mathrm{B}}T$, where ϵ_{PP} denotes the monomer-monomer interaction strength, ϵ_{00} the interaction between solvent components, ϵ_{P0} the interaction between monomers and solvent, and z is a value related to the number of neighbors in each lattice site, for a cubic lattice $z = 6$. The logarithmic terms represent the mixing entropy which favors the homogenization of the polymer-solvent mixture. In contrast, the term which contains the χ parameter battles against mixing, this is due to attractive interactions among components of the same kind, which is the most frequent case. The basic mechanisms driving mixing and demixing are captured by these two terms. Before we discuss how a system undergoes liquid phase separation, we introduce a general multicomponent mixture description based on the Flory-Huggins free energy.

1.8   Thermodynamics of multicomponent mixtures

Our main interest in studying the thermodynamics of multicomponent mixtures lies in the fact that the interior of a cell has a multicomponent nature. For example, the cytosol is composed of many different macromolecules (e.g., proteins, RNA and lipids) and ions immersed in water [61, 62]. Although the cytosol is a complex fluid where many of its components are taking part in out-of-equilibrium processes, we consider a multicomponent mixture with a well defined thermodynamic potential as a first step towards understanding some of the processes taking place inside the cell.

For this purpose, let us discuss the thermodynamics of incompressible multicomponent mixtures based on the well-known mean field Flory-Huggins free energy (1.1). The entropy is modeled by the Flory-Huggins mixing entropy and the interactions are considered to be short-ranged [59, 60, 63]. We begin our discussion by considering the multicomponent mixture in the $(T, P, \{N_i\})$ ensemble, where the temperature is denoted by T, the pressure by P, the particle number of molecules of type i by N_i and $\{N_i\}$ is the set of all particle numbers corresponding to the different types of molecules in the system. The corresponding thermodynamic potential is the Gibbs free energy $G(T, P, \{N_i\})$ and its derivatives define the thermodynamic conjugate variables. The entropy is $S = -\left.\partial G/\partial T\right|_{P,\{N_i\}}$, the chemical potential is $\mu_i = \left.\partial G/\partial N_i\right|_{T,P,\{N_{j\neq i}\}}$ and the volume of the system is $V = \left.\partial G/\partial P\right|_{T,\{N_i\}}$ [64]. An incompressible system satisfies $-(1/V)\left.\partial V/\partial P\right|_{T,\{N_i\}} = 0$, which states that the volume of the system is independent of pressure. The volume of an incompressible mixture is thus given by $V = \sum_i v_i N_i$, where the molecular volumes $v_i = \left.\partial V/\partial N_i\right|_{T,\{N_{j\neq i}\}}$ are constants. The Gibbs free energy describing a multicomponent mixture composed of $M + 1$ components can be written as follows [63]

$$G(T, p, \{N_i\}) = k_{\mathrm{B}}T \left(\sum_{i=0}^{M} N_i \ln \phi_i + \frac{V}{2} \sum_{i,j=0}^{M} \frac{\chi_{ij}}{v_0} \phi_i \phi_j \right) + \sum_{i=0}^{M} w_i N_i + p\,V \quad , \qquad (1.2)$$

where k_{B} denotes the Boltzmann constant, $\phi_i = v_i N_i/V$ is the volume fraction of component i, $\chi_{i,j}$ stands for the interaction strength between components of type i and j, w_i are the free energies due to internal degrees of freedom of molecules of component i, and v_0 denotes the molecular volume of the solvent. In its simplest form, the interaction parameters $\chi_{i,j}$ can be expressed as the sum of two contributions $\chi_{ij} = \chi_{ij}^S + \chi_{ij}^H/T$, where χ_{ij}^S is an entropic contribution and χ_{ij}^H is an enthalpic (energetic) contribution. These interaction parameters can be related to the commonly used Flory-Huggins interaction parameter by a simple transformation (see Appendix A). The chemical potentials of the system are given

by

$$\mu_i = k_{\mathrm{B}}T \left(\ln\phi_i + 1 - v_i \sum_{j=0}^{M} n_j + v_i \sum_{j=0}^{M} \frac{\chi_{ij}}{v_0}\phi_j - \frac{v_i}{2} \sum_{j,l=0}^{M} \frac{\chi_{jl}}{v_0}\phi_j\phi_l \right) + w_i + v_i\,P\,, \quad (1.3)$$

where we introduced $n_i = N_i/V$ as the concentration of molecules of type i.

We now define the corresponding free energy $F = G + pV$ in the ensemble where the volume is kept fixed. One subtlety arises from the fact that the system is incompressible. Since the volume of the system is independent of pressure and it only depends on the number of particles, we can arbitrarily eliminate one of the particle numbers when performing the corresponding Legendre transform which changes the description from the $(T, P, \{N_i\})$ ensemble to the $(T, V, \{N_i\})$ ensemble. A common choice and the one that we also use throughout this work is to eliminate the solvent particle number N_0 from the description and express it as a function of the volume and the rest of the particle numbers as $N_0 = (V - \sum_{i=1}^{M} v_i N_i)/v_0$. An immediate consequence of eliminating the solvent component is the appearance of new thermodynamic conjugate variables in the ensemble of fixed volume. To show that this is the case, let us first remember that the differential of the Gibbs free energy is $dG = -SdT + VdP + \sum_{i=0}^{M} \mu_i dN_i$ [64]. Using this expression we can calculate the differential $dF = d(G + pV)$ by eliminating the solvent from the description we obtain

$$dF(T, V, \{N_{i\neq0}\}, N_0(V, \{N_{i\neq0}\})) = -SdT - \left(P - \frac{\mu_0}{v_0} \right) dV$$
$$+ \sum_{i=1}^{M} \left(\mu_i - \frac{v_i}{v_0}\mu_0 \right) dN_i \quad, \qquad (1.4)$$

where on the left hand side we show the dependence of the solvent particle number on the volume and the rest of the particle numbers as $N_0(V, \{N_{i\neq0}\})$, and on the right hand side we use the relation $dN_0 = (dV - \sum_{i\neq0} v_i N_i)/v_0$. From Eq. (1.4) we identify the new thermodynamic conjugate variables of the volume and the particle numbers (the conjugate variable to the temperature T does not change, i.e. it still is the entropy S). These thermodynamic conjugate variables are

$$\Pi = P - \mu_0/v_0 \quad, \qquad\qquad (1.5)$$
$$\bar{\mu}_i = \mu_i - v_i\,\mu_0/v_0 \quad, \qquad\qquad (1.6)$$

where we introduced the osmotic pressure Π [65] and the exchange chemical potential $\bar{\mu}_i$ of molecules of type i with respect to the solvent. To clarify the meaning of the osmotic pressure Π, let us think of a vessel divided into two compartments separated by a semi-permeable membrane that only allows solvent to pass through. One of the compartments

is filled with pure solvent, the other one with a mixture containing other molecules and solvent. The osmotic pressure is then defined as the pressure that needs to be applied to the mixture to stop the solvent from flowing through the membrane [66]. The exchange chemical potential of component i, $\bar{\mu}_i$, defines the free energy cost of exchanging one molecule of component i for a number of solvent molecules occupying the same volume. Using Eqs. (1.5)-(1.6) we then write the differential of the free energy at fixed volume, Eq. (1.4), in a compact form as

$$dF(T, V, \{N_i\}) = -SdT - \Pi dV + \sum_i \bar{\mu}_i dN_i \quad , \tag{1.7}$$

where the index i runs through all components except for the solvent. From now on, when we discuss a system in the fixed volume ensemble, the indices in the sums will be taken to run through all components except for the solvent, unless otherwise stated. We define the free energy density at fixed volume f in the following equation

$$F(T, V, \{N_i\}) = V f(T, \{n_i\}) \quad , \tag{1.8}$$

where we used the fact that the free energy F is a homogeneous function of the first order (i.e. an extensive quantity) [64]. In our studies of phase separation we mainly focus on the properties of the free energy density f, because its convexity defines the stability of the system. We use Eq. (1.7) and Eq. (1.8) to express the osmotic pressure and the exchange chemical potential in terms of the free energy density as

$$\Pi(T, \{n_i\}) = \sum_i n_i \left. \frac{\partial f}{\partial n_i} \right|_{T, \{n_{j \neq i}\}} - f \quad , \tag{1.9}$$

$$\bar{\mu}_i(T, \{n_i\}) = \left. \frac{\partial f}{\partial n_i} \right|_{T, \{n_{j \neq i}\}} . \tag{1.10}$$

We now explicitly write the free energy density $f = G/V + P$ with G given by Eq. (1.2), which reads

$$\begin{aligned} f(T, \{n_i\}) = {} & k_{\mathrm{B}}T \left(\sum_i n_i \ln \phi_i + \frac{1}{2v_0} \sum_{i,j} \chi_{ij} \phi_i \phi_j \right) + \sum_i w_i \phi_i \\ & + {} k_{\mathrm{B}}T \left(\frac{(1 - \sum_i \phi_i)}{v_0} \ln \left(1 - \sum_i \phi_i \right) + \sum_i \frac{\chi_{i0}}{v_0} \phi_i \left(1 - \sum_i \phi_i \right) \right) \\ & + {} w_0 \left(\frac{1 - \sum_i \phi_i}{v_0} \right) \quad . \end{aligned} \tag{1.11}$$

The second and third lines show contributions where the solvent is involved, in such contributions the volume fraction of the solvent has been expressed in terms of the rest of the

components as $\phi_0 = 1 - \sum_i \phi_i$ to show explicitly that the solvent has been eliminated. We would like to stress that in an incompressible system the volume fractions and concentrations can be used interchangeably to describe the system since they are connected via $\phi_i = v_i n_i$.

In this section we have introduced all the relevant thermodynamic concepts for studying multicomponent mixtures. In the next section we use these concepts to discuss how liquid phase separation emerges in a multicomponent description.

1.9   Liquid phase separation as a thermodynamic instability

The origin of liquid-liquid phase separation can be understood using a simple argument in terms of the free energy F of the system. The system will separate into two different liquid phases if the sum of the free energies of both phases is lower than that of a well mixed homogeneous phase. The previous statement can be expressed in terms of the free energy density as

$$V^{\mathrm{I}} f\left(T, \{n_i^{\mathrm{I}}\}\right) + V^{\mathrm{II}} f\left(T, \{n_i^{\mathrm{II}}\}\right) < V f\left(T, \{n_i\}\right) \quad , \tag{1.12}$$

where T is the temperature, V^{I} and V^{II} are the volumes of the two different phases which add up to the total volume $V = V^{\mathrm{I}} + V^{\mathrm{II}}$ and the number of particles of each component in the two different phases $N_i^{\mathrm{I}} = V^{\mathrm{I}} n_i^{\mathrm{I}}$ and $N_i^{\mathrm{II}} = V^{\mathrm{II}} n_i^{\mathrm{II}}$ add up to the total particle number $N_i = V n_i$. We can interpret Eq. (1.12) as a violation of the convexity of the free energy, which leads to a thermodynamic instability [64]. The local stability of the free energy in concentration space can be discussed in terms of the second derivative matrix of the free energy density

$$\mathbf{H} = \frac{\partial^2 f}{\partial n_i \partial n_j} \quad , \tag{1.13}$$

where $\mathbf{H}$ is the Hessian of the free energy density. A composition (a point in the concentration space) is considered locally stable if $\mathbf{H}$ evaluated at that composition has all eigenvalues being positive. If one of the eigenvalues becomes zero for a given composition, the system is locally unstable and exhibits a thermodynamic instability [67]. The loci of points for which $\det \mathbf{H} = 0$, is called the spinodal and can be understood as the boundary between locally stable and locally unstable points.

The convexity of the free energy defines which compositions are thermodynamically stable and which ones are not, however, the convexity criterion alone is not enough for finding the compositions that coexist with each other. In the following we find the conditions defining the equilibrium between coexisting phases.

1.10   Phase coexistence conditions

We now derive the conditions for phase coexistence in a multicomponent mixture. We begin by specifying a functional $\mathcal{F}$, which accounts for the free energy contributions from the two different phases in addition to the conservation of particles and the conservation of volume included as constraints enforced by two Lagrange multipliers. The functional to be minimized reads

$$
\begin{aligned}
\mathcal{F} &= V^{\mathrm{I}} f(T, \{n_i^{\mathrm{I}}\}) + V^{\mathrm{II}} f(T, \{n_i^{\mathrm{II}}\}) \\
&+ \sum_i \bar{\mu}_i (N - V^{\mathrm{I}} n_i^{\mathrm{I}} - V^{\mathrm{II}} n_i^{\mathrm{II}}) - \Pi(V - V^{\mathrm{I}} - V^{\mathrm{II}}) \quad ,
\end{aligned}
\tag{1.14}
$$

where the Lagrange multipliers $\bar{\mu}_i$ and Π enforce the constraints of conservation of particles and conservation of volume, respectively. We specifically choose the symbols denoting the exchange chemical potential and the osmotic pressure precisely because, as we show below, the Lagrange multipliers enforcing the constraints are defined by the exchange chemical potential and the osmotic pressure. In order to find the coexistence conditions for a multicomponent mixture with fixed volume V and a fixed number of particles of each component N_i, we must find the minimum of the functional $\mathcal{F}$. The conditions for a minimum of $\mathcal{F}$ are

$$
\frac{1}{V^{\mathrm{I}}} \frac{\partial \mathcal{F}}{\partial n_i^{\mathrm{I}}} = \frac{\partial f^{\mathrm{I}}}{\partial n_i^{\mathrm{I}}} - \bar{\mu}_i = 0 \quad ,
\tag{1.15}
$$

$$
\frac{1}{V^{\mathrm{II}}} \frac{\partial \mathcal{F}}{\partial n_i^{\mathrm{II}}} = \frac{\partial f^{\mathrm{II}}}{\partial n_i^{\mathrm{II}}} - \bar{\mu}_i = 0 \quad ,
\tag{1.16}
$$

$$
\frac{\partial \mathcal{F}}{\partial V^{\mathrm{I}}} = f^{\mathrm{I}} - \sum_i \bar{\mu}_i n_i^{\mathrm{I}} + \Pi = 0 \quad ,
\tag{1.17}
$$

$$
\frac{\partial \mathcal{F}}{\partial V^{\mathrm{II}}} = f^{\mathrm{II}} - \sum_i \bar{\mu}_i n_i^{\mathrm{II}} + \Pi = 0 \quad ,
\tag{1.18}
$$

where for brevity we used $f^{\mathrm{I}} = f(T, \{n_i^{\mathrm{I}}\})$ and $f^{\mathrm{II}} = f(T, \{n_i^{\mathrm{II}}\})$. We do not include the derivatives with respect to the Lagrange multipliers since these derivatives are the conservation laws. By solving for a minimum of $\mathcal{F}$ we find the following conditions

$$
\bar{\mu}_i = \frac{\partial f^{\mathrm{I}}}{\partial n_i^{\mathrm{I}}} = \frac{\partial f^{\mathrm{II}}}{\partial n_i^{\mathrm{II}}} \quad ,
\tag{1.19a}
$$

$$
\Pi = \sum_i \frac{\partial f^{\mathrm{I}}}{\partial n_i^{\mathrm{I}}} n_i^{\mathrm{I}} - f^{\mathrm{I}} = \sum_i \frac{\partial f^{\mathrm{II}}}{\partial n_i^{\mathrm{II}}} n_i^{\mathrm{II}} - f^{\mathrm{II}} \quad .
\tag{1.19b}
$$

These conditions have a simple interpretation. Eq. (1.19a) states that the exchange chemical potential for every component must be the same in both phases and Eq. (1.19b) states

that the osmotic pressure in both phases must be balanced. We could add one more inter-
pretation of the coexistence conditions. In order to do this, let us rewrite Eq. (1.19b) in
the following way

$$f^{\mathrm{II}} - f^{\mathrm{I}} - \sum_i \bar{\mu}_i (n_i^{\mathrm{II}} - n_i^{\mathrm{I}}) = 0 \quad , \tag{1.20}$$

where we write $\bar{\mu}_i$ without any distinction between phases because it has the same value
in both of them. Interestingly, Eq. (1.20) can be interpreted geometrically as the equation
defining a common tangent hyper-plane in the concentration space. This shows that phases
which coexist live in a common tangent space of the free energy density f. For a system
with one independent component, Eq. (1.20) reduces to the well known common tangent
construction [68], which is the case that is usually presented in textbooks. Another relevant
concept is that of the binodal, which for a fixed set of intensive parameters (in this case
temperature) is defined as the set of all possible coexisting phases and it can be found by
solving for all the compositions fulfilling the coexistence conditions (1.19). Any composi-
tion lying inside of the region bounded by the binodal will undergo phase separation and
any composition outside of this region will remain homogeneously mixed. The particular
composition where the binodal and the spinodal meet, is called the critical point. The
critical point is the particular point in which phases that coexist become so similar that
they are completely indistinguishable from each other. In Appendix B we discuss general
local stability conditions and their connection to the criterion for finding the critical points
in multicomponent mixtures.

1.11   Phase diagram for a binary mixture

Let us now give a simple example to illustrate how the convexity of the free energy is
affected by changing the thermodynamic state (temperature or composition) of a binary
mixture. In this example we consider a binary mixture composed of a molecule with volume
fraction ϕ and of solvent (e.g. water) with volume fraction $1 - \phi$. For convenience, we
define the free energy per site $f_s = v_0 f$ (a common name coming from considering the
system in a lattice), where v_0 is the molecular volume of the solvent and f is the free
energy density describing the binary mixture. The free energy density per site reads [65]

$$f_s = k_{\mathrm{B}} T \left[\epsilon \phi \ln \phi + (1 - \phi) \ln(1 - \phi) + \chi \phi (1 - \phi) \right] \quad , \tag{1.21}$$

where χ is the Flory-Huggins interaction parameter and $\epsilon = v_0/v$ is the ratio between the
molecular volume of the solvent and the molecular volume of the macromolecule v. For
simplicity, we discuss the case for $\epsilon = 1$, which is called a regular solution. The system we
consider is one where there are only enthalpic interactions, i.e., $\chi = \chi^H/T$. For this system
high temperatures lead to mixing whereas low temperatures lead to phase separation. The

critical point can be found by solving the critical conditions, which for a one-component system are given by: $\partial^2 f / \partial \phi^2 = 0$ and $\partial^3 f / \partial \phi^3 = 0$ (see Appendix B). Solving these conditions one obtains a critical value for the temperature T_c and the volume fraction ϕ_c, which for our simple case are given by $T_c = \chi^H / 2$, and $\phi_c = 1/2$. The osmotic pressure Π and the exchange chemical potential of the molecule $\bar{\mu}$, can be calculated by means of Eqs. (1.9) and (1.10).

We now discuss the phase behavior of the system. We do so by analyzing $\bar{\mu}$ (Fig. 1.6(a)), Π (Fig. 1.6(b)) and the free energy density per site (Fig. 1.6(c)) as a function of the volume fraction ϕ for different temperature values. For $T > T_c$, the exchange chemical potential and the osmotic pressure are monotonic increasing functions of ϕ (blue lines in Fig. 1.6(a,b)). In this case, the free energy is convex for all values of ϕ (blue line in Fig. 1.6(c)), thus the system is always homogeneously mixed. For $T = T_c$, the free energy (black line in Fig. 1.6(c)) has zero curvature at the critical volume fraction $\phi = \phi_c$ (black point in Fig. 1.6(c)), which is reflected in inflection points in both $\bar{\mu}$ and Π (black points on top of the black lines Fig. 1.6(a,b)). This is the signature of a thermodynamic instability. For $T < T_c$, there is a non-monotonic behavior of $\bar{\mu}$ (orange line in Fig. 1.6(a)) and Π (orange line in Fig. 1.6(b)) which means that there are regions where the free energy is locally non-convex. In this case, we can perform a common tangent construction (green line in Fig. 1.6(c)) on the free energy density (orange line in Fig. 1.6(c)) and find the coexisting phases (green points in Fig. 1.6(c)). At this temperature, the system becomes locally unstable for a volume fraction range where the curvature of the free energy is negative, the boundary between the locally stable and unstable volume fractions gives the spinodal points (stars in Fig. 1.6(c)). These points (stars in Fig. 1.6(a,b)) bound the regions for which the slope of both $\bar{\mu}$ and Π, is negative. The coexisting phases have equal values of exchange chemical potentials and osmotic pressures, which ensures thermodynamic equilibrium between the phases.

The phase behavior of the system can be summarized in a phase diagram (Fig. 1.6(d)). The curve defined by all the coexisting phases is called the binodal line (green solid line in Fig. 1.6(d)), everything outside the binodal is a stable one-phase mixture, whereas everything within the binodal undergoes a demixing transition. The coexisting phases (green points in Fig. 1.6(d)) are connected by tie lines, which in this case are horizontal lines (green dotted line Fig. 1.6(d)) since the temperature must be the same in both phases. Any volume fraction lying on a tie line will split into the corresponding coexisting phases from which the tie line is spanned. The volume occupied by each of the phases can be calculated using the lever rule [65] which can be understood as follows: the fraction of the system in each phase bears a relation to the distance of the initial homogeneous volume fraction to the volume fraction of each phase. The largest fraction will be in the phase whose volume fraction is closer to the initial volume fraction, thus if the binodal have just been crossed, there will be an incipient phase (the one whose volume fraction is further
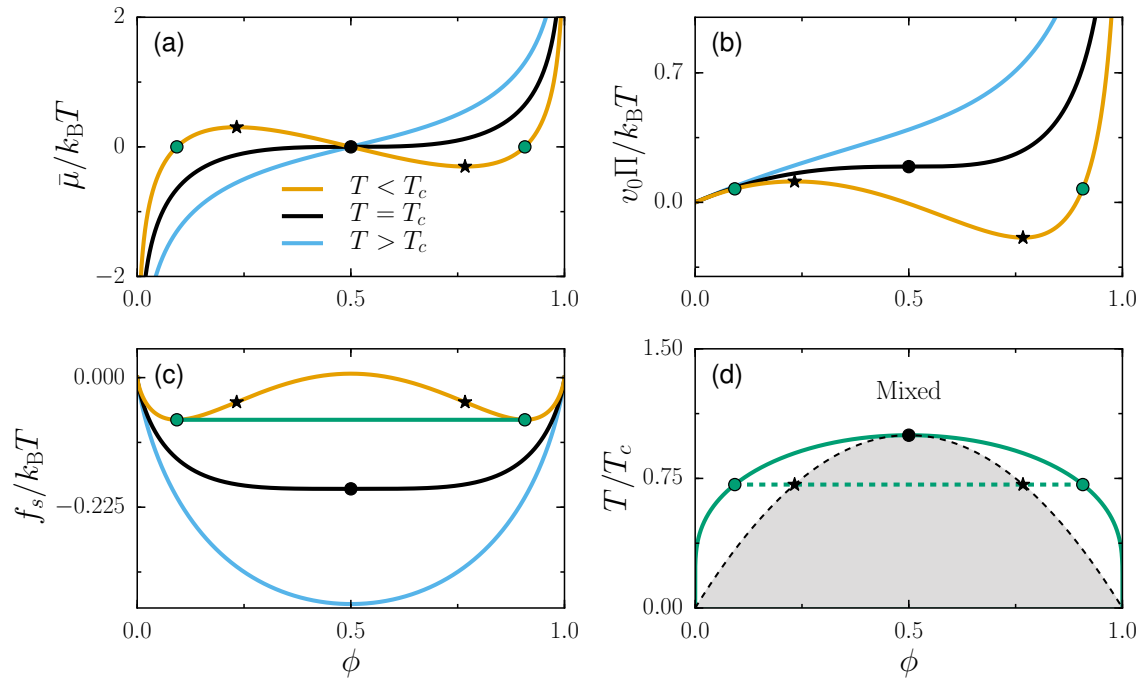
FIGURE 1.6: Phase behavior of a binary mixture. (a) Exchange chemical potential. (b) Osmotic pressure. (c) Free energy density per site. All three quantities are plotted as a function of the volume fraction for three different temperatures: $T < T_c$ (orange lines), $T = T_c$ (black lines), and $T > T_c$ (blue lines). In (c) the common tangent (green line) connecting the coexisting phases is shown. (d) Phase diagram for a binary mixture. The metastable region (light green shaded region), the spinodal region (dark green shaded region), and the corresponding tie line for $T < T_c$ (green dotted line) connecting two coexisting phases are shown. In all panels, the black circles correspond to a critical point, the black stars correspond to points located on the spinodal and the green points represent coexisting phases.

away) and one that will almost occupy the whole volume. The critical point is located at the meeting point of the binodal and the spinodal (dotted black line in Fig. 1.6(d)). It has the unique property that if one approaches the binodal from below the critical temperature for a system with critical volume fraction ϕ_c, the coexisting phases become more and more similar until the critical point is reached and the two phases become indistinguishable. This transition is called a continuous phase transition, or a second order phase transition. If we were to do the same procedure with any other volume fraction located within the region bounded by the binodal, going from low temperature to high temperature, what we would find is a sudden change from having two coexisting phases, where one of the phases will suddenly vanish and the other will occupy the whole volume in the system, this is called a discontinuous transition or a first order transition.

For each pair of coexisting phases, the one with the smallest volume fraction is called the dilute phase whereas the one with the largest volume fraction is called condensed phase. Thus, in a simple binary mixture, the system separates in a molecule poor phase and a molecule rich phase (in general terms, a macromolecule poor phase and a macromolecule

rich phase)[9]. Furthermore, inside the binodal we can distinguish two different regions, the spinodal region (dark green shaded region in Fig. 1.6(d)), bounded by the spinodal curve (dotted black line in Fig. 1.6(d)) and the region between the binodal curve and the spinodal curve, which is called metastable region (light green shaded region in Fig. 1.6(d)). In this thesis the difference between these two regions is not relevant since we only study systems that have reached thermodynamic equilibrium. The differences between these regions can be observed in the dynamic relaxation towards equilibrium for a system that is initially set in a volume fraction located within the binodal. The two regimes show very different growth of the coexisting phases. In the metastable region, spherical droplets of the condensed phase are nucleated and grow by both fusion and a process called Ostwald ripening where big droplets grow at the expense of the smaller ones due to chemical potential differences. In contrast, if the volume fraction is inside the spinodal region, the system undergoes rapid changes in what is called spinodal decomposition until it reaches thermodynamic equilibrium where there are two phases defined by the tie line which crosses the initial composition.

We decided to present this lengthy discussion on how to read a phase diagram since they will be heavily used throughout this work. Besides, by constructing the phase diagram of a system one can predict its behavior for changes in compositions or other variables such as temperature or pH. In Chapter 2 we discuss how to numerically construct phase diagrams.

To summarize the previous three sections, we have analyzed the thermodynamics of multicomponent mixtures, presented the thermodynamic criterion for the occurrence of a demixing transition and derived the coexistence condition. We believe that this thermodynamic framework can shed light on the mechanisms driving phase separation in the cell, particularly, in this work we use equilibrium thermodynamics to study protein phase separation on different scenarios where different reactions such as binding or protonation and deprotonation of a protein are relevant. For this reason, we now briefly introduce the standard knowledge of chemical reactions and chemical equilibrium.

1.12   Chemical equilibrium in multicomponent mixtures

Mixtures composed of many different components often undergo chemical reactions. We define a set of chemical reactions by a stoichiometry matrix with elements ν_{ij}. This contains all the information about changes in compositions due to chemical reactions. More precisely, the matrix element ν_{ij} corresponds to the number of molecules of component j that are transformed in the i-th chemical reaction. The convention we use is that ν_{ij} is negative if the component j is a reactant and negative if it is a product (the signs are just a convention and can be reversed). For a system with s components, the chemical reactions

---

[9]This is not always the case for multicomponent systems, where the composition of the phases strongly depend on the interactions in the system.

can be expressed as

$$\sum_{j=1}^{s} \nu_{ij} A_j \rightleftharpoons 0 \quad , \quad \text{for } i \in \{1, \ldots, r\} \quad , \tag{1.22}$$

where A_j denotes the chemical symbol for component j, and r is the number of independent chemical reactions in the system.

A system undergoing any type of reversible reactions attains chemical equilibrium when the system reaches its free energy minimum, for which the net change in the number of components is zero on average. The conditions defining the chemical equilibrium can be found by different methods [64, 69]. In particular, in Chapter 3 we discuss the chemical equilibrium conditions in terms of conservation laws associated to the stoichiometry matrix for an arbitrary number of components and reactions. However, for the time being, we just state the chemical equilibrium conditions, which are given by

$$\sum_{j=1}^{s} \nu_{ij} \mu_j = 0 \quad , \quad \text{for } i \in \{1, \ldots, r\} \quad , \tag{1.23}$$

where $\mu_j{}^{10}$ is the chemical potential of component j. Considering the concentration dependence of the chemical potentials (see Eq. (1.3)), it follows that the chemical equilibrium conditions given in Eq. (1.23) provide r relations between the concentrations of the s components of the system. The compositions that the system can attain at chemical equilibrium define a surface (hyper-surface or manifold) in the s-dimensional concentration space, which we refer to as the chemical equilibrium surface.

In the next section we show that in phase-separating systems which undergo chemical reactions, not all the coexisting phases can be reached nor all the chemical equilibrium states, the chemical equilibrium conditions and the chemical reactions must be simultaneously fulfilled in order for different phases to coexist.

1.13   Phase coexistence at chemical equilibrium

Many systems capable of exhibiting phase separation may also undergo different types of reactions [70–73]. In particular we have seen in Sect. 1.2 and in Sect. 1.3 that inside the cell, many biochemical processes are organized in biomolecular condensates which seem to form via liquid-liquid phase separation [3]. In addition, it is known that the interior of a cell is constantly undergoing chemical changes [74, 75], some of which are clearly out-of-equilibrium processes that we cannot address using the techniques developed in this work. However, there might be some processes that are mostly driven by local

---

[10]In the case of components with a well defined chemical symbol like H_2O, the chemical potential will be expressed as μ_{H_2O} for clarity.

equilibrium thermodynamics which can be then studied using a thermodynamic equilibrium perspective.

In Sect. 1.10 we have derived the coexistence conditions, which essentially select those compositions of the system fulfilling a balance between the exchange chemical potentials and osmotic pressures of the two phases. Since chemical equilibrium imposes extra constraints to the system, the system will demix only if both the chemical equilibrium conditions and the phase coexistence conditions are simultaneously fulfilled.

In order to show an example of a system that undergoes simultaneous chemical reactions and phase separation, we consider a ternary mixture model with only one chemical reaction. The system we consider is composed of a solvent and a molecule that exists into two different states, one where the molecule does not have any tendency to phase separate and another where the molecule can interact with molecules in the same state. We denote these two states by A and A', respectively. They transform obeying the following reaction

$$A \rightleftharpoons A' \quad . \tag{1.24}$$

This ternary system can be described using the following free energy density per site

$$
\begin{aligned}
f_s \;=\; & k_\mathrm{B}T \left(\phi_A \ln \phi_A + \phi_{A'} \ln \phi_{A'} + (1 - \phi_A - \phi_{A'}) \ln(1 - \phi_A - \phi_{A'}) + \chi \phi_{A'}^2 \right) \\
& + w_A \phi_A + w_{A'} \phi_{A'} + w_0 (1 - \phi_A - \phi_{A'}) \quad ,
\end{aligned}
\tag{1.25}
$$

where ϕ_A is the volume fraction of A, ϕ'_A is the volume fraction of A', χ is the interaction strength between A' molecules; w_A, $w_{A'}$ and w_0 are the corresponding free energies due to internal degrees of freedom of A, A', and the solvent, respectively. The chemical equilibrium condition for the reaction (1.24) is simply: $\mu_A = \mu_{A'}$, which leads to[11]

$$\phi_A = \phi_{A'} \exp \left(\frac{w_{A'} - w_A}{k_\mathrm{B}T} + \chi \phi_{A'} \right) \quad , \tag{1.26}$$

where we can see that the interplay between the free energies due to internal degrees of freedom and the interaction strength defines the compositions at chemical equilibrium.

For a fixed set of parameters, T, w_A, w'_A and χ, we first find the binodal by solving the coexistence conditions given in Eq. (1.19) (green solid line in Fig. 1.7) as if there were no chemical reactions. In the absence of chemical reactions, the system would undergo phase separation for any composition located within the binodal (green shaded region in Fig. 1.7)), whereas if the system can react via Eq. (1.24), the only phases that will coexist are those found at the intersection between the chemical equilibrium curve (black dotted line) and the binodal (green points in Fig. 1.7). We see then how chemical equilibrium

---

[11]For an incompressible system, the chemical equilibrium conditions can also be expressed by substituting μ_i for the exchange chemical potentials $\bar{\mu}_i$ without further change.
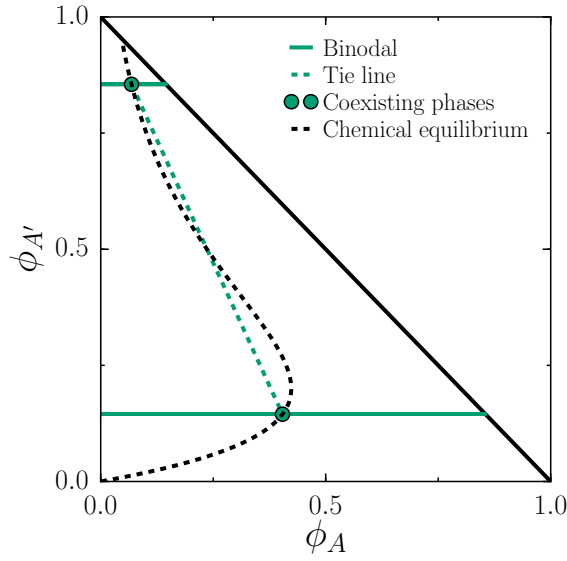
FIGURE 1.7: Phase separation at chemical equilibrium. The binodal line (green solid line) and chemical equilibrium curve (dotted black line) are shown. Their intersections define the coexisting phases of the system at chemical equilibrium (green points). One tie line (green dotted line) connecting the coexisting phases and the demixing region (green shaded area) for the system in the absence of chemical reactions are shown. Parameters: $w_0 = 0$, $w_A - w'_A = 7\,k_B T_0$, $\chi^S = 1.5$, $\chi^H = 4\,T_0$ and $T/T_0 = 4$. With $\chi = \chi^S + \chi^H/T$ and T_0 a reference temperature.

heavily restricts the compositions that the system can attain, nevertheless, as we have shown here, we can find coexisting phases for a system undergoing chemical reactions.

The procedure to find coexisting phases for a system undergoing chemical reactions is as follows: one first needs to construct the equilibrium surface and the binodal corresponding to the free energy as if there were no chemical reactions. Then, one needs to find the intersections between the binodal and the chemical equilibrium curve (surface) in order to find the coexisting phases. Therefore, the first step becomes increasingly harder if the system has a large number of components, because one has to deal with constructing a phase diagram in a high dimensional space. In Chapter 3 we deal with this issue from a different perspective. We base our description on the conserved components associated to the stoichiometry matrix [69, 76, 77], using their corresponding particle numbers as natural variables, the free energy can be analyzed in a lower dimensional space compared to the original one, which is equal to the number of chemical components. To put it clearly, if a system has s components and r independent chemical reactions, the construction of the coexisting phases can be done in the reduced dimension $c = s - r$, where c is the number of independent conserved components.

1.14   Overview of the Thesis

In this thesis we study the effects of chemical and binding reactions on liquid-liquid phase separation at chemical equilibrium. To do so, we develop a generic framework to study liquid phase separation in systems undergoing different reactions. One may wonder, is it relevant to study systems at thermodynamic equilibrium in order to understand what happens inside of a living cell? We are tempted to give a positive answer. We back this claim by showing that using minimal models we can account for diverse phenomena such as the mechanism behind the localization of P granules explained in Sect. 1.3 and the pH-dependent response of different proteins which undergo phase separation as suggested in Sect. 1.4. We expect to convey this message in the rest of the thesis in the clearest possible way.

This thesis is organized as follows: In Chapter 2 we describe the numerical methods that we use to construct phase diagrams for multicomponent mixtures. We then discuss what are the effects that different parameters such as molecular volumes and interaction strength have in phase diagrams of binary and ternary mixtures. We conclude by discussing two realistic scenarios that could be relevant in the context of protein condensates. In Chapter 3, we study liquid-liquid phase separation at chemical equilibrium. In particular, we discuss the conservation laws in chemical reacting systems and construct independent conserved components starting from the stoichiometry matrix. Then, using the particle numbers corresponding to these conserved components as independent variables, we show how new thermodynamic conjugate variables can be suitably defined at chemical equilibrium. We conclude by discussing the construction of phase diagrams for chemically reacting systems and giving a simple example. In Chapter 4, we study the effect of pH on macromolecular phase separation. We do so by defining a model in which the charge state of the macromolecules is controlled by means of protonation and deprotonation reactions, which in turn are set by the pH of the solution (mixture). Using this scenario where macromolecules can be protonated and deprotonated, the phase-separating tendency of the macromolecules is modulated by the composition of the mixture, which is in turn controlled by the pH. Building on the concepts developed in Chapter 4, we define the corresponding free energy for a fixed pH ensemble and study phase diagrams as a function of temperature and pH. We conclude this chapter by discussing a more realistic scenario which seems to be in agreement with the observations of phase separation inside yeast cells discussed in Sect. 1.4. In Chapter 5, we focus on studying the phase separation behavior of PGL-3, an important component of P granules, as well as the effect of RNA on its tendency to phase separate. In this chapter we show experimental evidence of the liquid-like properties of PGL-3 drops and discuss the RNA-binding properties of PGL-3 and MEX-5. Guided by the experimental data, we construct a minimal model where we couple a phase-separating system with RNA-binding reactions accounting for the abilities

of PGL-3 and MEX-5 to bind to RNA. We then compare this model with the experimental data and obtain the parameters that best describe the experimental observations. What we find is that our minimal model has the necessary elements to explain the P granule localization to the posterior side in the *C. elegans* embryo. In Chapter 6, we discuss the findings of this work and discuss possible future directions of research that could take our work as a starting point.

Chapter 2

Construction of Phase Diagrams

Phase diagrams are useful representations of the thermodynamic behavior of a system. The construction of phase diagrams allow us to predict what would happen when some thermodynamic variables of the system are changed. In the context of protein phase separation, there have been studies of protein solutions *in vitro* were the authors have measured phase diagrams of protein solutions [23, 28, 30, 34, 78] as a function of salt concentration, mRNA and pH, among others. Thus, we are interested in describing how to construct a phase diagram, not only by stating the conditions that must be fulfilled but by describing the numerical method that we use in order to construct phase diagrams.

In this chapter we discuss numerical methods to find the coexisting phases based on the convexity of the free energy and the coexistence conditions presented in Chapter 1. We use these methods to discuss the properties of binary and ternary macromolecular mixtures by showing a variety of phase diagrams and their dependence on the intrinsic properties of macromolecules like their size or the interaction strength between them. We then conclude this chapter by discussing possible phase diagrams for key proteins in the formation of P granules, namely PGL-3 and PGL-1. We use previous results published in [28] and predict possible binodal lines for a PGL-3 solution and for a mixture containing PGL-3 and PGL-1.

2.1   Convex hull construction to find coexisting phases

A standard procedure to find coexisting phases is to use a root finder and solve for the concentrations that fulfill the coexistence conditions given in Eqs. (1.19a) and (1.19b) [79, 80]. This approach needs an initial guess. For systems with many different components a good initial guess is a requirement for convergence of the root solver. We instead find the coexisting phases by means of a geometrical construction called the convex hull [81, 82]. The convex hull of a set of points is defined as the smallest convex set that contains all the points of the set [83]. In our case, the convex hull of the free energy density determines

the region where the homogeneous states, which are defined by the concentration of each component, are thermodynamically stable. We call these states the mixed states. The part of the free energy which is not part of the convex hull represents the states that undergo a demixing transition and therefore we call them demixed states.

We now show how to construct coexisting phases by means of the convex hull for a binary mixture described by Eq. (1.21). As a reminder, the free energy per site is defined as $f_s = v_0 f$ with f the free energy density and v_0 the molecular volume of the solvent, and it reads

$$f_s = k_{\mathrm{B}}T \left[\epsilon\phi \ln \phi + (1 - \phi)\ln(1 - \phi) + \chi\phi(1 - \phi) \right] \quad , \tag{2.1}$$

where T is the temperature, ϕ is the volume fraction of the macromolecule (intrinsically disordered protein), $\epsilon = v/v_0$ is the ratio of the molecular volumes of the macromolecule v, and that of the solvent, and $\chi = \chi^S + \chi^H/T$ is a Flory-Huggins parameter with an entropic contribution χ^S, and an enthalpic contribution χ^H. The free energy is shown in Fig. 2.1. Choosing the temperature $T > T_c$, or in terms of the interaction parameter $\chi > \chi_c$, results in a non-convex free energy, see Fig. 2.1(a). The states in the convex hull of the free energy density are the states which remain mixed (blue line in Fig. 2.1(b)). In order to construct the coexisting phases, drawing a line connecting the states at the interior boundaries of the convex hull gives a natural construction of the common tangent for coexisting phases. In one dimension, the construction of the convex hull and the identification of its internal boundaries is sufficient to find the coexisting phases. The simplicity of this approach is that no prior knowledge is needed besides the values of the free energy as a function of the composition variable (in this case the volume fraction ϕ). This construction does not require any derivatives or initial guesses in order to find a pair of coexisting phases. We use the *Quickhull* algorithm routine implemented in *Scipy* to construct the convex hull. A complete description of the algorithm can be found in [83] and a simple example is given in Appendix C.

Constructing coexisting phases for a multicomponent mixture becomes increasingly difficult because of the growing number of variables and equations that need to be solved. For a ternary mixture one would need to solve three non-linear equations simultaneously, in this case, a good initial guess becomes crucial. Using the convex hull can provide us with a good initial guess to solve the coexistence conditions and a combination of a convex hull construction and a root solver allows us to construct the full phase diagram. Let us now examine how to use the convex hull to construct the coexisting phases in the case of a ternary mixture described by:

$$f_s/k_{\mathrm{B}}T = \phi_1 \ln \phi_1 + \phi_2 \ln \phi_2 + (1 - \phi_1 - \phi_2)\ln(1 - \phi_1 - \phi_2) + \chi\phi_2(1 - \phi_1 - \phi_2) \quad , \tag{2.2}$$

where ϕ_1 and ϕ_2 are the volume fractions of two different components, and where we
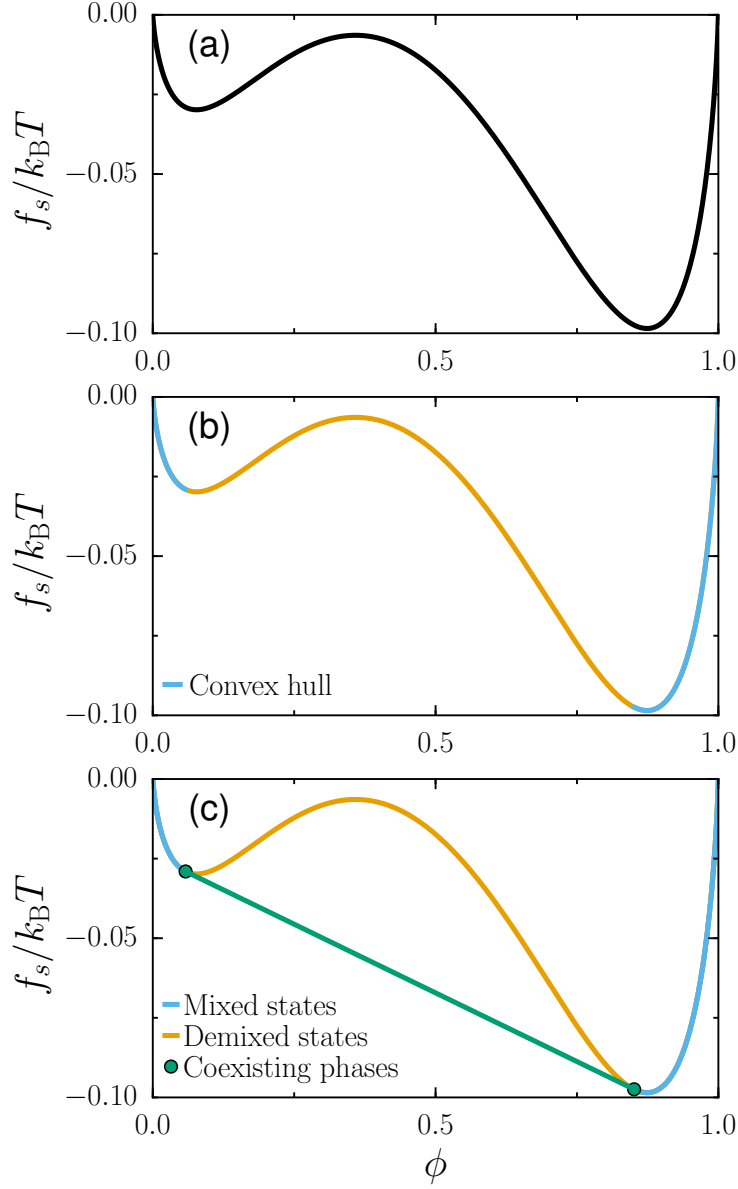
FIGURE 2.1: Common tangent construction using the convex hull. (a) The dimensionless free energy per site $f_s/k_\mathrm{B}T$ from Eq. (1.21) is shown for $\chi > \chi_c$. (b) Convex hull of the dimensionless free energy per site (blue line), the rest of the states do not belong to the convex hull (orange line). (c) The common tangent is constructed by connecting the points in the interior boundary of the convex hull, the points that are connected by the common tangent line (green line) are the coexisting phases (green circles). We choose the molecular volume ratio $\epsilon = 1/2$.

considered only one Flory-Huggins parameter describing the tendency of ϕ_2 to phase-separate from the solvent. In this case, the free energy density represents a surface which depends on two composition variables (ϕ_1 and ϕ_2), see Fig. 2.2(a). We first construct the convex hull of the free energy density, which is the set of states representing the convex region of the free energy density, or equivalently the set of states which would remain homogeneously mixed. These states correspond to the blue points in Fig. 2.2(a), the orange points denote the set of states which will undergo a demixing transition. After constructing the convex hull of the free energy, we project the mixing and demixing states into the composition plane $\phi_1 - \phi_2$, see Fig. 2.2(b). If we were only interested in finding the region where the system will undergo a demixing transition but not necessarily in knowing the composition of the phases that coexist, the convex hull alone provides this information. We may go further and find the boundary between the mixing and demixing regions given by the green line in Fig. 2.2(c). This green line corresponds to an approximated binodal. So far, we have found an approximation for the binodal line but we do not know which phases coexist with each other. The final step is to construct the tie lines connecting the coexisting phases. We construct them by using the approximate binodal as initial guess for a root solver method (implemented in *Scipy*). This allow us to solve the coexistence conditions and obtain the coexisting phases. Using these solutions we then construct the tie lines, some examples are drawn in Fig. 2.2(d) as green dotted lines connecting green circles on the binodal. These steps are generic and can be used for any functional form of the free energy density and can easily be extended to higher dimensions.

2.2 Characteristics of phase diagrams

Phase diagrams reflect the phase behavior of the system as a function of different thermodynamic variables or molecular properties, e.g. temperature, composition, molecular size. For instance, in protein solutions it might be relevant to understand how changing the concentration of a binding partner of a phase separating protein can affect its tendency to phase separate. It might also be important to study how sudden changes in temperature affect the concentrations of the coexisting phases. All this information is encoded in the phase diagram. Developing an understanding of the different shapes that phase diagrams can take, may help us shed some light on the possible underlying mechanisms which generate the responses in protein solutions. In this section we discuss how certain molecular features, such as the molecular volume or the interaction between the components, affect the phase diagrams.

2.2.1 Binary mixtures

Phase diagrams for binary mixtures have been used for a long time to study the properties of polymers in solution. In recent years, due to the discovery of the so called membraneless
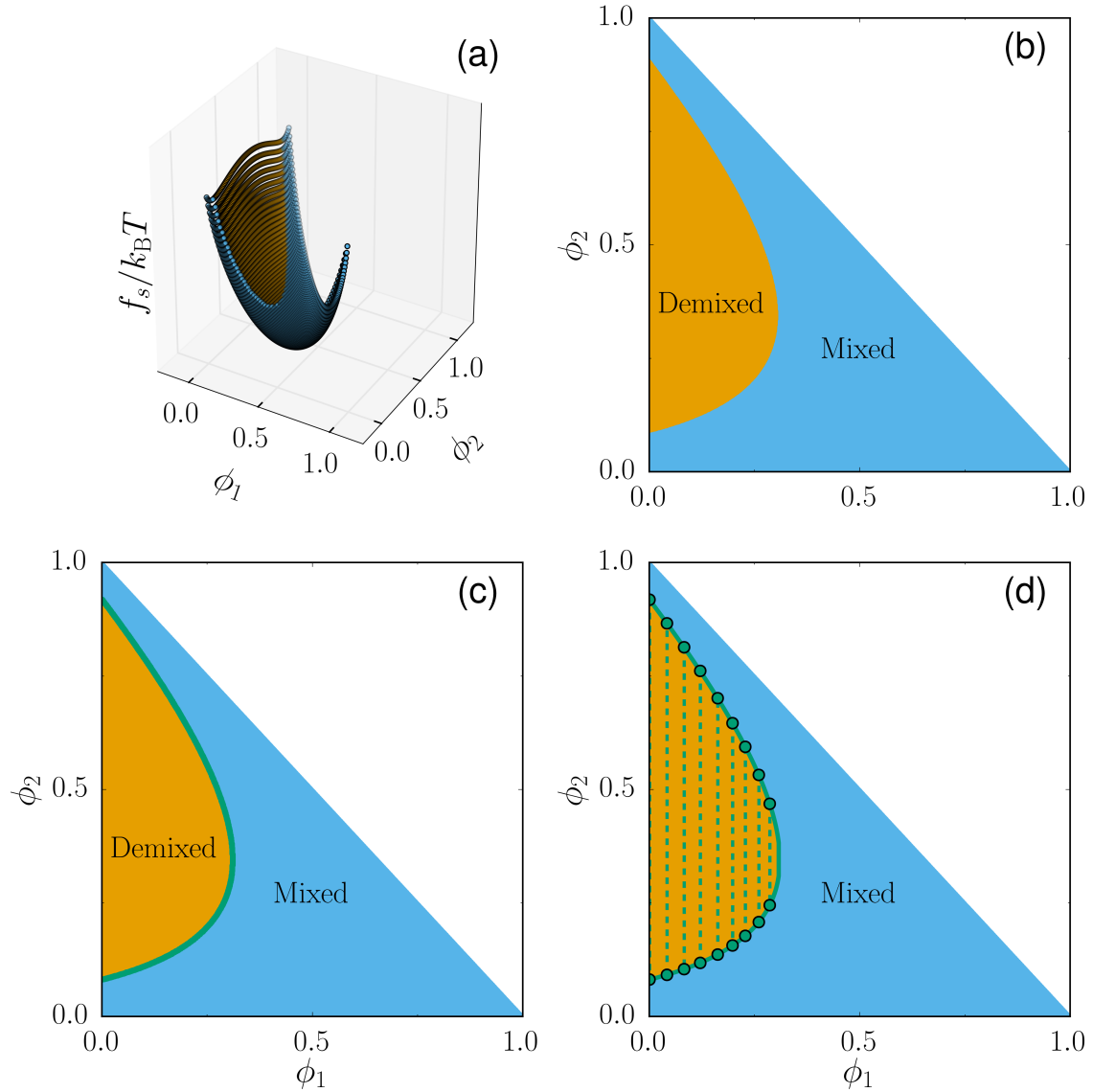
FIGURE 2.2: Constructing coexisting phases for a ternary mixture using a dimensionless free energy per site. (a) Dimensionless free energy per site $f_s/k_\mathrm{B}T$ as a function of the volume fractions ϕ_1 and ϕ_2. Blue points belong to the convex hull of $f_s/k_\mathrm{B}T$ whereas the orange points do not. (b) Projection of the mixed states and demixed states obtained from (a) into the $\phi_1 - \phi_2$ plane. (c) The binodal line (green line) is constructed by finding the boundary between the mixed states and demixed states of the system. (d) The tie lines (dotted green lines) connecting coexisting phases (green circles) are constructed by solving the coexistence conditions. The Flory-Huggins parameter used here is $\chi = 2.9$

organelles whose formation appears to be driven by liquid-liquid phase separation, understanding the properties of protein solutions has gained attention [23, 28, 84]. As explained in Chapter 1, membraneless organelles are formed by RNA and many different proteins which have intrinsically disordered regions. Some of these proteins have been shown to exhibit phase separation *in vitro* [23, 28, 30, 84]. In the absence of any extra component (like RNA) we consider these protein solutions to be simple binary mixtures. We now revisit some of the properties that could be observed in experiments performed with protein solutions.

Let us consider a binary mixture described by Eq. (2.1). For this system, the critical point is found to be by [65]:

$$\phi_c = \frac{\sqrt{\epsilon}}{1 + \sqrt{\epsilon}} \quad , \tag{2.3a}$$

$$T_c = \frac{\chi^H}{\frac{1}{2}\left(1 + \sqrt{\epsilon}\right)^2 - \chi^S} \quad . \tag{2.3b}$$

where ϕ_c is the critical volume fraction of the macromolecule (intrinsically disordered protein) and T_c is the critical temperature. Looking at the critical values (2.3), we see that the critical volume fraction decreases with increasing size of the macromolecules (decreasing ϵ), whereas the critical temperature increases with increasing size of the macromolecules. The critical volume fraction in this minimal model is independent of the interaction strength and provides the following qualitative picture for intrinsically disordered proteins: proteins which have large intrinsically disordered regions should have a low critical concentration ($n_c = \epsilon\phi_c/v_0$), therefore, phase separation is expected to happen at very low concentrations for large intrinsically disordered proteins.

We now study two different scenarios, one in which the Flory-Huggins interaction parameter χ is only of enthalpic nature (Fig. 2.3(a,b)) and one in which there is a significant entropic contribution (Fig. 2.3(c,d)). Using Eq. (2.3b), where we chose χ^H and χ^S by fixing the value of the critical temperature to be $T_c = 372.5$K for $\epsilon = 10^{-4}$ in both scenarios. We construct binodals as a function of the volume fraction and temperature for different values of the molecular volume ratio ϵ, the volume fractions which lie inside of the binodal are those which demix into two different phases, those who lie outside remain homogeneously mixed. In these phase diagrams, tie lines are horizontal lines connecting the phases which coexist at a given temperature. A general feature of these phase diagrams is that for a fixed value of χ^S and χ^H, the tendency to phase separate increases with the size of the macromolecules (decreases with ϵ). The asymmetry of the binodal stems from the large differences in molecular volumes, whereas for a regular mixture ($\epsilon = 1$) the binodal is always symmetric (green light in Fig. 2.3(a)). One interesting feature that emerges when there is a large entropic contribution to the interaction parameter, is that for small ϵ there is a significant temperature range where the phase diagram becomes narrow (dark green line

in Fig. 2.3(c)) as compared to a system in which the only contribution to the interaction parameter is enthalpic (dark green line in Fig. 2.3(a)). Since we want to stress the relation between the phase diagrams of binary mixtures and those that could be experimentally obtained in protein solutions, we show the same phase diagrams as a function of the molar concentration in Fig 2.3(b,d). There is evidence that some phase separating proteins have saturation concentrations between $(0.5 - 2)\mu$M [28], so if we seek to reproduce this behavior, we should study the regions of the binodal where the concentration of the low density branch is of the order of $1\,\mu$M. Let us define the partitioning coefficient $\mathcal{P} = \phi^{\mathrm{II}}/\phi^{\mathrm{I}}$, where ϕ^{II} and ϕ^{I}, are the volume fractions of the high and low density phases respectively. If we now analyze the partitioning coefficient for a low density phase corresponding to $1\,\mu$M for the different values of ϵ, we observe that the partitioning coefficient increases for decreasing values of ϵ, ranging from $\mathcal{P} \approx 10^8$, for $\epsilon = 1$, to $\mathcal{P} \approx 5 \times 10^2$, for $\epsilon = 10^{-4}$ (see Fig. 2.3(b,d)). The saturation concentration and the partitioning coefficient can then be used to estimate the molecular volume ratio ϵ and the interaction parameter χ of a given protein in solution. By studying the temperature response of protein phase separation, we may be able to learn which type of processes drive phase separation, i.e. whether they are mostly driven by direct attractive interactions between components or if the processes are driven by entropic effects.

In this section we have seen that addressing the phase diagrams of protein solutions as a binary mixture provides us with a qualitative picture of the phase behavior and gives us some insight into the basic mechanisms driving phase separation. We could say that if the coexisting phases do not vary drastically in response to temperature changes one may guess that there might be an entropic contribution which plays the main role in the system, whereas for strong variations in the corresponding coexisting phases for temperature changes one would expect that direct interactions play a significant role.

2.2.2 Phase diagram topologies of ternary mixtures

Ternary mixtures have a rich phase behavior. For a fixed temperature one can find a diversity of phase diagrams with different topologies even for molecules of the same size. In this section, using the numerical tools discussed in Sect. 2.1, we construct phase diagrams which include all the different topologies that can be found in a ternary mixture. The free energy per site for a regular ternary mixture reads

$$
\begin{aligned}
\frac{f_s}{k_{\mathrm{B}}T} =\ & \phi_1 \ln \phi_1 + \phi_2 \ln \phi_2 + (1 - \phi_1 - \phi_2)\ln(1 - \phi_1 - \phi_2) \\
& + \chi_{12}\phi_1\phi_2 + \chi_{10}\phi_1(1 - \phi_1 - \phi_2) + \chi_{20}\phi_2(1 - \phi_1 - \phi_2)\quad,
\end{aligned}
\tag{2.4}
$$

where χ_{ij} denotes the Flory-Huggins parameter interaction between i and j. We now explain the different topologies of the phase diagrams for this ternary mixture. One topology is given by a two-phase region which emerges from one of the axis which is delimited by one
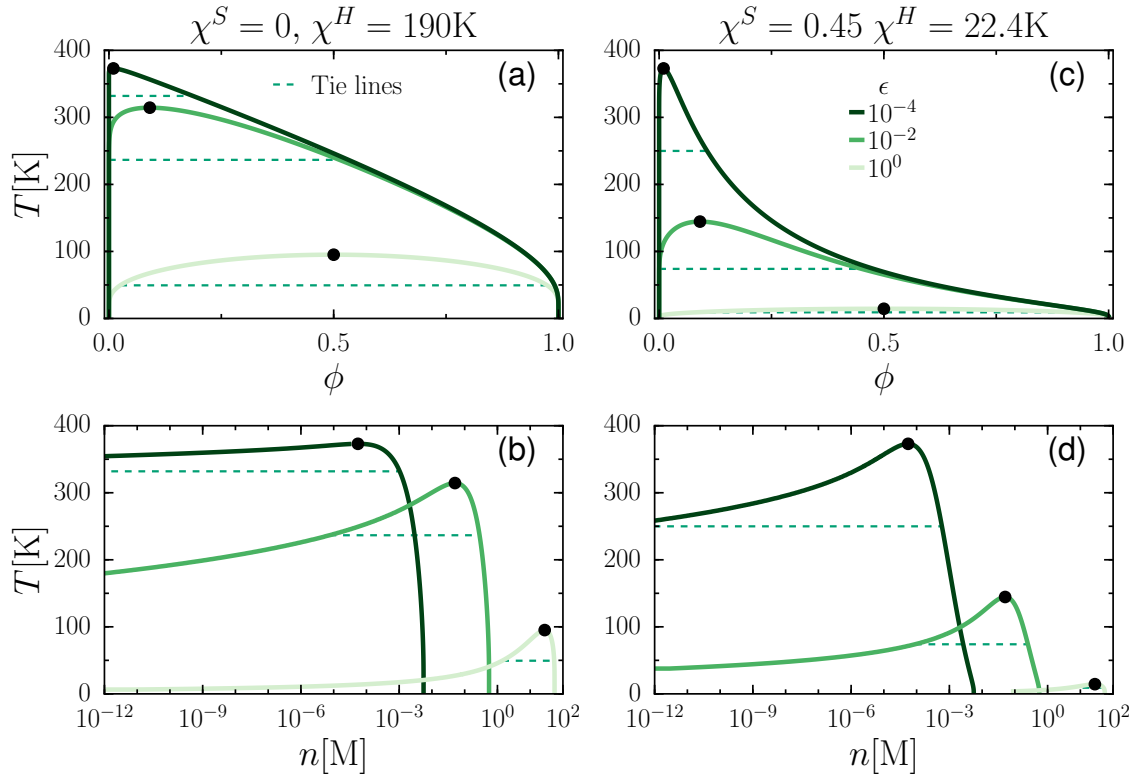
FIGURE 2.3: Phase diagrams of a binary mixture for varying molecular volumes ratio ϵ. (a) Binodals for a system with an interaction parameter which only has an enthalpic contribution. (b) Same binodals as in (a) but drawn as a function of molar concentration. (c) Binodals for a system with an interaction parameter which has a significant entropic. (d) Binodals from (c) plotted as a function of molar concentrations. Critical points are shown as black circles and tie lines are represented as dotted green lines connecting coexisting phases. Legends in (a) and (c) apply to all plots in this figure.

critical point (black point), see Fig. 2.4(a). This is obtained when only one of the interaction parameters is above the critical value for binary phase separation and the other two are below the critical threshold (for the ternary mixture described by Eq. (2.4) the critical threshold is $\chi_{ij}^c = 2$). If there are two interaction parameters which are above the critical treshold, depending on the third interaction parameter, there are two possible topologies (Fig. 2.4(b,c)). One topology is given by a two-phase region spanning from one volume fraction axis to the other volume fraction axis with no critical points (Fig. 2.4(b)). The second topology is given by two disconnected two-phase regions each one delimited by a critical point (black points in Fig. 2.4(c)). For systems where also the third interaction parameter is above the critical threshold, that is, all three parameters are above their critical values for binary phase separation, we may get three different topologies depending on the strength of the interactions. One of the possible topologies is given by two disconnected regions, one of them spanning from one axis to the other without any critical point and the other region starting at one axis and ending at a critical point (Fig. 2.4(c)). Another topology is given by three disconnected regions, all of them spanning from a volume fraction axis and each one of them delimited by a critical point (Fig. 2.4(d)). The last of these three topologies has no critical points, instead the two-phase regions which span from each of the three axis merge in a three-phase coexistence region (Fig. 2.4(e)). The last topology that we find is that of an isolated region of phase separation which ends at two critical points and is not connected to any of the volume fraction axes (Fig. 2.4(f)). This topology can be obtained for interaction parameters which are negative and below a certain threshold. The critical value for the free energy given by Eq. (2.4) for negative Flory-Huggins parameters is given by $\chi_{ij}^c = -8$, which can be calculated using the conditions given in Appendix B.

There is one feature that we would like to highlight regarding the construction of these phase diagrams. In order to construct a global phase diagram, sometimes it is necessary to know how many phases will there be in certain composition region (for example, where does a three phase region exists), using the convex hull construction as a first approximation to the binodal, this information is not needed. The approximated binodal obtained by means of the convex hull drastically reduces the search space of possible initial conditions for solving the coexistence conditions using any numerical root solver.

We gave this short exposition of possible topologies in ternary phase diagrams to illustrate how the complexity of diagrams grows with the number of components. Moreover, according to the Gibbs phase rule [64], the possibility of a higher number of coexisting phases also grows with the number of components. That is why in this ternary mixture one may observe a three phase region whereas in a binary mixture there is no such region. In order to connect this discussion with what is observed in a living cell, we need to remember that the cytosol is multicomponent in nature, giving rise to the possibility of the simultaneous emergence of several distinct phases. Interestingly, this seems to be the case
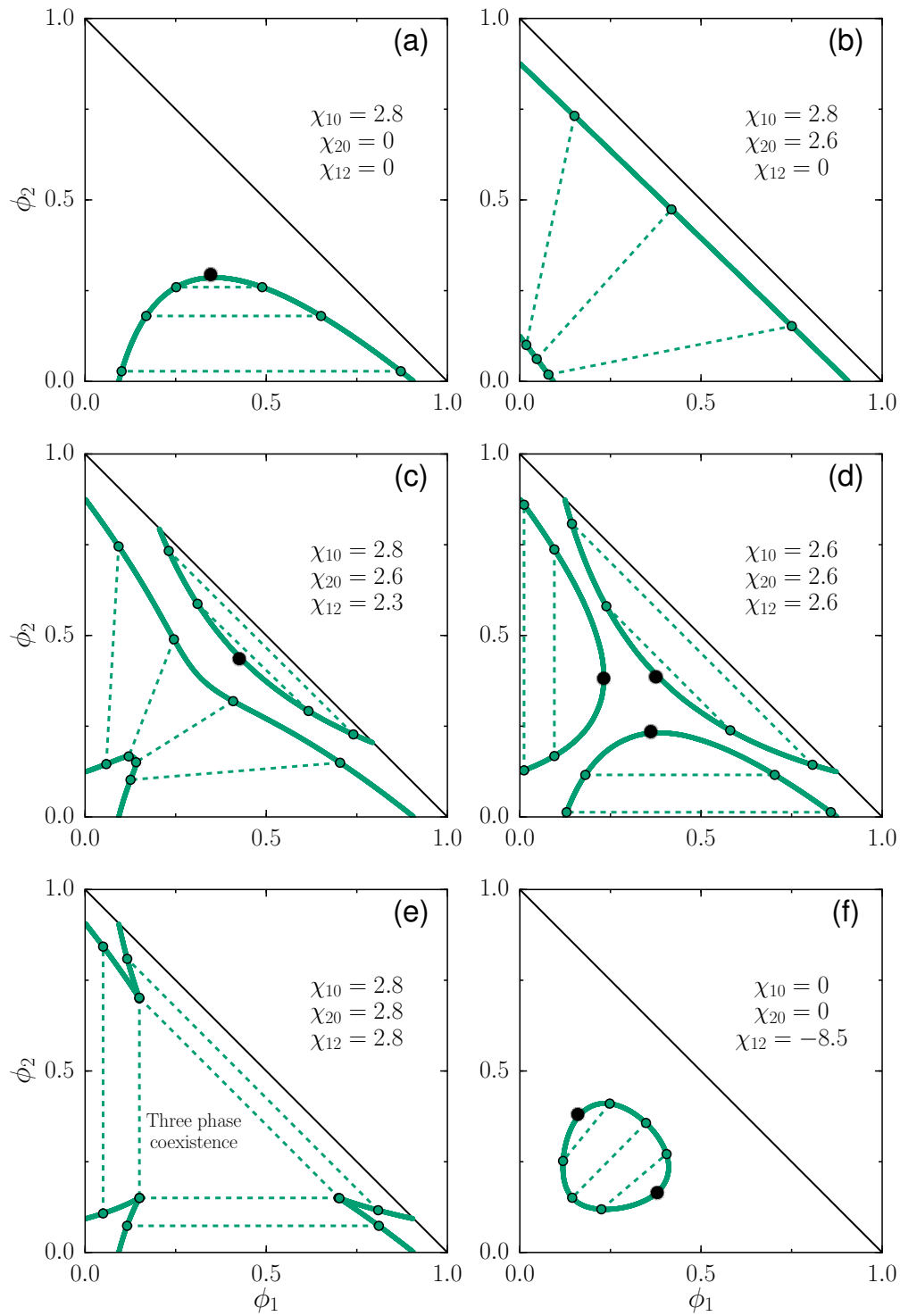
FIGURE 2.4: Phase diagrams showing generic topologies of the phase diagrams for ternary mixtures. Binodals (green solid lines), tie lines (green dotted lines) connecting coexisting phases (green points) and critical points (black points) are shown.

since there are different membraneless organelles present simultaneously in the cells [3] as well as evidence of multiphase coexistence within a single biomolecular condensate [18], this gives us a hint that a multicomponent description of the cytosol is strictly necessary to understand its full complexity.

2.3   Protein condensates *in vitro*

In the previous sections we discussed generic phase diagrams for binary and ternary mixtures. We now present possible phase diagrams describing protein phase separation *in vitro*. In particular we discuss what was observed regarding phase separation of the intrinsically disordered protein PGL-3, a key component of P granules. Experiments where done to assess the phase separation of the PGL-3 protein, with PGL-3 tagged with a fluorescent protein to keep track of its concentration [28]. It was shown by the authors that PGL-3 undergoes phase separation at a saturation concentration of $n_{\mathrm{p3}} \approx 2\,\mu\mathrm{M}$ under physiological conditions. The condensed phase was found to be approximately 50 times more concentrated. This might be an underestimation due to the non-linearity relation between the emitted fluorescence and the concentration of the fluorescent label in the concentrated phase. Recent developments have shown that the partitioning coefficient might be of the order of hundreds[1], but for the time being we use parameters close to the ones found in [28].

We first analyze two possible phase diagrams for the binary mixture of PGL-3 and solvent (in the experiments, the protein is immersed in a buffer). The free energy per site is given by Eq. 1.21 with a molecular volume ratio between solvent and PGL-3, $\epsilon_{\mathrm{p3}} = v_0/v_{\mathrm{p3}} = 1/15000$, and the Flory-Huggins parameter capturing the interactions between PGL-3 and the solvent is $\chi_{\mathrm{p3}} = 0.512$ at $T = 298\mathrm{K}$. These values were chosen to match the experimental observations of the saturation concentration of PGL-3, $n_{\mathrm{p3}} \approx 2\mu\mathrm{M}$ and the partitioning coefficient of PGL-3, $\mathcal{P}_{\mathrm{p3}} \approx 50$ [28]. The phase diagrams are constructed for two different cases with different functional forms of χ_{p3} as a function of temperature. The first is an interaction parameter which only has an enthalpic contribution. Using the value of $\chi_{\mathrm{p3}} = 0.512$ at $T = 298\mathrm{K}$ we estimate the enthalpic contribution to be $\chi_{\mathrm{p3}}^{H} = 298 \times (0.512)\mathrm{K}$. For the second case, to have a contrast in the phase diagrams between the previous enthalpic dominated interaction and an entropic dominated interaction, we fix $\chi_{\mathrm{p3}}^{S} = 0.505$ as the main contribution and use again the value of $\chi_{\mathrm{p3}} = 0.512$ at $T = 298\mathrm{K}$, we estimate the value of the enthalpic contribution as $\chi_{\mathrm{p3}}^{H} = 298 \times (0.512 - 0.505)$. Using these two different choices of parameters, we construct the corresponding binodals (Fig. 2.5 (a)). Even though the two binodals intersect at $T = 298\mathrm{K}$ by construction, we observe a marked difference between the two cases. Let us focus on a temperature range between 280K and 320K, which is a temperature range where we expect PGL-3 phase separation to be relevant in the *C. elegans*. For the entropic case, we observe a very weak

---

[1]Private communication with Dr. Patrick M. McCall

dependence of the composition of the coexisting phases on temperature changes (blue solid line in the inset of Fig. 2.5 (a)). For the enthalpic case instead, we observe big changes in the concentration of the diluted branch of the binodal as temperature varies (orange solid line in Fig. 2.5 (a)). This might be relevant for protein condensates, as drastic changes in concentration as a function of temperature may result in lack of robustness of their assembly.

We are not aware of any conclusive data regarding the temperature dependence of the concentrations of the coexisting phases in these protein solutions. However, the simple approach presented here would allow us to discriminate between cases in which entropic interactions dominate and cases in which enthalpic interactions dominate based only on the temperature dependence of the concentrations in the coexisting phases of protein solutions.

Let us now study how the system behaves if we consider another key protein of P granules, namely the PGL-1 protein. For this purpose, we analyze a ternary mixture composed of PGL-3 with volume fraction ϕ_{p_3}, PGL-1 with volume fraction ϕ_{p_1}, and solvent with volume fraction $(1 - \phi_{p_3} - \phi_{p_1})$. We describe the system with the following free energy per site

$$
\begin{aligned}
\frac{f_s}{k_B T} &= \epsilon_{p_3} \phi_{p_3} \ln \phi_{p_3} + \epsilon_{p_1} \phi_{p_1} \ln \phi_{p_1} + (1 - \phi_{p_3} - \phi_{p_1}) \ln(1 - \phi_{p_3} - \phi_{p_1}) \\
&+ (\chi_{p_3} \phi_{p_3} + \chi_{p_1} \phi_{p_1})(1 - \phi_{p_3} - \phi_{p_1}) \quad ,
\end{aligned}
\tag{2.5}
$$

where we introduced χ_{p_1} as the interaction between PGL-1 and solvent, and ϵ_{p_1} is the molecular volume ratio between the solvent and PGL-1. In oder to construct a phase diagram at fixed temperature $T = 298$K for this ternary mixture, we consider the experimental observation that PGL-1 does not exhibit phase separation in the μM concentration range in the absence of PGL-3 and choose χ_{p_1} in such a way that phase separation is not found in the μM range. In addition, we choose a value for $\epsilon_{p_1} = 1/17000$, based on the fact that PGL-1 has a longer aminoacid sequence than PGL-3 (771 amino acids for PGL-1 compared to 693 for PGL-3). We then construct the phase diagram corresponding to this choice of parameters (Fig. 2.5 (b)). We can appreciate from the phase diagram that once the concentration of PLG-3 crosses the threshold for binary phase separation ($n_{p_3} > 2\mu$M), adding PGL-1 does not change the concentration of PGL-3 in the coexisting phases if the concentration of PGL-1 is $n_{p_1} \lesssim 3\mu$M. The condensed phase has a PGL-1 concentration approximately 10 times larger than that of the concentration in the dilute phase.

In this section we have presented possible phase diagrams for the proteins PGL-3 and PGL-1 *in vitro*. This short presentation of phase diagrams is intended to give a quantitative idea of the parameters describing the phase behavior of protein solutions. The large asymmetry between molecular volumes of the proteins compared to the solvent (water in this case), provides an intuitive idea of why some proteins already phase-separate at very low concentrations.
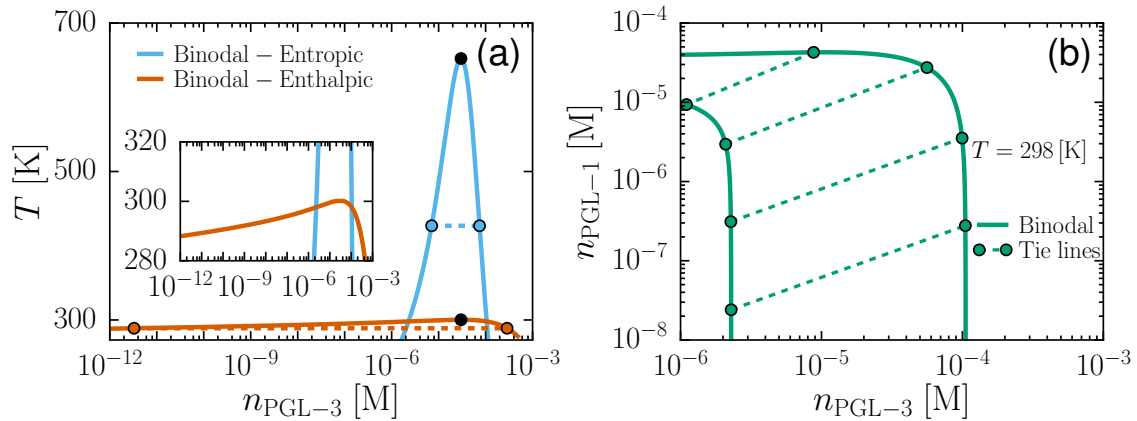
FIGURE 2.5: Possible phase diagrams for the phase separating protein PGL-3 in a binary mixture (a) and for the ternary mixture of PGL-3, PGL-1 and solvent (b). (a) Comparison between the binodals for systems with interactions which are dominantly enthalpic (blue line) or entropic (orange line). Tie lines are shown as dotted lines connecting coexisting phases which are blue and orange points for the entropic and enthalpic case respectively. The critical points are indicated by black points. The inset in (a) shows a zoom in into a temperature range (y axis of the inset) where PGL-3 phase separation is relevant in *C. elegans*. (b) Phase diagram of a ternary solution made of PGL-3, PGL-1 and solvent. The green solid line is the binodal, characteristic tie lines are shown as green dotted lines connecting two coexisting phases (green points). The parameters used are described in the main text.

2.4   Summary

In this chapter we showed how the convex hull construction can be used to determine the stable compositions of a multicomponent free energy. We presented a numerical method to construct phase diagrams by means of the convex hull alone or in combination with a numerical root solver to find coexisting phases. We then constructed examples of phase diagrams of binary mixtures where we highlighted the marked differences between phase diagrams where the interaction is either of entropic or enthalpic nature. For ternary mixtures we discussed all the possible phase diagram topologies. We concluded the chapter by showing possible phase diagrams for real protein solutions, namely, a binary mixture made of the protein PGL-3 and solvent, as well as a ternary mixture of proteins PGL-3 and PGL-1 immersed in a solvent. The main goal of this chapter was to establish the numerical methods that we use to study liquid phase separation in multicomponents systems and to discuss minimal binary and ternary models for protein solutions.

Chapter 3

Generic Theory of Phase Separation at Chemical Equilibrium

Multicomponent mixtures with chemical reactions exhibiting phase separation have been studied in the past [70–72, 76, 85, 86], mostly focused on particular models. Here we present a generic approach describing how to proceed in studying a system with many components undergoing several chemical reactions and phase separation simultaneously. Since there is a growing interest in having a deeper understanding on what are the effects of binding reactions in protein phase separation, developing such a thermodynamic description may be useful. Many if not all key components of protein condensates, bind to RNA, forming protein-RNA complexes which undergo phase separation [52]. It has also been shown *in vitro* that the presence of RNA lowers the saturation concentration of the PGL-3 protein [28]. Another example where studying phase separation in systems with chemical reactions could be useful is that of macromolecules undergoing phase separation depending on the pH of the solution where they are immersed. Since the charge state of a macromolecule is controlled by the pH of the solution, in order to describe this phenomenon, one could couple a model of a phase separating system together with chemical reactions describing changes in the charge state of the macromolecule. We devote Chapter 4 to address this question. In Chapter 1, we discussed a multicomponent mixture as a minimal model for studying the cytosol, since the interior of a cell is constantly undergoing chemical changes, we complement this approach by studying the influence of chemical reactions in mixtures of interacting components. In this case, we do not only consider short range interactions which can be captured by the Flory-Huggins parameter, but also highly directional interactions which allow for the formation of new chemical components or molecular complexes.

This chapter is organized as follows: we begin discussing the role of conservation laws in systems with chemical reactions. We then derive the chemical equilibrium conditions

based on these conservation laws. This then allow us to define the corresponding thermo-dynamic potentials at chemical equilibrium, where the conserved quantities are chosen as independent composition variables. Using these thermodynamic potentials we discuss the Maxwell construction for systems at chemical equilibrium. We conclude the chapter by giving a concrete example of an interacting system, and show that there might be more than one solution to the chemical equilibrium conditions, this allows the system to have coexisting phases which come from different branches of the equilibrium compositions. This is drastically different from the case of an ideal mixture, where it has been shown that there is a unique set of solutions for the chemical equilibrium conditions [87]. We show how to construct a full phase diagram as a function of temperature using the convex hull construction, which regardless of the shape of the free energy still gives the correct criterion in selecting the compositions where the free energy is a convex function, hence thermodynamically stable.

The thermodynamic framework developed here can be applied to any system for which the free energy of the system is known. There are similarities in our approach to study conservation laws in chemical reactions to previous work developed by R.A. Alberty [69]. In this chapter, we give our own account on how to define the chemical potentials of conserved components, how to derive the chemical equilibrium conditions as a consequence of the conservation laws and we go a step further and show how to construct the corresponding coexisting phases for systems which undergo simultaneous phase separation and chemical reactions.

3.1   Conservation laws in chemical reactions

Conservation laws in chemical reactions specify the collection of atoms, elements and charges that remain unchanged during chemical reaction events. Hereafter, we refer to these collection of matter as conserved components. For instance, if we consider the self-ionization of water given by

$$\mathrm{H_3O^+ + OH^- \rightleftharpoons 2\,H_2O} \quad , \tag{3.1}$$

the hydrogen and the oxygen atoms are independent conserved components, while we refer to the hydronium $\mathrm{H_3O^+}$, the hydroxide $\mathrm{OH^-}$ and the water molecule $\mathrm{H_2O}$ as chemical components.

In this section we show that if the stoichiometry matrix of the system is known, we can systematically construct a full set of independent conserved components to describe the system and importantly, also their corresponding particle numbers. For this purpose, let us consider a set of r reactions in a system with s chemical components. As discussed

in Chapter 1, these reactions are described in terms of the stoichiometry matrix ν by

$$\sum_{j=1}^{s} \nu_{ij} A_j \rightleftharpoons 0 \quad , \quad \text{for } i \in \{1, \dots, r\} \quad , \tag{3.2}$$

where ν_{ij} are matrix elements of ν giving the number of molecules of the j-th chemical component that are transformed in the i-th reaction, while A_j denotes the chemical symbol for the j-th chemical component. The stoichiometry matrix ν is thus an $r \times s$ matrix, where the number of rows and columns are defined by the number of independent reactions and the number of chemical components, respectively. Here, reactions 3.2 provide r constraints that lead to a reduced number $c = s - r$ of independent components in the system. We now argue that if the stoichiometry matrix ν is known, the c independent components of the system can be chosen as conserved components [69]. Let us define the conservation laws by

$$\sum_{j=1}^{s} \nu_{ij} C_j^{(k)} = 0 \quad , \quad \text{for } i \in \{1, \dots, r\} \quad , \tag{3.3}$$

where $C_j^{(k)}$ defines the corresponding amount of the k-th conserved component in the j-th chemical component[1]. The vector $\vec{C}^{(k)} = (C_1^{(k)}, \dots, C_j^{(k)} \dots, C_s^{(k)})$ defines the composition of the k-th conserved component in terms of the s original chemical components. By solving Eqs. (3.3), it can be shown that there are precisely c linearly independent solutions given by a set of vectors $\{\vec{C}^{(1)}, \vec{C}^{(2)}, \dots, \vec{C}^{(c)}\}$ [88][2]. Placing the elements of this set as columns of a matrix, defines what is known as the conservation matrix $\mathcal{C}$ [69]. The elements of this matrix obey

$$\sum_{j=1}^{s} \nu_{ij} \mathcal{C}_{jl} = 0 \qquad \text{for } i \in \{1, \dots, r\} \text{ and } l \in \{1, \dots, c\} \quad , \tag{3.4}$$

which is basically a restatement of the conservation laws. The matrix form of the conservation laws nicely conveys the message of having c independent conserved components in each of the r reactions. So far, we have shown that given a stoichiometry matrix, one can find c independent components which satisfy the conservation laws given by Eq. (3.3). Let us now discuss the particle numbers associated to the independent conserved components, if N_i denotes the particle number of the i-th chemical component and N_j^c denotes the particle number of the j-th conserved component, which hereafter we call the j-th conserved

---

[1]For example, if $A_j = H_2O$ and the k-th conserved component refers to the hydrogen atoms, $\vec{C}_j^{(k)} = 2$.

[2]A matrix with r rows and s columns denoted by A, with $r < s$, has a rank$(A) = r$. The null space of A is defined as the set of vectors $\vec{x}$, which satisfy $A\vec{x} = 0$. The rank of the null space is given by $c = s - r$, which means that there are c linear independent vectors satisfying $A\vec{v} = 0$, and they span the full null space [88]

particle numbers. We can relate them by means of the conservation matrix as follows

$$\sum_{i=1}^{s} N_i \mathcal{C}_{ij} = N_j^c \qquad \text{for } i \in \{1, \ldots, r\} \quad , \tag{3.5}$$

which states the global conservation laws in a chemically reacting system. Thus, we have seen that the conservation matrix $\mathcal{C}_{ij}$ encodes all the information needed to define a set of independent conserved components and their corresponding particle numbers. For completeness and clarity, an example discussing the construction of the conserved components and their corresponding particle numbers is given in Appendix D.

In this section we showed in general, that the c independent components in a system composed of s different chemical components undergoing r chemical reactions can always be chosen to be conserved components. The relevance of this result will become clearer in the next sections.

3.2 Chemical equilibrium conditions

We discuss the chemical equilibrium conditions in the $(T, P, \{N_i\})$ ensemble, where T is the temperature, P is the pressure and $\{N_i\}$ is the set of all different chemical components of the system. The corresponding thermodynamic potential for this ensemble is the Gibbs free energy $G(T, P, \{N_i\})$. The differential dG is given by [64]

$$dG = -SdT + VdP + \sum_{i=1}^{s} \mu_i dN_i \quad . \tag{3.6}$$

If the system undergoes r independent reactions and the total number of chemical components is s, in the previous section we showed that we can construct c independent conserved components. We name the c particle numbers associated to the independent conserved components by, $\{N_1^c, N_2^c, \ldots, N_c^c\}$. We now express c particle numbers of the chemical components as a function of the c conserved particle numbers and of the $r = s - c$ remaining particle numbers of the chemical components. Without any loss of generality, we choose to eliminate the first c particle numbers of the chemical components [3], i.e. we express the particle numbers $N_i = N_i(N_1^c, \ldots, N_c^c, N_{c+1}, \ldots, N_s)$, for $i = \{1, \ldots, c\}$. We then express the differential of the Gibbs free energy Eq. (3.6) in terms of the new variables as

$$dG = -SdT + VdP + \sum_{j=1}^{c} \left(\sum_{i=1}^{c} \mu_i \frac{\partial N_i}{\partial N_j^c} \right) dN_j^c + \sum_{j=c+1}^{s} \left(\mu_j + \sum_{i=1}^{c} \mu_i \frac{\partial N_i}{\partial N_j} \right) dN_j \quad . \tag{3.7}$$

---

[3]If one of the chemical species is a conserved component the variable transformation would lead to a trivial renaming of its corresponding particle number

Thermodynamic equilibrium demands $dG = 0$. Since T, P and the particle numbers of the conserved components are fixed, the first three terms on the right hand side of Eq. (3.7) vanish. This then leads to the following condition describing the chemical equilibrium conditions,

$$\mu_j + \sum_{i=1}^{c} \mu_i \frac{\partial N_i}{\partial N_j} = 0 \quad , \quad \text{for } j = \{c+1, \dots, s\} \quad , \tag{3.8}$$

This is not just a different way of deriving the well known chemical equilibrium conditions, it is one way to define new thermodynamic conjugate variables for a system that is at chemical equilibrium[69, 76, 77]. We now define the differential of the Gibbs free energy at chemical equilibrium as

$$d\bar{G} = -SdT + VdP + \sum_{j=i}^{c} \mu_i^c dN_i^c \quad , \tag{3.9}$$

where we introduced the Gibbs free energy at chemical equilibrium $\bar{G}$ and the chemical potentials of the conserved components $\mu_j^c = \partial \bar{G} / \partial N_i^c \big|_{T,P,\{N_{j\neq i}^c\}}$, which are the thermodynamic conjugate variables to the conserved particle numbers, defined by

$$\mu_j^c = \sum_{i=1}^{c} \mu_i \frac{\partial N_i}{\partial N_j^c} \quad . \tag{3.10}$$

In this section, we have showed that a system in the $(T, P, \{N_i\})$ ensemble at chemical equilibrium is fully described, in addition to the temperature and pressure, by the c particle numbers corresponding independent conserved components. Moreover, if we want to use a different thermodynamic description with fixed values of a given chemical potential (or a combination of them), we would need to use the newly defined chemical potentials in Eq. (3.10) when performing a Legendre transform to exchange thermodynamic conjugate variables. This allow us to define other thermodynamic potentials that also describe the system at chemical equilibrium. We will make use of this when discussing the phase behavior of a system in the ensemble where the pH is fixed in Chapter 4. A thorough study of Legendre transforms in systems with chemical reactions can be found in the literature [76].

3.3 Maxwell construction at chemical equilibrium

We are interested in studying the phase separation behavior of systems that undergo chemical reactions. In order to do so, we have shown that the description of a system at chemical equilibrium can be done by identifying a set of independent conserved components and using them as independent composition variables. We are now interested in describing what are the conditions to find coexisting phases at chemical equilibrium. If we describe the

system in the ensemble of fixed temperature and pressure with the corresponding thermodynamic potential $\bar{G}$, the coexistence conditions are simply given by

$$\mu_i^c(T, P, \{N_i^{cI}\}) = \mu_i^c(T, P, \{N_i^{cII}\}) \tag{3.11}$$

where the superscripts I and II refer to two coexisting phases. Since the system is considered to be already at chemical equilibrium, the only chemical potentials that need to be equilibrated for phases to coexist are those of the conserved components.

Let us now discuss the case for a system described in the fixed volume ensemble. We must first remember that in an incompressible system the volume is given by $V = \sum_{i=1}^{s} v_i N_i$, where the molecular volumes v_i are independent of composition and pressure. The volume of incompressible systems does not change when undergoing chemical reactions (this is just an approximation because volume is not conserved in general [89]). If the volume is fixed, then we can express it as a linear combination of the conserved components, $V = \sum_{i=1}^{c} v_i^c N_i^c$, where we introduced the molecular volumes of the conserved components v_i^c. Taking the differential of the volume dV gives

$$dV = \sum_{i=1}^{c} v_i^c dN_i^c \quad . \tag{3.12}$$

We may now use Eq. (3.12) to eliminate one of the conserved components from the description. The corresponding free energy in the fixed volume ensemble is defined by $F = G + PV$, we follow the same convention at chemical equilibrium and define the corresponding free energy at chemical equilibrium as: $\bar{F} = \bar{G} + PV$. In order to define independent thermodynamic variables, we calculate the differential of the free energy $d\bar{F}$. To do so, we eliminate N_k^c by using $N_k^c = (V - \sum_{i \neq k} v_i^c N_i^c)/v_k^c$. This leads to

$$d\bar{F} = -SdT + \left(P - \frac{\mu_k^c}{v_k^c}\right)dV + \sum_{i \neq k}\left(\mu_i^c - \frac{v_i^c}{v_k^c}\mu_k^c\right)dN_i^c \quad . \tag{3.13}$$

We see that Eq. (3.13) is basically the same as Eq. (1.4). The only difference stems from the fact that in Eq. (3.13), we eliminated an arbitrary conserved component that is not necessarily the solvent. In systems where the solvent is a conserved component, we would choose to eliminate the solvent and the thermodynamic conjugate variables would be the osmotic pressure Π (Eq. (1.5)) and the exchange chemical potentials $\bar{\mu}$ (Eq. (1.6)) of each conserved component.

Let us now define the free energy density at chemical equilibrium as $\bar{f} = \bar{F}/V$, where $\bar{f} = \bar{f}(T, \{n_{i \neq k}^c\})$. The Maxwell construction for $\bar{f}$ can be done in the same way as described in Chapter 1, for this reason we restrict ourselves to showing the coexistence

conditions explicitly for this case, which are given by

$$\frac{\partial \bar{f}^{\text{I}}}{\partial n_i^{c\text{I}}} = \frac{\partial \bar{f}^{\text{II}}}{\partial n_i^{c\text{II}}} \quad \text{for } i \neq k \quad , \tag{3.14}$$

$$\bar{f}^{\text{I}} - \sum_{i \neq k} \frac{\partial \bar{f}^{\text{I}}}{\partial n_i^{c\text{I}}} = \bar{f}^{\text{II}} - \sum_{i \neq k} \frac{\partial \bar{f}^{\text{II}}}{\partial n_i^{c\text{II}}} \tag{3.15}$$

where $\bar{f}^{\text{I}}$ and $\bar{f}^{\text{II}}$ denote the free energy at chemical equilibrium in phases I and II respectively, and $n_i^{c\text{I}}$ and $n_i^{c\text{II}}$, the concentration of the conserved components in each of the two phases.

In the following, we describe a minimal model consisting of three components and one chemical reaction. We will explicitly show how using a conserved quantity one can analyze the phase behavior of the system at chemical equilibrium using the same approach as that for a mixture without chemical reactions. The only requirement is to first solve the chemical equilibrium conditions, express the solutions in terms of the conserved component and evaluate the free energy in these states.

3.4   Minimal model for heating induced phase separation

Here, we revisit the model introduced in Chapter 1 to discuss phase separation at chemical equilibrium. However, we discuss it in light of what we have derived so far in this chapter. Recapitulating, we propose a ternary model with only one chemical reaction. The system is composed of a solvent with particle number N_0 and a molecule that exists in two states, one where the molecule does not have any tendency to phase separate and another where the molecule can interact with the rest of the molecules in the same state. These two states are denoted by A and A', respectively. They transform obeying the following reaction

$$A \rightleftharpoons A' \quad . \tag{3.16}$$

In this system, it is easy to identify an independent set of conserved particle numbers. The solvent is not involved in the chemical reaction, thus is conserved, as well as the sum of the particle numbers in the state A and state A', given by $N_{A_T} = N_A + N_{A'}$, where N_A stands for the particle number of molecules in state A and $N_{A'}$ denotes the particle number of molecules in state A'. We describe this system at temperature T and pressure P using the following Gibbs free energy

$$
\begin{aligned}
G &= k_{\text{B}}T \left(N_A \ln \phi_A + N_{A'} \ln \phi_{A'} + N_0 \ln \phi_0 + \frac{\chi}{v_0} v_A N_{A'} \phi_{A'} \right) \\
&\quad + w_A N_A + w_{A'} N_{A'} + w_0 N_0 + P(v_A N_A + v_{A'} N_{A'}) + v_0 N_0) \quad , \tag{3.17}
\end{aligned}
$$

where χ is the interaction strength between molecules in state A', w_i is the free energy contribution due to internal degrees of freedom of each of the chemical components, v_i denotes the molecular volume and ϕ_i the volume fraction. For simplicity, we only consider interactions between molecules in the state A'. We further consider that the volume of the molecules do not change during chemical reactions, thus $v_A = v_{A'}$.

We now follow the steps described in the previous sections. We first eliminate one of the particle numbers using $N_A = N_{A_T} - N_{A'}$ and substitute it in the differential of the free energy dG, leading to

$$dG = -SdT + VdP + \mu_A dN_{A_T} + (\mu_{A'} - \mu_A)dN_{A'} \quad , \qquad (3.18)$$

where we can readily identify μ_A as the conjugate chemical potential to the conserved particle number N_{A_T}. Thermodynamic equilibrium implies the chemical equilibrium condition $\mu_A = \mu_{A'}$, using Eq. (3.17) the chemical equilibrium condition reads

$$\phi_A = \phi_{A'} \exp\left(\frac{w_{A'} - w_A}{k_B T \epsilon_A} + \frac{2\chi\phi_{A'}}{\epsilon_A}\right) \quad , \qquad (3.19)$$

where we introduced the molecular volume ratio $\epsilon_A = v_0/v_A$.

For an ideal system ($\chi = 0$) the chemical equilibrium condition Eq. (3.19) has only one possible chemical equilibrium. However, if interactions are present ($\chi \neq 0$), there might be more than one solution. We show a comparison of the chemical equilibrium condition Eq. (3.19) for an ideal and a nonideal case in Fig. 3.1(a), the dotted line corresponds to the ideal case and the solid line corresponds to the nonideal case, (the parameters used for this example can be found in the caption). The multiple solutions to the chemical equilibrium condition can be better observed by analyzing ϕ_A or $\phi_{A'}$ as a function of the total molecular volume fraction ϕ_{A_T}. We clearly see in Fig. 3.1(b,c) that they are not uniquely valued. The different chemical equilibrium branches are shown with different colors in Fig. 3.1(a,b,c).

Let us now consider the Gibbs free energy at chemical equilibrium, $\bar{G}$, which is obtained by evaluating Eq. (3.17) as a function of the solutions obtained from Eq. (3.19). The free energy $\bar{G}$ depends on N_{A_T}, N_0, T and P. Expressing the volume of the system in terms of the conserved components as $V = v_A N_{A_T} + v_0 N_0$ and changing our description to the fixed volume ensemble we can eliminate the solvent from the description. We discuss the phase behavior in terms of the free energy per site at chemical equilibrium, $\bar{f}_s = v_0(\bar{G}/V - P)$, which is given by

$$\begin{aligned}\bar{f}_s &= k_B T\left(\epsilon_A \phi_A \ln\phi_A + \epsilon_A \phi_{A'} \ln\phi_{A'} + (1 - \phi_A - \phi_{A'})\ln(1 - \phi_A - \phi_{A'}) + \chi\phi_{A'}^2\right) \\ &+ w_A \phi_A + w_{A'}\phi_{A'} + w_0(1 - \phi_A - \phi_{A'}) \quad , \end{aligned} \qquad (3.20)$$

where $\phi_A = \phi_A(\phi_{A_T}, T)$ and $\phi_{A'} = \phi_{A'}(\phi_{A_T}, T)$ are obtained from the chemical equilibrium condition Eq. (3.19). The free energy density per site $\bar{f}_s$ is shown in Fig. 3.1(d), where we see three distinct branches of the free energy coming from the three chemical branches. Interestingly, the ternary system analysis reduces to that of a binary mixture whose free energy has different branches. Nevertheless, the convex hull construction can be used to identify the coexisting phases.

As an illustration, we construct a full phase diagram as a function of the total molecular volume fraction ϕ_{A_T} and a scaled temperature $\widetilde{T} = T/T_0$, with T_0 being an arbitrary reference temperature with the rest of the parameters fixed, see caption in Fig. 3.2 for a complete list of the parameters used. The solutions to the chemical equilibrium conditions are shown as a function of $\widetilde{T}$ in Fig. 3.2(a,b). We see that for low temperatures (dark blue line), there are three different branches. Increasing the temperature leads to a unique solution to the chemical equilibrium conditions as a function of the total molecular volume fraction ϕ_{A_T}. We then do the Maxwell construction for the free energy per site for different temperatures, see Fig. 3.2(c). For high temperatures (dark red line), we see that there is no phase separation because the free energy is convex for all values of ϕ_{A_T}. Decreasing the temperature leads to the appearance of a region of phase separation, where the coexisting phases are denoted by green points, these points are joined by a common tangent and all compositions within the region spanned by the common tangent will demix. Decreasing further the temperature leads to the appearance of a small loop (see the line corresponding to $\widetilde{T} = 3.3$). For this case, the coexisting phases belong to different chemical equilibrium branches. The loops associated to the different branches of the free energy move further to the right for decreasing temperatures, until the free energy branches stop crossing for $\widetilde{T} < 2$ (not shown) and phase separation is no longer observed since the chemical branch that remains is fully convex. This results from the parameters choice $w_A < w_{A'}$, which favors the state A of the molecule at low temperatures, whereas for higher temperatures A' is favored. The phase behavior corresponding to the free energy is shown in Fig. 3.2(d), where tie lines are shown as dotted green lines connecting coexisting phases (green points) and the binodal is drawn as a solid green line. This system has a region of phase separation bounded by an upper temperature for which there is a critical point (black point) and a lower temperature, where the binodal connects with the $\phi_{A_T} = 1$ in a first order transition point where the values of ϕ_A and $\phi_{A'}$ are discontinuous. One interesting feature of this system that arises naturally in chemically reacting systems is reentrancy [71, 73, 85], that is, a system that is mixed for low temperatures, demixes if the temperature is increased, and then becomes mixed again if the temperature is raised even further. For example, hydrogen bonding has been shown to lead to interesting phase-separating behavior of macromolecules that are well solvated at low temperatures and then release all the hydrogen bonds if the temperature is increased, leading to phase separation due to hydrophobic interactions, if the temperature is raised even further the mixing entropy dominates the behavior of the
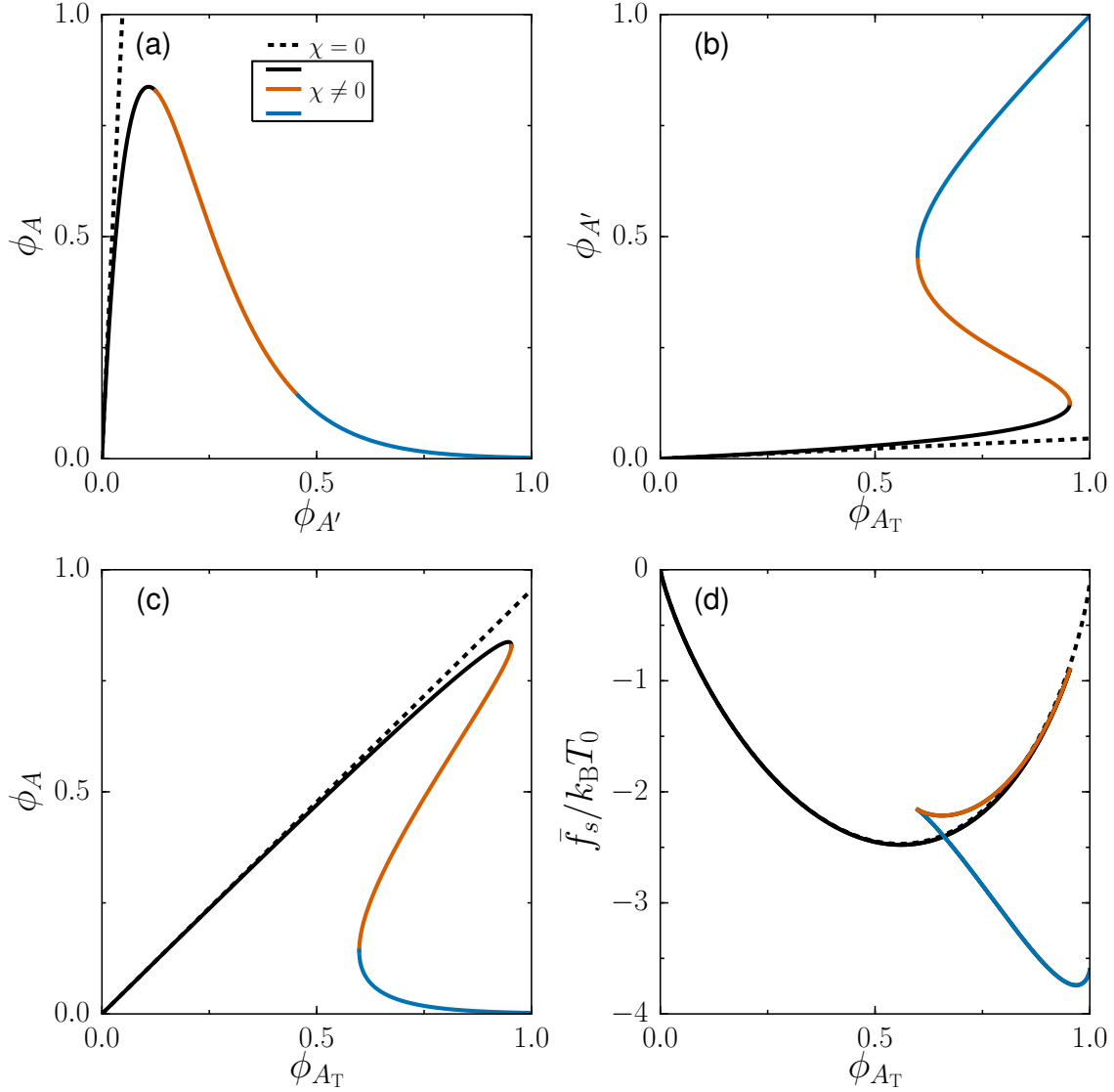
FIGURE 3.1: Comparison between an ideal and a nonideal mixture at chemical equilibrium. (a) Law of mass action given by Eq. (3.19) for an ideal system (dotted line) and for a nonideal system (solid line colored in black, orange and blue). (b) Volume fraction of the molecule in state A as a function of the total molecular volume fraction ϕ_{A_T}. The ideal case has a unique solution (dotted line) whereas the nonideal case, has three different chemical equilibrium branches, these are shown as black, orange and blue solid lines. (c) Volume fraction of the molecule in state A' as a function of the total molecular volume fraction ϕ_{A_T}, with the same description as in (b). (d) Free energy density at chemical equilibrium per site as a function of the total molecular volume fraction ϕ_{A_T}. Three distinct branches of the free energy density are shown, corresponding to the three chemical equilibrium branches. We observe that for the ideal case, the free energy remains convex (dotted line), in contrast, for the nonideal case, there is a region between two of the chemical branches (black and blue solid lines) which is clearly non-convex. The interaction parameter is given by $\chi = \chi^S + \chi^H/T$. Parameters: $\epsilon_A = 0.5$, $\chi^S = 1$, $\chi^H = 6T_0$, $\widetilde{T} = 4.6$ and $w_{A'} - w_A = 7k_\mathrm{B}T_0$. Legends apply to all plots.

system leading again to a well-mixed system [73, 90].

This model is useful to illustrate the concepts discussed in this chapter, and shows that describing the system in terms of its conserved components allow us to interpret and predict the phase behavior of the system in a rather straightforward manner by using the methods developed in Chapters 1 and 2.

3.5   Summary

In this chapter, we introduced a generic thermodynamic framework to study multicomponent systems in which their components undergo chemical reactions. We first showed how to construct a set of conserved components if the stoichiometry matrix of the system is known. We continued by showing that this set of conserved components fully describes the system at chemical equilibrium. Using the conserved particle numbers as independent variables in the free energy at chemical equilibrium, we derived an expression for their corresponding conjugate chemical potentials. We discussed the Maxwell construction in systems at chemical equilibrium and showed that the only chemical potentials that are left to equilibrate are those corresponding to the conserved components. In the case of the ensemble of fixed volume, we saw that if we use the conserved particle numbers as independent variables, the Maxwell construction holds for the corresponding free energy density at chemical equilibrium in the same way as in the case without reactions. We concluded the chapter by presenting a minimal model in which we showed explicitly how to use a conservation law and chemical equilibrium to construct a full phase diagram as a function of temperature. In brief, we showed that if one uses conserved particle numbers as independent variables in a system at chemical equilibrium, the numerical techniques from Chapter 2 can be applied directly to the system in question.

We would like to highlight that the standard procedure to construct coexisting phases in systems with chemical reactions usually starts from constructing the phase diagrams associated to the system in the absence of chemical reactions. Then what follows is to solve the corresponding chemical equilibrium conditions for each reaction. Finally, in order to find coexisting phases one needs to find the intersections between the coexistence manifold and the chemical equilibrium manifold [70, 71, 85]. We instead, presented here a different method, we first solve the corresponding chemical equilibrium condition for each reaction and then evaluate the free energy on these solutions (using the conserved particle numbers as independent variables). We then find the corresponding coexisting phases by the means exposed in the Chapter 2. Although both methods result in the same phase diagrams, there is a high complexity associated to the former. Finding the phase behavior for the system without considering the chemical reactions may be difficult due to its high dimension in composition space, which would be equivalent to the number of different chemical components s. In contrast, in our approach, we only need to analyze the reduced
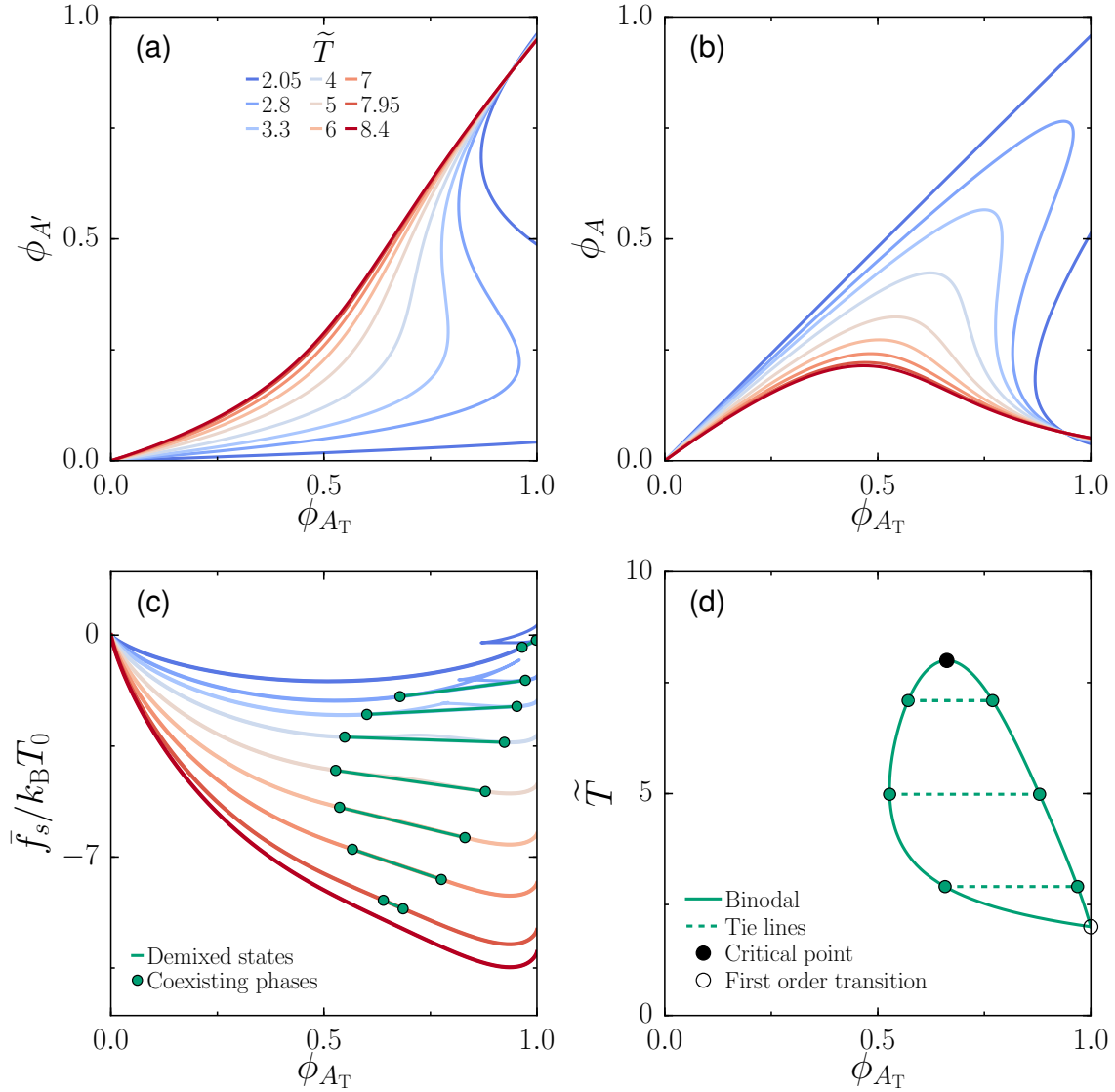
FIGURE 3.2: Phase diagram construction for a chemically reacting system. (a),(b) Volume fractions at chemical equilibrium for different temperatures as a function of ϕ_{A_T}, (a) ϕ'_A and (b) ϕ_A. Colored lines indicate the temperature values ranging from lower (dark blue) to higher (dark red). (c) Dimensionless free energy $\bar{f}_s/k_B T_0$ for different temperatures (same color code as in (a) and (b)). The common tangent (green lines) connecting coexisting phases (green points) are constructed by means of the convex hull. (d) Reentrant phase diagram showing the region of phase separation delimited by the binodal (green solid line). The binodal connects to the $\phi_{A_T} = 1$ line in a first order transition point (open circle). The tie lines are horizontal lines (dotted green lines) which connect the coexisting phases (green points). The binodal ends in the upper part at a critical point (black point). The interaction parameter is given by $\chi = \chi^S + \chi^H/T$. Parameters: $\epsilon_A = 1$, $w_{A'} - w_A = 7k_B T_0$, $\chi^S = 1.5$ and $\chi^H = 4T_0$. Legends in panel (a) also apply to panels (b) and (c).

dimension space of the conserved components, which is given by the number of independent conserved components c, which becomes smaller as the number of independent reactions r grows larger.

To conclude, developing a generic thermodynamic framework at chemical equilibrium opens up the possibility to assess the effect of chemical reactions on systems exhibiting phase separation in a systematic manner. At the same time, making use of the conservation laws allow us to understand which are the conjugate thermodynamic variables at chemical equilibrium. This turns out to be crucial in order to perform Legendre transforms in systems with chemical reactions.

Chapter 4

Liquid Phase Separation Controlled by the Acidity of the Environment

The acidity, which is quantified by the pH, is an important determinant of the cells viability. Because of this, cells stringently control its cytosolic pH by different mechanisms [2]. One of the reasons of this control is the fact that cytosolic proteins perform various physiological processes depending on their charge state, which is dictated by the cytosolic pH [91]. Additionally, the energetics of the cell are based on proton pumps which are tuned to respond to pH gradients across the cell membrane [1, 92]. An interesting example where pH changes have a strong effect, is the transition from a dynamic to an arrested state of the cytoplasm of yeast cells during nutrient depletion [16] introduced in Chapter 1. This transition is reversible and provides a protective mechanism that helps cells to survive periods of nutrient depletion until conditions improve and nutrients become available again. The biophysical mechanisms responsible for this transition have been linked to changes in the cytosolic pH. The effects of pH on yeast cells can be described as follows. During nutrient depletion, a yeast cell does not have sufficient resources to supply proton pumps that are responsible for regulating the intracellular pH. As a result, in an acidic environment the pH of the cytoplasm of cells drops and some proteins become insoluble. If the resulting protein condensates occupy a large volume the mechanical properties of the cytoplasm can transition from a fluid to an arrested state. This observation suggests that the reduction of pH triggers phase separation of proteins from solution. Interestingly, in this case, phase separation is triggered as the pH of the solutions moves closer to the isoelectric points of many constituent proteins. This raises the question why pH changes, and in particular, pH values in the vicinity of the isoelectric point promote phase separation. More generally we want to understand how the formation of protein condensates can be regulated by changes in pH.
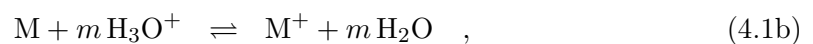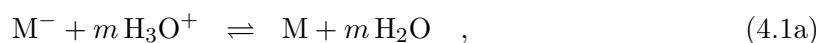
To address this question we present a generic thermodynamic framework to study the

influence of pH on liquid-liquid phase separation. The key idea is to couple a system capable of undergoing phase separation with a set of chemical reactions corresponding to the protonation/deprotonation of water components and macromolecules such as proteins. We consider two types of interactions, namely attractive interactions between oppositely charged macromolecules in the presence of counterions and salt, and attractive interactions among neutral macromolecules that could be mediated for example by Van der Waals or hydrophobic interactions. Using conservation laws and chemical equilibrium conditions [76], we construct an effective thermodynamic potential describing a system with pH as a thermodynamic variable. We then use this thermodynamic potential to determine the phase behaviour of the system as a function of the molecular properties and pH. We find coexisting phases of different compositions of charged and uncharged macromolecules. We show that the compositions of the coexisting phases can be controlled by changing pH.

This chapter is organised as follows. We first introduce a set of chemical reactions in which the charge state of a macromolecule is fixed by the pH of the system. We continue by defining the pH and showing its relation to the previously introduced chemical reactions. We then present the thermodynamics of multicomponent mixtures and discuss the parameter choices for our study. Afterwards, we study the thermodynamic equilibrium for a system with fixed pH, whereby using conservation laws, we identify the thermodynamic conjugate variables of the system which we use to construct the corresponding thermodynamic potential for fixed pH. In order to discuss chemical and phase equilibrium in our model, we introduce new composition variables and thermodynamic fields. We then analyze the phase behaviour of the system, first at the isoelectric point and then as pH deviates from it. We briefly discuss the consequences of considering electrical neutrality in the system and show that this leads to a natural difference of pH and electric potential across coexisting phases. Finally, we conclude with a discussion of our results. The content of this chapter is closely related to our preprint [93].

4.1 Chemical reactions and pH in macromolecular systems

We study a multicomponent mixture of macromolecules which can exist in three different charge states. Macromolecules with a maximal positive net charge $+m$ are denoted by M^+, those with maximal negative net charge $-m$ by M^- and neutral macromolecules are denoted by M. We also consider water molecules H_2O, hydronium ions H_3O^+ and hydroxide ions OH^-. We describe both protonation and deprotonation of the macromolecules as well as the self-ionization of water with the following chemical reactions

$$M^- + m\,H_3O^+ \;\rightleftharpoons\; M + m\,H_2O \quad, \tag{4.1a}$$

$$M + m\,H_3O^+ \;\rightleftharpoons\; M^+ + m\,H_2O \quad, \tag{4.1b}$$

$$H_3O^+ + OH^- \;\rightleftharpoons\; 2\,H_2O \quad. \tag{4.1c}$$

The average charge state of the macromolecules determined from the reactions defined in Eqs. (4.1) is controlled by the pH of the mixture. The pH is defined as [94]

$$\text{pH} = -\log_{10} a_{\text{H}+} \quad , \tag{4.2}$$

with the relative activity of the proton $a_{\text{H}+}$ given by

$$a_{\text{H}+} = \exp\left(\frac{\mu_{\text{H}+} - \mu_{\text{H}+}^0}{k_{\text{B}}T}\right) \quad , \tag{4.3}$$

where k_{B} is the Boltzmann constant, $\mu_{\text{H}+}$ is the chemical potential of protons in the system and $\mu_{\text{H}+}^0$ denotes the chemical potential of protons in a reference state [95]. The definition of pH in Eq. (4.2) refers to the proton activity. Protons in water are typically hydrated [96, 97]. At chemical equilibrium, the proton hydration reaction $\text{H}^+ + \text{H}_2\text{O} \rightleftharpoons \text{H}_3\text{O}^+$, implies the relation $\mu_{\text{H}+} = \mu_{\text{H}_3\text{O}+} - \mu_{\text{H}_2\text{O}}$, where $\mu_{\text{H}_3\text{O}+}$ is the chemical potential of the hydronium ions and $\mu_{\text{H}_2\text{O}}$ is the chemical potential of water. Therefore, the proton activity Eq. (4.3) can be written as

$$a_{\text{H}+} = \exp\frac{\left(\mu_{\text{H}_3\text{O}+} - \mu_{\text{H}_2\text{O}}\right) - \left(\mu_{\text{H}_3\text{O}+}^0 - \mu_{\text{H}_2\text{O}}^0\right)}{k_{\text{B}}T} \, , \tag{4.4}$$

where the reference chemical potentials, $\mu_{\text{H}_3\text{O}+}^0$ and $\mu_{\text{H}_2\text{O}}^0$ define the pH scale. A standard choice for the reference chemical potentials are the chemical potential of the hydronium ions $\mu_{\text{H}_3\text{O}+}^0$ at strong dilution evaluated at a standard concentration ($n_{\text{H}_3\text{O}+}^0 = 1\,\text{M}$) and the chemical potential of pure water $\mu_{\text{H}_2\text{O}}^0$ [95]. In the strong dilution limit, the proton activity is $a_{\text{H}+} \simeq n_{\text{H}_3\text{O}+}/n_{\text{H}_3\text{O}+}^0$, leading to $\text{pH} \simeq -\log_{10} n_{\text{H}_3\text{O}+}$, where $n_{\text{H}_3\text{O}+}$ denotes the concentration of the H_3O^+ ions (see Appendix). In this paper we use Eqs. (4.2) and (4.4) to define the pH. Next, we describe a thermodynamic framework to quantify the effect of pH on phase separation behaviour in this macromolecular system.

4.2 Thermodynamics of macromolecular systems

For the sake of clarity, we briefly re-introduce the thermodynamics of an incompressible multicomponent mixture. We consider an incompressible multicomponent mixture in the (T, P, N_i) ensemble with temperature T, pressure P and N_i denoting the number of particles of component i in the mixture. The corresponding thermodynamic potential is the Gibbs free energy $G(T, P, N_i)$. The chemical potential is defined by $\mu_i = \partial G/\partial N_i|_{T,P,N_{j\neq i}}$, the entropy is $S = -\partial G/\partial T|_{P,N_i}$ and the volume of the system is $V = \partial G/\partial P|_{T,N_i}$. By incompressibility we mean that the molecular volumes of each component, $v_i = \partial V/\partial N_i|_{T,P,N_{j\neq i}}$ are independent of pressure and composition. The volume density of the Gibbs free energy is given by $g(T, P, n_i) = G(T, P, N_i)/V(N_i)$, where

we have introduced the concentrations $n_i = N_i/V$ and the volume $V(N_i) = \sum_i v_i N_i$. The chemical potentials can then be calculated from the Gibbs free energy density by

$$\mu_i = v_i \left(g - \sum_k \frac{\partial g}{\partial n_k} n_k \right) + \frac{\partial g}{\partial n_i} \quad . \tag{4.5}$$

We study the multicomponent mixture using a Flory-Huggins mean field free energy model where the Gibbs free energy density reads [59, 60]:

$$g = k_B T \sum_k n_k \ln(n_k v_k) + \sum_k w_k n_k + \sum_{kl} \frac{\Lambda_{kl}}{2} n_k n_l + P \quad . \tag{4.6}$$

The logarithmic terms stem from the mixing entropy, w_k denote internal free energies of molecules of type k and the interaction parameters Λ_{kl} describe the contribution to the free energy due to molecular interactions. Molecular interactions can outcompete the mixing entropy and cause the emergence of coexisting phases. Using the free energy density Eq. (4.6), the chemical potentials are

$$\mu_i = v_i(P - \Sigma) + w_i + k_B T(\ln(n_i v_i) + 1) + \sum_k \Lambda_{ik} n_k \quad , \tag{4.7}$$

where Σ is defined by

$$\Sigma = \sum_{kl} \frac{\Lambda_{kl}}{2} n_k n_l + k_B T \sum_k n_k \quad . \tag{4.8}$$

In the multicomponent mixture we consider that the indices i, k and l run over the six components of the system, which are the three charge states of the macromolecules M, M^+, M^- as well as the three charge states of water H_2O, H_3O^+ and OH^-. We further consider the molecular volumes of the macromolecules to be all equal, $v = v_M = v_{M^+} = v_{M^-}$, where we have introduced the macromolecular volume v. We also consider the molecular volumes of water and water ions to be the same, $v_0 = v_{H_2O} = v_{H_3O^+} = v_{OH^-}$, and denote them by v_0. In order to have a minimal number of components, the presence of salt and counterions which neutralize our solution are taken into account implicitly. The presence of salt and counterions screen the electrostatic interactions, thus providing a characteristic length scale of the interaction potentials between charged species, while the counterions mediate the interactions between oppositely charged macromolecules, hence giving an effective interaction between them. The effective interactions in our system are captured by the interaction matrix Λ_{ij}. We consider all effective interactions $\Lambda_{ij} = 0$ except for those between positively and negatively charged macromolecules which we choose as $\Lambda_{M^-M^+} = \Lambda_{M^+M^-} = v\chi_e/\epsilon$ and the effective interaction between neutral macromolecules given by $\Lambda_{MM} = 2v\chi_n/\epsilon$. Here we have introduced the molecular volumes ratio $\epsilon = v_0/v$ as

well as χ_e and χ_n which are interaction parameters characterizing the strength of charge-charge and neutral-neutral interactions respectively, with this choice, the interaction parameters describe the interaction scale corresponding to a water molecule. Attractive interactions are described by negative values of these interaction parameters, which we will vary in our study of phase behaviour. For these interaction parameters, we obtain the following chemical potentials

$$\mu_{\mathrm{M}} = k_{\mathrm{B}}T\left(\ln(v\,n_{\mathrm{M}})+1\right)+2\frac{v}{\epsilon}\chi_{\mathrm{n}}n_{\mathrm{M}}+w_{\mathrm{M}}+v(P-\Sigma) \quad, \tag{4.9a}$$

$$\mu_{\mathrm{M^+}} = k_{\mathrm{B}}T\left(\ln(v\,n_{\mathrm{M^+}})+1\right)+\frac{v}{\epsilon}\chi_{\mathrm{e}}n_{\mathrm{M^-}}+w_{\mathrm{M^+}}+v(P-\Sigma) \quad, \tag{4.9b}$$

$$\mu_{\mathrm{M^-}} = k_{\mathrm{B}}T\left(\ln(v\,n_{\mathrm{M^-}})+1\right)+\frac{v}{\epsilon}\chi_{\mathrm{e}}n_{\mathrm{M^+}}+w_{\mathrm{M^-}}+v(P-\Sigma) \quad, \tag{4.9c}$$

$$\mu_{\mathrm{H_3O^+}} = k_BT\left(\ln\left(v_0 n_{\mathrm{H_3O^+}}\right)+1\right)+w_{\mathrm{H_3O^+}}+v_0(P-\Sigma) \quad, \tag{4.9d}$$

$$\mu_{\mathrm{OH^-}} = k_BT\left(\ln\left(v_0 n_{\mathrm{OH^-}}\right)+1\right)+w_{\mathrm{OH^-}}+v_0(P-\Sigma) \quad, \tag{4.9e}$$

$$\mu_{\mathrm{H_2O}} = k_BT\left(\ln\left(v_0 n_{\mathrm{H_2O}}\right)+1\right)+w_{\mathrm{H_2O}}+v_0(P-\Sigma) \quad. \tag{4.9f}$$

In the following, we study the system at chemical equilibrium for a fixed pH.

4.3 Chemical equilibrium at fixed pH

We start by stating the conservation laws for a system undergoing the reactions given in Eqs. (4.1) and choose the conserved quantities as independent composition variables at chemical equilibrium. We then use these independent composition variables to identify the thermodynamic conjugated variables at chemical equilibrium. We finalize the section by constructing a thermodynamic potential which describes the system at a fixed pH value.

4.3.1 Chemical conservation laws

We consider a system with s different molecular species. If there are r different chemical reactions taking place, there exist $c = s - r$ independent composition variables. Here $s = 6$, the number of independent reactions is $r = 3$, therefore the number of independent composition variables is $c = 3$. These independent composition variables can be chosen as conserved quantities during chemical reaction events. Although the number of independent composition variables is fixed, there is not a unique choice of conserved variables [76]. From the reaction scheme (4.1) we readily identify three conserved components in every chemical reaction, their associated particle numbers are

$$N = N_{\mathrm{M^+}} + N_{\mathrm{M^-}} + N_{\mathrm{M}} \quad, \tag{4.10}$$

$$N_{\mathrm{H}} = 3N_{\mathrm{H_3O^+}} + m(2N_{\mathrm{M^+}} + N_{\mathrm{M}}) + N_{\mathrm{OH^-}} + 2N_{\mathrm{H_2O}} \quad, \tag{4.11}$$

$$N_s = N_{\mathrm{H_3O^+}} + N_{\mathrm{OH^-}} + N_{\mathrm{H_2O}} \quad. \tag{4.12}$$

The three conserved particle numbers are: N the total number of macromolecules, N_H the number of hydrogen atoms (more precisely, protons) and N_s is the number of oxygen atoms. We can combine these equations to show that the partial charge involved in the chemical reactions $N_q = N_H - 2N_s - mN$ is also a constant given by

$$N_q = N_{\mathrm{H_3O^+}} + m(N_{\mathrm{M^+}} - N_{\mathrm{M^-}}) - N_{\mathrm{OH^-}} \quad . \tag{4.13}$$

In the next section it will become clear why do we use the net charge involved in the chemical reactions N_q instead of N_H. Please note that we consider the net charge to be neutralized by counterions that are not explicitly considered in the simplified model, considering explicitly the electrical neutrality condition leads to interesting conclusions which will be discussed in the next chapter of this Thesis.

4.3.2   Conjugate thermodynamic variables at chemical equilibrium

Here, we obtain conditions for chemical equilibrium in terms of the conserved variables defined in Eqs. (4.10), (4.12) and (4.13). To do so, we use the variable transformation $(N_{\mathrm{M}}, N_{\mathrm{H_3O^+}}, N_{\mathrm{H_2O}}) \to (N, N_s, N_q)$ to eliminate N_{M}, $N_{\mathrm{H_3O^+}}$ and $N_{\mathrm{H_2O}}$. The differential of the Gibbs free energy is given by $dG = -SdT + VdP + \sum_i \mu_i dN_i$, which after this variable transformation becomes:

$$
\begin{aligned}
dG \;=\; & -S\,dT + V\,dP + \mu_{\mathrm{M}}\,dN + \mu_{\mathrm{H_2O}}\,dN_s + (\mu_{H_3O^+} - \mu_{\mathrm{H_2O}})dN_q \\
& + (\mu_{\mathrm{M^-}} + m\,\mu_{\mathrm{H_3O^+}} - \mu_{\mathrm{M}} - m\,\mu_{\mathrm{H_2O}})\,dN_{\mathrm{M^-}} + (\mu_{\mathrm{M^+}} \\
& + m\,\mu_{\mathrm{H_2O}} - \mu_{\mathrm{M}} - m\,\mu_{\mathrm{H_3O^+}})\,dN_{\mathrm{M^+}} \\
& + (\mu_{\mathrm{H_3O^+}} + \mu_{\mathrm{OH^-}} - 2\mu_{\mathrm{H_2O}})\,dN_{\mathrm{OH^-}} \quad .
\end{aligned}
\tag{4.14}
$$

If the system has reached chemical equilibrium, the variations dG with respect to changes in the non-conserved composition variables $N_{\mathrm{M^+}}$, $N_{\mathrm{M^-}}$ and $N_{\mathrm{OH^-}}$ must vanish. This then leads to the following chemical equilibrium conditions [64]:

$$\mu_{\mathrm{M^-}} + m\,\mu_{\mathrm{H_3O^+}} \;=\; \mu_{\mathrm{M}} + m\,\mu_{\mathrm{H_2O}} \quad , \tag{4.15a}$$

$$\mu_{\mathrm{M}} + m\,\mu_{\mathrm{H_3O^+}} \;=\; \mu_{\mathrm{M^+}} + m\,\mu_{\mathrm{H_2O}} \quad , \tag{4.15b}$$

$$\mu_{\mathrm{H_3O^+}} + \mu_{\mathrm{OH^-}} \;=\; 2\,\mu_{\mathrm{H_2O}} \quad . \tag{4.15c}$$

Using the chemical equilibrium conditions, the differential of the Gibbs free energy at chemical equilibrium is given by

$$dG \;=\; -S\,dT + V\,dP + \mu_{\mathrm{M}}\,dN + \mu_{\mathrm{H_2O}}\,dN_s + (\mu_{H_3O^+} - \mu_{\mathrm{H_2O}})dN_q \quad , \tag{4.16}$$

which explicitly shows that the Gibbs free energy at chemical equilibrium has the dependence $G(T, P, N, N_s, N_q)$. This allows us to identify pairs of conjugate thermodynamic variables. From Eq. (4.16) we identify the conjugate thermodynamic variables to the composition variables (N, N_s, N_q) as $(\mu_{\mathrm{M}}, \mu_{\mathrm{H_2O}}, \mu_{H_3O^+} - \mu_{H_2O})$ respectively. These conjugate variables have to be used to obtain Legendre transforms of the thermodynamic potentials at chemical equilibrium [76].

4.3.3 Thermodynamic ensemble for fixed pH

In order to describe the system in an ensemble with pH as a variable, we perform a Legendre transform to construct a thermodynamic potential which depends on $\mu_{\mathrm{H_3O^+}} - \mu_{\mathrm{H_2O}}$[1]. This thermodynamic potential is given by the following Legendre transform

$$\bar{G}(T, P, N, N_s, \mu_{\mathrm{H_3O^+}} - \mu_{\mathrm{H_2O}}) = G - (\mu_{\mathrm{H_3O^+}} - \mu_{\mathrm{H_2O}})N_q \quad . \tag{4.17}$$

The differential of $\bar{G}$ reads

$$\begin{aligned} d\bar{G} &= -SdT + VdP + \mu_{\mathrm{M}}\, dN + \mu_{\mathrm{H_2O}}\, dN_s \\ &\quad - N_q\, d(\mu_{\mathrm{H_3O^+}} - \mu_{\mathrm{H_2O}}) \quad . \end{aligned} \tag{4.18}$$

We now clarify why fixing this chemical potential difference and the temperature sets the pH value of the system. Using Eq. (4.2) and Eq. (4.4), we can express the pH as

$$\mathrm{pH} = \frac{\left(\mu_{\mathrm{H_3O^+}} - \mu_{\mathrm{H_2O}}\right) - \left(\mu_{\mathrm{H_3O^+}}^0 - \mu_{\mathrm{H_2O}}^0\right)}{k_{\mathrm{B}}T} \log_{10} e \quad , \tag{4.19}$$

where it is explicitly shown that the pH of the system is set by the relative chemical potential of hydronium ions with respect to water, $\mu_{\mathrm{H_3O^+}} - \mu_{\mathrm{H_2O}}$, and the temperature T of the system. We can also define the corresponding free energy in the isochoric ensemble as

$$\bar{F} = G - (\mu_{\mathrm{H_3O^+}} - \mu_{\mathrm{H_2O}})N_q - PV \quad , \tag{4.20}$$

which has the following differential form

$$\begin{aligned} d\bar{F} &= -SdT - PdV + \mu_{\mathrm{M}}\, dN + \mu_{\mathrm{H_2O}}\, dN_s \\ &\quad - N_q\, d(\mu_{\mathrm{H_3O^+}} - \mu_{\mathrm{H_2O}}) \quad . \end{aligned} \tag{4.21}$$

In our system, the volume V can be expressed in terms of the conserved quantities as $V = vN + v_0 N_s$, leading to $dN_s = dV/v_0 - vdN/v_0$. We can then reduce the number of

---

[1]A free energy for fixed pH has been proposed before [69, 76, 77] but instead of using a definition of pH which considers the chemical potential of protons as we do here, the pH was considered only in the dilute limit for which a definition based on the concentration of hydronium ions is considered.

independent variables and rewrite the differential form of $\bar{F}$ as

$$d\bar{F} \quad = \quad -S\,dT - \Pi\,dV + \bar{\mu}_{\mathrm{M}}\,dN - N_q\,d\bar{\mu}_{\mathrm{H_3O^+}} \quad , \tag{4.22}$$

where we have introduced the exchange chemical potentials $\bar{\mu}_{\mathrm{M}}$ of neutral macromolecules and $\bar{\mu}_{\mathrm{H_3O^+}}$ of hydronium ions as well as the osmotic pressure Π, which are defined by

$$\bar{\mu}_i \quad = \quad \mu_i - \frac{v_i}{v_0}\mu_{\mathrm{H_2O}} \quad , \tag{4.23}$$

$$\Pi \quad = \quad P - \frac{\mu_{\mathrm{H_2O}}}{v_0} \quad , \tag{4.24}$$

where $i = \mathrm{M}$ or $\mathrm{H_3O^+}$. The free energy $\bar{F}(T, V, N, \bar{\mu}_{\mathrm{H_3O^+}})$ depends only on the temperature T, the volume V, the total macromolecule particle number N and the exchange chemical potential of the hydrogen ions $\bar{\mu}_{\mathrm{H_3O^+}}$. In addition to introducing the corresponding thermodynamic potential $\bar{F}$ which describes an incompressible system with a fixed pH value, we have reduced our multicomponent system description from six components undergoing three independent chemical reactions to an effective binary mixture with the total macromolecule density $n = N/V$ as the only relevant composition variable. We can now ask how the pH affects phase separation.

4.4   Control of phase separation by pH

In the following, we discuss how the pH controls both, chemical and phase equilibrium in our system. We start by discussing the chemical equilibrium conditions Eqs. (4.15) in terms of newly defined composition variables and thermodynamic field. After discussing the chemical equilibrium conditions in terms of the pH and the newly defined fields, we provide the conditions for phase equilibrium for a system described by the corresponding thermodynamic potential defined by Eq. (4.20). We end the section by showing a construction of the coexisting phases for a given choice of parameters.

4.4.1   Thermodynamic fields controlling chemical equilibrium

We are now interested in discussing how the pH and other parameters influence the chemical equilibrium described by Eqs. (4.15). In order to do so, we introduce thermodynamic fields and composition variables which allow us to interpret the chemical equilibrium conditions intuitively. Let us first introduce the following composition variables

$$n \quad = \quad n_{\mathrm{M^+}} + n_{\mathrm{M^-}} + n_{\mathrm{M}} \quad , \tag{4.25}$$

$$\phi \quad = \quad \frac{n_{\mathrm{M^+}} + n_{\mathrm{M^-}}}{2n} \quad , \tag{4.26}$$

$$\psi \quad = \quad \frac{n_{\mathrm{M^+}} - n_{\mathrm{M^-}}}{2n} \quad . \tag{4.27}$$

These composition variables express the total concentration of macromolecules n, the fraction of charged macromolecules ϕ and the difference between concentrations of oppositely charged macromolecules relative to the total number of macromolecules ψ.

It is convenient to make a rearrangement of the chemical equilibrium conditions Eqs. (4.15) as follows

$$\mu_{\mathrm{M}^+} + \mu_{\mathrm{M}^-} = 2\,\mu_{\mathrm{M}} \quad , \tag{4.28a}$$

$$\mu_{\mathrm{M}^+} - \mu_{\mathrm{M}^-} = 2m\,(\mu_{\mathrm{H_3O}^+} - \mu_{\mathrm{H_2O}}) \quad , \tag{4.28b}$$

$$\mu_{\mathrm{H_2O}} - \mu_{\mathrm{OH}^-} = \mu_{\mathrm{H_3O}^+} - \mu_{\mathrm{H_2O}} \quad , \tag{4.28c}$$

where we see that the right hand sides of Eqs. (4.28b) and (4.28c) are determined by the pH. We now write the conditions Eqs. (4.28a) and (4.28b) in terms of the composition variables using the expressions of the chemical potentials (Eqs. (4.9) given in the Appendix) leading to

$$k_{\mathrm{B}}T \ln \frac{\phi^2 - \psi^2}{(1 - 2\phi)^2} + \frac{2v\chi_{\mathrm{e}}n\phi}{\epsilon} - \frac{4v\chi_{\mathrm{n}}n(1 - 2\phi)}{\epsilon} = h_\phi \quad , \tag{4.29}$$

$$k_{\mathrm{B}}T \ln \frac{\phi + \psi}{\phi - \psi} - \frac{2v\chi_{\mathrm{e}}n\psi}{\epsilon} = h_\psi \quad , \tag{4.30}$$

where we have defined

$$h_\phi = 2w_{\mathrm{M}} - w_{\mathrm{M}^+} - w_{\mathrm{M}^-} \quad , \tag{4.31}$$

$$h_\psi = 2m\,(\mu_{\mathrm{H_3O}^+} - \mu_{\mathrm{H_2O}}) - w_{\mathrm{M}^+} + w_{\mathrm{M}^-} \quad . \tag{4.32}$$

These quantities play the role of fields controlling the chemical equilibrium and phase separation behaviour of the system. The molecular field h_ϕ characterizes which of the macromolecular charge states is energetically favoured due to their internal free energies w_i, while the field h_ψ expresses deviations of the pH from its value at the isoelectric point (pI) of the macromolecules. Remember that the isoelectric point is the value of the pH for which macromolecules are on average neutral. Thus, in our system, the isoelectric point is defined as the value pI of the pH at which the charged macromolecules obey, $n_{\mathrm{M}^+} = n_{\mathrm{M}^-}$, or equivalently $\psi = 0$, which implies $h_\psi = 0$ via Eq. (4.30). Using Eq. (4.19), we find an expression for the pI value:

$$\mathrm{pI} = \left(\frac{w_{\mathrm{M}^+} - w_{\mathrm{M}^-}}{2mk_{\mathrm{B}}T} - \frac{\mu^0_{\mathrm{H_3O}^+} - \mu^0_{\mathrm{H_2O}}}{k_{\mathrm{B}}T} \right) \log_{10} e \quad . \tag{4.33}$$

We can therefore express the field h_ψ in terms of the pH and the pI as follows

$$\frac{h_\psi}{k_{\mathrm{B}}T} = \frac{2m}{\log_{10} e}(\mathrm{pI} - \mathrm{pH}) \quad , \tag{4.34}$$

which explicitly shows how h_ψ characterizes deviations of the system from its isoelectric point. For given h_ϕ and h_ψ, the composition variables ϕ and ψ can be determined from Eqs. (4.29) and (4.30) as a function of the total macromolecule density n, the temperature T and the pH. One symmetry can be identified in Eqs. (4.28a) and (4.28b), namely that the system behaves identically under the transformation $\psi \to -\psi$ and $h_\psi \to -h_\psi$, which will be reflected in the phase diagrams as a function of deviations from the isoelectric point. This symmetry stems from both, considering that positively and negatively charged macromolecules have the same interaction with the rest of the components as well as from choosing their molecular volumes to be the same. Changing any of the two previously mentioned conditions would break this symmetry. We must note that in a more realistic scenario there would be differences in solvation of the two different charged states of the macromolecule [98]. Using Eqs. (4.9d)-(4.9f), we introduce two more fields controlling the relative concentrations of hydronium and hydroxide ions with respect to water molecules, these fields obey

$$\log\left(\frac{n_{\mathrm{H_3O^+}}}{n_{\mathrm{H_2O}}}\right) = \frac{h_{\mathrm{H}}}{k_{\mathrm{B}}T} + \frac{h_\psi}{2m} \quad , \tag{4.35}$$

$$\log\left(\frac{n_{\mathrm{OH^-}}}{n_{\mathrm{H_2O}}}\right) = \frac{h_{\mathrm{O}}}{k_{\mathrm{B}}T} - \frac{h_\psi}{2m} \quad , \tag{4.36}$$

where the fields h_{H} and h_{O} are defined by

$$h_{\mathrm{H}} = \frac{w_{\mathrm{M^+}} - w_{\mathrm{M^-}}}{2m} - w_{\mathrm{H_3O^+}} + w_{\mathrm{H_2O}} \quad , \tag{4.37}$$

$$h_{\mathrm{O}} = \frac{w_{\mathrm{M^-}} - w_{\mathrm{M^+}}}{2m} - w_{\mathrm{OH^-}} + w_{\mathrm{H_2O}} \quad . \tag{4.38}$$

We can go further and use the condition $v_0 n_s + vn = 1$ and Eqs. (4.35) and (4.36) to express the concentrations of hydronium ions, hydroxide ions and water molecules as a function of the total macromolecule density n, temperature T and pH (or h_ψ), given by

$$n_{\mathrm{H_2O}} = \frac{1 - vn}{v_0 \left(1 + e^{h_{\mathrm{H}}/k_{\mathrm{B}}T + h_\psi/2m} + e^{h_{\mathrm{O}}/k_{\mathrm{B}}T - h_\psi/2m}\right)} \quad , \tag{4.39}$$

$$n_{\mathrm{H_3O^+}} = \frac{(1 - vn)\, e^{h_{\mathrm{H}}/k_{\mathrm{B}}T + h_\psi/2m}}{v_0 \left(1 + e^{h_{\mathrm{H}}/k_{\mathrm{B}}T + h_\psi/2m} + e^{h_{\mathrm{O}}/k_{\mathrm{B}}T - h_\psi/2m}\right)} \quad , \tag{4.40}$$

$$n_{\mathrm{OH^-}} = \frac{(1 - vn)\, e^{h_{\mathrm{O}}/k_{\mathrm{B}}T - h_\psi/2m}}{v_0 \left(1 + e^{h_{\mathrm{H}}/k_{\mathrm{B}}T + h_\psi/2m} + e^{h_{\mathrm{O}}/k_{\mathrm{B}}T - h_\psi/2m}\right)} \quad . \tag{4.41}$$

We have shown that chemical equilibrium can be fully accounted for by the internal free energies of all species w_i which we take to be constant, the temperature T, the pH (or equivalently the chemical potential difference $\mu_{\mathrm{H_3O^+}} - \mu_{\mathrm{H_2O}}$) and the total macromolecule concentration n.

4.4.2   Phase coexistence in the pH ensemble

We are interested in describing the phase behaviour of the system in the pH ensemble. To this end we make use of the composition variables n, ϕ and ψ and the fields h_ϕ, h_ψ, h_H and h_O defined in the previous section. Using Eq. (4.20) we define the free energy density $\bar{f}(T, n, \bar{\mu}_{H_3O^+}) = \bar{F}(T, V, N, \bar{\mu}_{H_3O^+})/V$ which reads

$$
\begin{aligned}
\bar{f}(T, n, \bar{\mu}_{H_3O^+}) \;\; = \;\; & k_B T \Bigg[n(\phi + \psi) \ln(vn(\phi + \psi)) + n(\phi - \psi) \ln(vn(\phi - \psi)) \\
& + n(1 - 2\phi) \ln(vn(1 - 2\phi)) \\
& + \frac{1}{v_0}(1 - vn) \ln \left(\frac{1 - vn}{1 + e^{h_H/k_B T + h_\psi/2mk_B T} + e^{h_O/k_B T - h_\psi/2mk_B T}} \right) \Bigg] \\
& + \frac{v}{\epsilon} \chi_e n^2 (\phi^2 - \psi^2) + \frac{v}{\epsilon} \chi_n n^2 (1 - 2\phi)^2 - h_\phi n\phi - h_\psi n\psi \\
& + w_M n + \frac{w_{H_2O}}{v_0}(1 - vn) \quad ,
\end{aligned}
\tag{4.42}
$$

where the functions $\phi(T, n, \bar{\mu}_{H_3O^+})$ and $\psi(T, n, \bar{\mu}_{H_3O^+})$ are defined implicitly in Eqs. (4.29) and (4.30) in terms of the temperature T, the exchange chemical potential of hydronium ions $\bar{\mu}_{H_3O^+}$ and the total macromolecule density n.

We now discuss the phase coexistence conditions for this incompressible system at fixed temperature T and at a fixed pH value, which can be obtained by a Maxwell construction. It follows from Eq. (4.22) that the exchange chemical potential of neutral macromolecules is given by $\bar{\mu}_M = \partial \bar{f}/\partial n \big|_{T, \bar{\mu}_{H_3O^+}}$ and that the osmotic pressure is given by $\Pi = \bar{f} - n\, \partial \bar{f}/\partial n \big|_{T, \bar{\mu}_{H_3O^+}}$. Using the free energy density $\bar{f}$, we write the phase equilibrium conditions describing equal exchange chemical potentials of the neutral macromolecules and equal osmotic pressures in both phases [68]:

$$
\begin{aligned}
\bar{\mu}_M(n^I) \;\; &= \;\; \bar{\mu}_M(n^{II}) \quad , & (4.43a) \\
\bar{\mu}_M(n^I) \;\; &= \;\; \frac{\bar{f}(n^{II}) - \bar{f}(n^I)}{n^{II} - n^I} \quad , & (4.43b)
\end{aligned}
$$

where the superscripts I, II denote the two coexisting phases. We did not write the explicit dependence of the relative chemical potential $\bar{\mu}_M$ and of the free energy density $\bar{f}$ on the temperature T and on the relative chemical potential $\bar{\mu}_{H_3O^+}$. These conditions correspond to the common tangent construction [68]. We can find the coexisting phases by first calculating the values of ϕ and ψ using Eqs. (4.29) and (4.30) (see Fig. 4.1(a,b)), and then by doing a common tangent construction for the free energy density Eq. (4.42) evaluated as a function of the total macromolecule volume fraction $\bar{n} = vn$ (Fig. 4.1(c)). The phase diagrams can then be readily constructed by repeating the steps described above for different parameter values.
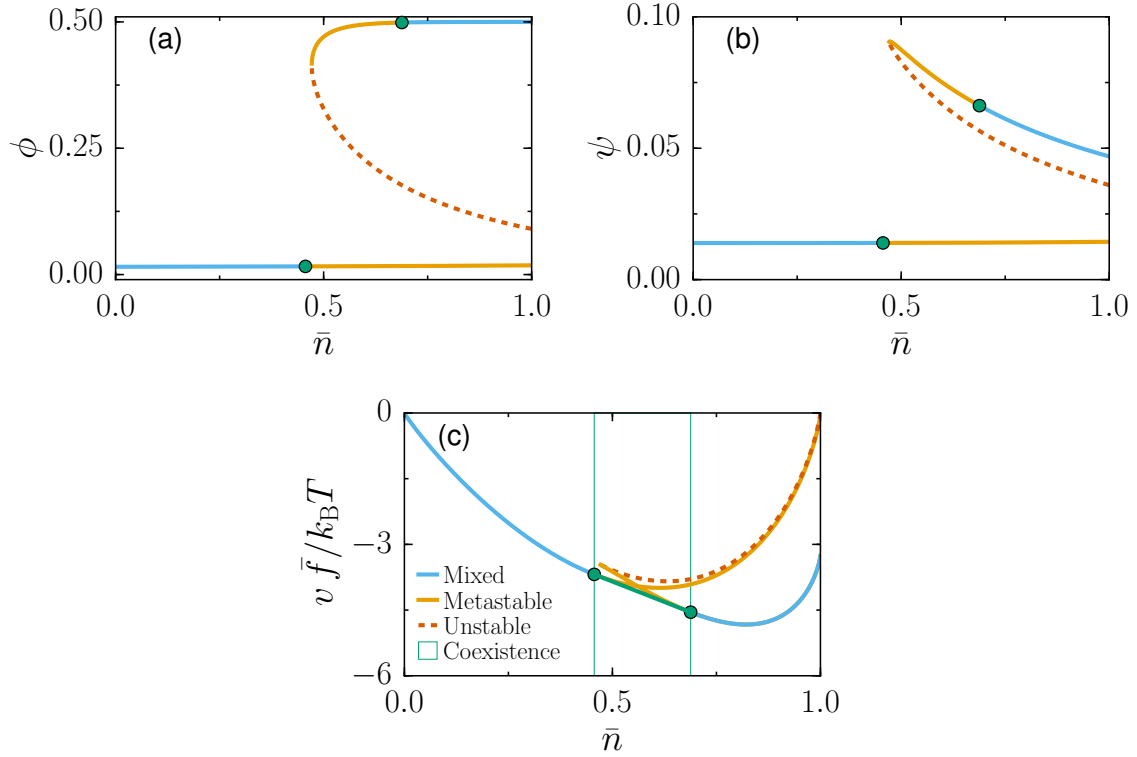
FIGURE 4.1: Chemical equilibrium conditions and free energy density at chemical equilibrium. Multiple solutions are found for ϕ (a) and ψ (b) as a function of the total macromolecule volume fraction $\bar{n}$, enabling the system to exhibit phase separation between different branches of the chemical equilibrium. The blue solid lines correspond to equilibrium concentrations where the system remains homogeneous, the orange solid lines represent solutions to the chemical equilibrium relations which are metastable states and the dotted red line shows the unstable states. (c) Maxwell construction for the dimensionless free energy density $v\bar{f}/k_BT$ as a function of the total macromolecule volume fraction, the green line describes the region of macromolecule volume fraction where the system split into two phases with different compositions given by the green circles. Parameters $\chi_e/k_BT = -3$, $\chi_n = 0$, $pI - pH = 0.2$, $h_\phi/k_BT = -10$ and $\epsilon = 0.1$, apply to all panels.

4.5   Phase diagrams at the isoelectric point

In this section, we investigate different phase diagrams which can be obtained by varying the temperature and the total number of macromolecules, keeping the system at its iso-electric point. In order to discuss the effects of temperature we use a weighted sum of the interaction parameters $\chi = \chi_e + 4\chi_n$, the ratio of interaction parameters $\lambda = \frac{2\chi_n}{\chi_e + 4\chi_n}$, the molecular field h_ϕ and rewrite Eqs. (4.29) and (4.42) in the following compact way:

$$h_\phi \quad = \quad 2k_B T \ln \frac{\phi}{(1 - 2\phi)} + \frac{2\chi \bar{n}(\phi - \lambda)}{\epsilon} \quad , \tag{4.44}$$

$$v\bar{f}(T, n, \bar{\mu}_{H_3O^+}) \quad = \quad k_B T \left[2\bar{n}\phi \ln(\bar{n}\phi) + \bar{n}(1 - 2\phi) \ln(\bar{n}(1 - 2\phi)) \right.$$
$$\left. + \frac{1}{\epsilon}(1 - \bar{n}) \ln \left(\frac{1 - \bar{n}}{1 + e^{h_H/k_B T} + e^{h_O/k_B T}} \right) \right]$$
$$+ \chi \epsilon^{-1} \bar{n}^2 (\phi^2 - 2\lambda\phi + \lambda/2) - h_\phi \bar{n}\phi$$
$$+ \left(w_M - \frac{w_{H_2O}}{\epsilon} \right) \bar{n} + \frac{w_{H_2O}}{\epsilon} \quad . \tag{4.45}$$

We further consider attractive interactions $\chi < 0$ and $\lambda > 0$. In order to construct phase diagrams as a function of temperature we rescale all the variables which have energy units with $k_B T_0$, where T_0 is a reference temperature. Free energy minimization at constant $h_\phi, \chi, \lambda, \epsilon, h_H, h_O$ and T together with a common tangent construction, lead to different possible topologies of phase diagrams which are summarized in Fig. 4.2 and Fig. 4.3.

   For $2\chi(1/2 - \lambda) > \epsilon h_\phi$, or $\epsilon h_\phi > -2\chi\lambda$, the diagram has the same topology as that of a simple two component mixture [65], see Fig. 4.2(a,d,e,h). At low temperature the system demixes in a low density and a high density phase. For $\epsilon h_\phi \ll 2\chi(1/2 - \lambda)$, the proportion of charged molecules is exponentially small and the system behaves like a neutral polymer solution. In this case the system separates into a low density phase and a high density phase composed essentially of neutral macromolecules $\phi \simeq 0$. The coexistence curve is bell shaped and by construction the tie lines are parallel to the $\bar{n}$ axis. The point at which the tangent is parallel to this axis is a critical point (Fig. 4.2(a,e)). It belongs to the same universality class as a liquid-vapor critical point. For $\epsilon h_\phi \gg -2\chi\lambda$, the concentration of neutral molecules is exponentially small and the system demixes between a low density phase and a high density phase composed essentially of charged molecules $\phi \simeq 1/2$. Again the coexistence curve is "bell shaped" and one observes the existence of an isolated critical point (Fig. 4.2(d,h)).

   In both limits, the mean-field calculation of the critical coordinates can be done an-alytically. The details are given in the Appendix. In these limits, the critical values are

given by:

$$\bar{n}_c^b \quad = \quad \frac{\sqrt{\epsilon}}{1+\sqrt{\epsilon}} \quad , \text{ for } \phi = 0 \text{ and } \phi = \frac{1}{2} \; , \tag{4.46a}$$

$$kT_c^b \quad = \quad -\frac{2\chi_n}{(1+\sqrt{\epsilon})^2} \quad , \text{ for } \phi = 0 \; , \tag{4.46b}$$

$$kT_c^b \quad = \quad -\frac{\chi_e}{2\,(1+\sqrt{\epsilon})^2} \quad , \text{ for } \phi = \frac{1}{2} \; , \tag{4.46c}$$

where we have used the condition for critical points, $\partial^2 \bar{f}/\partial n^2 = 0$ and $\partial^3 \bar{f}/\partial n^3 = 0$. We refer hereafter to these coexistence regions as quasi-binary regions.

For intermediate values of the molecular field, $2\chi(1/2 - \lambda) \lesssim \epsilon\, h_\phi \lesssim -2\chi\lambda$ the possible topologies of phase diagrams are more complex. Increasing the value of h_ϕ, from very large negative values towards positive values, leads to the emergence of a second coexistence region (Fig. 4.2(b,f)). This coexistence region, disconnected from the quasi-binary region, is bounded by a critical point and it is connected to the $\bar{n} = 1$ axis at a single point, where the tie line span vanishes. The transition point corresponds to a first order transition on the $\bar{n}$ line, where ϕ undergoes a discontinuity. A second coexistence region emerges via a critical point(Fig. 4.3(a,d)), whose values are given by:

$$\phi_c \quad = \quad \frac{1}{4} \; , \tag{4.47a}$$

$$kT_c \quad = \quad -\frac{\chi}{8\epsilon} \quad , \tag{4.47b}$$

$$h_{\phi,c} \quad = \quad \frac{\chi(2 + \ln(2) - 8\lambda)}{4\epsilon} \quad . \tag{4.47c}$$

The two phase region collapses to one point on the $\bar{n} = 1$ line, both in the first order and in the second order scenarios, this results from the existence of only one singularity in ϕ on the $\bar{n} = 1$ line. Note that $h_{\phi,c}$ can be either positive or negative, so that one can have a critical point on the $\bar{n} = 1$ line both in the neutral and in the charged regimes, see the Supporting Material.

For some values of h_ϕ, with $h_\phi > h_{\phi,c}$, the two coexistence regions merge (Fig. 4.2(c,g)). Depending on λ, there may be two different generic scenarios. We first explain what happens for $\lambda < 1/4$. In this case, the two regions merge, giving rise to two triple points (Fig. 4.3(f)). The two triple points have a low density phase enriched in neutral macromolecules coexisting with both, an intermediate phase with a large macromolecule concentration, which is also rich in neutral macromolecules and with a macromolecule dense phase of essentially charged macromolecules. For $\epsilon\, h_\phi > 2\chi(1/4 - \lambda)$, one triple point and the first order transition point vanish. This bound for h_ϕ is found by solving the coexistence conditions at $\bar{n} = 1$, for $T = 0$. For increasing values of h_ϕ, the remaining triple point moves towards and eventually merges with the critical point of the quasi-binary region,

leading to a coexistence region which has only one critical point (Fig. 4.2(c)). Finally, for larger values of h_ϕ, the system behaves as a binary mixture of charged macromolecules and solvent (Fig. 4.2(d)). In contrast, for $\lambda > 1/4$ and $h_\phi > h_{\phi,c}$, when the two regions merge, there is only one triple point, where two phases which are essentially enriched in charged macromolecules, coexist with a high concentration phase enriched in neutral macromolecules (Fig. 4.2(g)). For larger values of h_ϕ, the triple point vanishes together with the first order transition point at $T = 0$ and $\epsilon h_\phi = 2\chi(1/4 - \lambda)$. This vanishing leads again to the quasi-binary mixture of charged macromolecules and solvent (Fig. 4.2(h)).

The existence of a transition point on the $\bar{n} = 1$ line leads to the different topologies of the phase diagrams (Fig. 4.3). To understand the behaviour of this point, we analyze the derivative of the free energy density with respect to ϕ at $\bar{n} = 1$, in particular for $\lambda = 0.2$ (Fig. 4.3(a-c)). At the critical temperature, Eq. (4.47b), there is an inflection point (Fig. 4.3(a)), which translates into a critical point for $h_\phi = h_{\phi,c}$ at $T = T_c$ (Fig. 4.3(d)). For lower values of the temperature, $T < T_c$ one finds a first order transition point at $h_\phi > h_{\phi,c}$, in which two phases coexist at $\bar{n} = 1$ (Fig. 4.3(b,e)). For lower temperatures, the region which extends from the first order transition point merges with the rest of the phase diagram (Fig. 4.3(f)) and the derivative of the free energy density becomes increasingly dominated by a linear term in ϕ, given by $2\chi(\phi - \lambda)/\epsilon$ (Fig. 4.3(c)). For $T = 0$, there is a corresponding value of the molecular field, $\epsilon h_\phi = 2\chi(1/4 - \lambda)$, for which we find a solution to the coexistence conditions. For values $\epsilon h_\phi > 2\chi(1/4 - \lambda)$ there is no longer a transition point at $\bar{n} = 1$. We only focus on $\lambda = 0.2$ because the behaviour of the transition point at $\bar{n} = 1$ is similar for $\lambda > 0.25$. Now that we developed a detailed understanding of phase diagrams at the isoelectric point, we can use the developed framework to study the effects of varying pH.

4.6   Phase separation for varying pH

Here we study how deviations in pH with respect to the isoelectric point affect the phase behaviour of the system while keeping the temperature constant. To this end, we study phase diagrams as a function of $\mathrm{pI} - \mathrm{pH}$ (Eq. (4.34)) and the total macromolecular volume fraction $\bar{n}$. We construct phase diagrams for different values of the molecular field h_ϕ, the interaction strength among charged macromolecules χ_e and the interaction strength among neutral macromolecules χ_n. Varying the pH in our model leads to very characteristic features in the phase diagrams which allow us to distinguish the dominant interaction driving phase separation.

Let us first consider the case where neutral molecules are energetically favoured over charged molecules ($h_\phi = -8$, Fig. 4.4(a-c)). In this case, a system with only charge-charge interactions (Fig. 4.4(a), $\chi_\mathrm{n} = 0$) exhibits reentrant behaviour when changing the pH where the corresponding domain in the phase diagram is enclosed by two critical points.
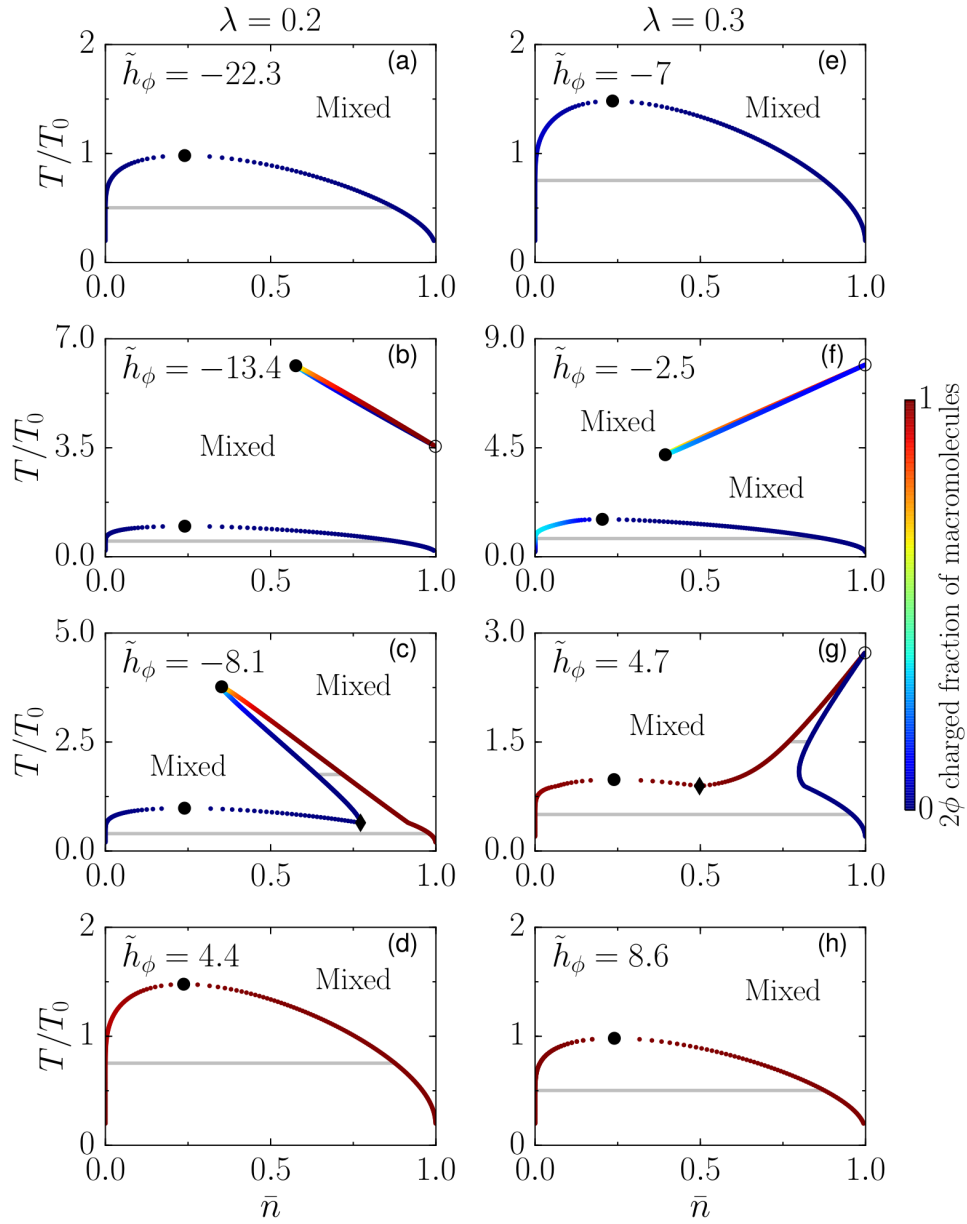
FIGURE 4.2: Topologies of the phase diagrams for varying values of the normalized molecular field $\tilde{h}_\phi = h_\phi/k_B T_0$ defined in Eq. (4.31) where T_0 is a reference temperature. (a-d) phase diagrams for a system in which charge-charge interactions are slightly stronger than neutral-neutral interactions. (e-h) Phase diagrams for a system in which neutral-neutral interactions are slightly stronger than charge-charge interactions. The binodals are given by the colored points which denote coexisting phases. Tie lines (grey solid lines) connect coexisting phases and are horizontal. The regions within the binodals undergo a demixing transition, whereas the regions outside the binodal lines remain well mixed. The critical points where phases become indistinguishable are denoted by black circles, first order transition points where there is a discontinuity in the value of ϕ are denoted by white circles and triple points are denoted by black diamonds. A thorough explanation of the phase diagrams is given in the main text. Parameters $\chi = -8.5$ and $\epsilon = 0.1$, apply to all panels. The colorbar indicates the value of the charged fraction of macromolecules 2ϕ.
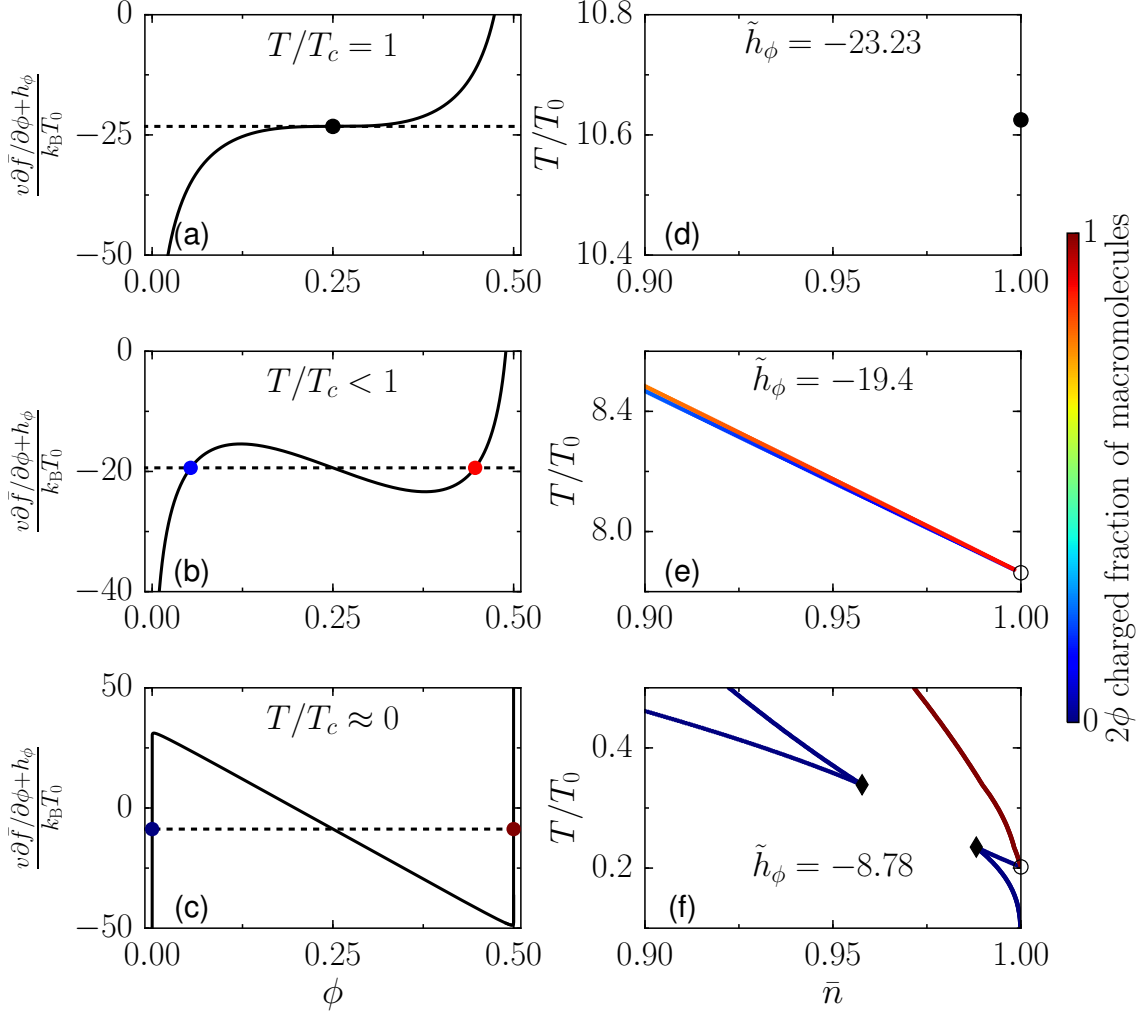
FIGURE 4.3: Critical behaviour on the $\bar{n} = 1$ line. (a-c) Derivative of the free energy with respect to ϕ for different temperature values . (a) Inflection point corresponding to the critical point defined in Eq. (4.47) shown as a filled black circle. The black dotted line shows the value of $\tilde{h}_\phi = \tilde{h}_\phi^c \simeq -23.23$. (b) Emergence of a maximum and a minimum for $T < T_c$, coexisting phases are shown as two colored circles (the color encodes their value of ϕ). (c) Derivative of the free energy density for $T/T_0 = 0.2$, implying $T/T_c \ll 1$. (d-f) Phase diagrams for fixed $\tilde{h}_\phi = h_\phi/k_B T_0$ in the vicinity of the transition point at $\bar{n} = 1$.(d) The filled black circle is the isolated critical point defined in Eq. (4.47) corresponding to (a). (e) The phase coexistence lines shown in blue and red end on the $\bar{n} = 1$ line at the first order transition point (open circle) defined in (b). (f) The coexistence region connected to the $\bar{n} = 1$ axis, merges with the quasi-binary region, leading to the appearance of two triple points (black diamonds in (f)). Parameters $\chi = -8.5$, $\lambda = 0.2$ and $\epsilon = 0.1$, apply to all panels. T_0 is a reference temperature with $T_c/T_0 \simeq 10.6$.
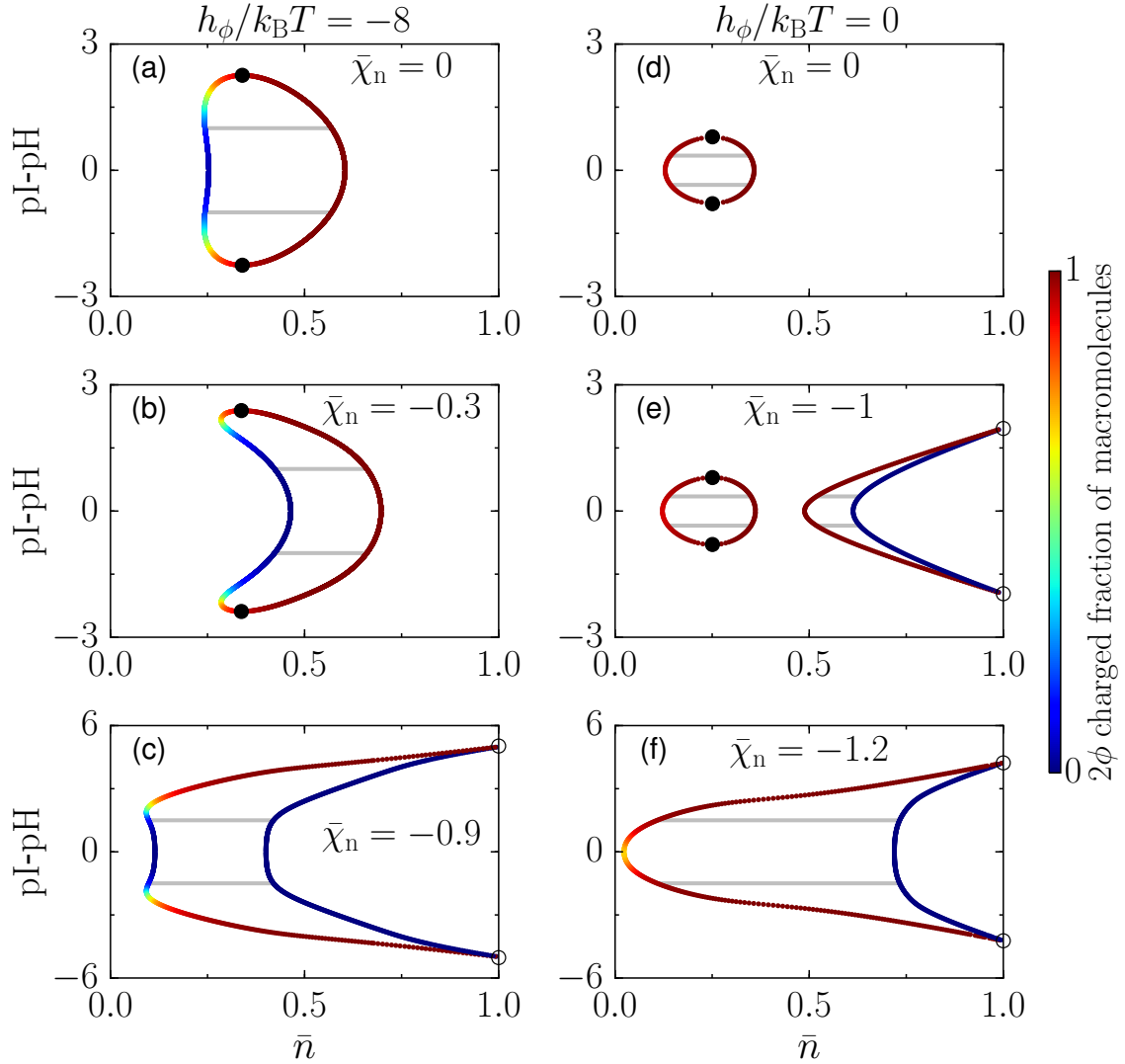
FIGURE 4.4: Phase behaviour as a function of pH for fixed interaction strength between charges $\chi_e/k_BT = -3.5$ and varying values of the interaction strength $\bar{\chi}_n = \chi_n/k_BT$ between neutral macromolecules. (a-c) Phase diagrams with neutral macromolecules energetically favoured $(h_\phi/k_BT = -8)$. (a) In the absence of neutral-neutral interactions there is a small region in the diagram where there is reentrant phase separation behaviour. (b) Small values of neutral-neutral interactions lead to a reduction of the demixing region. (c) Increasing the neutral-neutral attraction even further, the two critical points merge and two first order transition points appear at $\bar{n} = 1$, these points have a discontinuity in ϕ and ψ. (d-f) Phase diagrams with charged macromolecules energetically favoured $(h_\phi/k_BT = 0)$: (d) An effective binary mixture at the pI shows a simple mixing behaviour while deviating from the isoelectric point. (e) For large enough interactions between neutral macromolecules, a second disconnected region appears, such region ends in two first order transition points at $\bar{n} = 1$. (f) Increasing the neutral-neutral interactions further, the two regions merge giving rise to a broadening of the demixing region while the critical points vanish and the coexistence region connects to the $\bar{n} = 1$ line with two first order transition points in which ϕ and ψ have a discontinuity. Parameters: $\epsilon = 0.1$ and $m = 1$, apply to all panels. The colorbar indicates the value of the charged fraction of macromolecules 2ϕ.

Beyond these points, phase separation is not possible for any value of the macromolecular volume fraction $\bar{n}$ while between the critical points, there is a range in $\bar{n}$ where phase separation can occur. The degree of such phase separation is maximal at the isoelectric point pH=pI, which is characterized by the largest difference between coexisting phases in their macromolecular volume fraction, as well as in their charged fraction between coexisting phases. Deviating from the isoelectric point corresponds to lowering the amount of one of the charged components ($\psi \neq 0$, Eq. (4.27)). This change in the relative composition between charged macromolecules decreases the interaction term among charged components (proportional to $n_+ n_-$) and thereby lowers their propensity to phase separate. There is a small range in macromolecular volume fractions where phase separation is absent at the isoelectric point but can be triggered by changing the pH value away from pI. This range strongly increases for stronger interactions among neutral macromolecules (χ_n more negative, Fig. 4.4(b)). Such behaviour for phase separation is unexpected because phase separation occurs despite of an asymmetric ratio of the charged macromolecules. It emerges as a consequence of a reduction in the mixing entropy of the macromolecules by moving away from the isoelectric point which in turn allows the system to phase separate at lower values of charged fraction of macromolecules ϕ. Even though the system shows phase separation at lower macromolecular volume fraction, the region of phase separation decreases for increasing deviations from the isoelectric point. Increasing the attraction among neutral macromolecules even further leads to coexisting phases which are approximately composed of neutral macromolecules and solvent in a range of pH close to the isoelectric point (Fig. 4.4(c)). Moreover, two discontinuous phase transition points emerge while the two critical points merge and vanish. In contrast to the previous two cases (a,b) the broadest range in $\bar{n}$ where phase separation occurs is not located at the isoelectric point (c). We must recognize that having such symmetric phase diagrams for pH deviations below and above the pI is a consequence of our choice of parameters, i.e. the charged macromolecules M^+ and M^- have equal molecular volumes and no interactions with the remaining components. Note that the internal free energies w_{M+} and w_{M-} only affect the value of the pI but not the symmetry of the phase diagrams around the pI.

We now discuss the effects of pH variations for a system in which charged macromolecules are energetically favoured over neutral ones. We start considering a mixture without neutral-neutral interactions (Fig. 4.4(d)) that exhibits a behaviour at the pI resembling a binary mixture of charged macromolecules and solvent (Fig. 4.2(d,h)). For values of pH away from the isoelectric point, we observe a monotonic decrease of the macromolecular order parameter, as well as a fairly constant charged fraction composition in both phases until they meet at two symmetric critical points. After switching on an attractive interaction among neutral macromolecules a second phase separation region appears at larger values of the total macromolecular volume fraction $\bar{n}$ (Fig. 4.4(e)).

This region is characterized by a high density phase mostly composed of neutral macro-molecules coexisting with a phase rich in charged macromolecules. These two phases meet at two first order transition points (open symbol, Fig. 4.4(e)). The appearance of this region is a consequence of having an attraction among neutral macromolecules and charged macromolecules, respectively, favoring phase separation dominantly between both while the solvent is of rather similar concentration in the coexisting phases. Interestingly, the two regions behave independently from each other when increasing the attraction between macromolecules because each region is associated to a different solution of chemical equilibrium (Fig. 4.1). While increasing the attraction further, the two regions merge (Fig. 4.4(f)). This merging leads to a broad region of phase separation corresponding to a large difference in the fraction of charged macromolecules and solvent as well as the vanishing of the two critical points. The high density phase is made of neutral macromolecules which coexist with a low density phase composed of solvent and charged macromolecules. We show the behaviour of ψ along the binodal lines in the same phase diagrams as in Fig. 4.4 in the Appendix.

One typical feature of most phase diagrams (Fig. 4.4(a,b,d-f)) is that the broader region of phase separation exists in the vicinity of the isoelectric point. The region of phase separation shrinks when deviating from the pI due to a decrease in interaction energy among charged macromolecules (Fig. 4.4(a,d)) or in the interaction energy among both charged macromolecules and neutral molecules (Fig. 4.4(b,e,f)). There is only clear exception among these phase diagrams (Fig. 4.4(c)), where the phase separation region slightly increases for pH values away from the isoelectric point. The increase is due to the emergence of another stable chemical branch which lowers the free energy by an increase in the mixing entropy in the low density phase and increasing the interaction among neutral macromolecules in the high density phase. One interesting feature of the phase diagrams, is that when the neutral-neutral interactions become dominant, i.e., the phase behaviour at the isoelectric point is mainly driven by the interaction among neutral macromolecules, we observe the vanishing of the critical points, giving rise to phase diagrams which only have first order transitions. The dominant interaction thus defines the topology of the phase diagram as a function of pH.

Finally, we study a more realistic scenario where the maximal net charge of the macromolecules m is chosen to be $m = 50$, which is close to the net maximal charge of some proteins that respond to pH changes and that are found in the so called stress granules [30, 34][2]. For simplicity we only consider interactions between oppositely charged macromolecules, which in this case are given by $\chi_e/\epsilon = -\alpha z k_B T$, where z is the total number of charges of the macromolecule and α is a factor describing the contribution of each fixed charge of the macromolecules to the interaction (we use the value $\alpha = 7.5$ as

---

[2]The proteins Sup35 and Pab1, which respond to pH changes have net charges within this range and the total number of charges is approximately as high as 184 for Sup35.
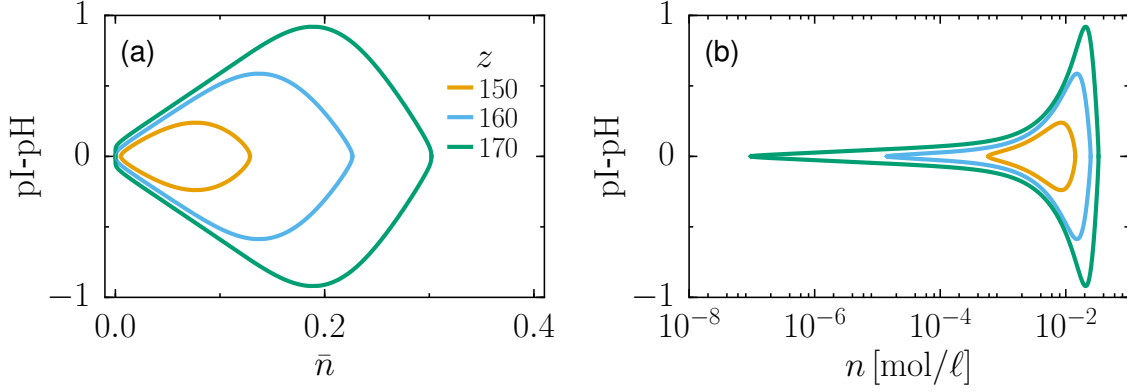
FIGURE 4.5: Binodal lines for different choices of the total number of charges on the macro-molecules z with interactions $\chi_e/k_B T = -\alpha z \epsilon$. The shaded region within the binodals is the region where macromolecules undergo a demixing transition, whereas the region outside is where the system remains homogeneously mixed. (a) Phase diagram as a function of the total macro-molecule volume fraction $\bar{n}$ and deviations from the isoelectric point $pH - pI$. (b) Same diagram as in (a) but as a function of the total macromolecule molar concentration n. Parameters $\epsilon = 0.002$, $h_\phi/k_B T = 10$ and $\alpha = 7.5$, apply to both panels.

reported in [99]). We study the system for three different values of the total number of charges in the macromolecule z and choosing $\epsilon = 0.002$ (Fig. 4.5) motivated by a volume ratio of water molecules and a typical protein. For all values of z considered, there exists a broad region of phase separation at the isoelectric point (Fig. 4.5). The coexistence becomes broader for increasing values of z due to an increase in the interaction strength. Our minimal model predicts that at the isoelectric point, a mixture of macromolecules with a large total number of charges, will phase separate over a large concentration range. Finally, we also show that reducing z can lead to a drastic reduction of the concentration range in which the system undergoes phase separation (Fig. 4.5(b)). So far we have focused on systems where one can impose the value of the pH, which remains unchanged for if the system undergoes phase separation. In the following, we briefly discuss the consequences of explicitly considering electrically neutrality in the system on the pH values in the coexisting phases.

4.7 Electric potential and pH differences across phases

In the previous sections we discussed the effect of pH on phase separation. In particular, we studied systems in which the pH is externally imposed and does not change for a system undergoing phase separation. These conditions would be in agreement with the idea that biomolecular condensates inside of the cell are not able to maintain a pH gradient across their interface due to the lack of a membrane [3]. In this section however, we show in very general terms that a natural consequence of considering systems that are electrically neutral, is the emergence of an electric potential difference across the interfaces

of liquid coexisting phases [100, 101]. This electric potential difference is concomitant with a difference of pH across phases (the chemical potential of the proton differs in different phases). It is well known that different membrane-bound compartments like mitochondria, Golgi-network, nucleus and lysosome among others [2, 102–104], have very different pH values which correlate with their functions. In contrast, this has not been studied nor suggested in protein condensates, we therefore believe that developing an understanding of possible electric potential and pH differences between protein condensates and the cytosol might help unraveling their function.

A system at thermodynamic equilibrium is considered to be electrically neutral, otherwise, there would be a current caused by the net charge in the system. The condition of zero net charge in the system is called the charge neutrality condition [105, 106]. In this section, we answer show how charge neutrality modifies the coexistence conditions. In order to so, let us consider a system at fixed temperature T and pressure P, composed of M different components with particle numbers $\{N_1, \ldots, N_M\}$. Each of these components carry a charge z_i ($z_i = 0$ is included), hence, the charge neutrality condition can be expressed as

$$\sum_{i=1}^{M} z_i N_i = 0 \quad . \tag{4.48}$$

We now proceed in a similar way as in Chapter 1 and 3 to derive the coexistence conditions. If the system splits in two phases, we need to find the conditions for a minimum of the Gibbs free energy $G(T, P, \{N_i\})$ subject to the constraints of charge neutrality in each of the phases and the conservation of the particle number of each of the different components. This can be done by minimizing the following functional

$$\mathcal{G} = G^{\mathrm{I}} + G^{\mathrm{II}} + \sum_{i=1}^{M} \mu_i^{\mathrm{el}} \left(N_i - N_i^{\mathrm{I}} - N_i^{\mathrm{II}} \right) + \Phi^{\mathrm{I}} \sum_{i=1}^{M} z_i N_i^{\mathrm{I}} + \Phi^{\mathrm{II}} \sum_{i=1}^{M} z_i N_i^{\mathrm{II}} \quad , \tag{4.49}$$

where G^{I} and G^{II} are the Gibbs free energies contributions from the two different phases, N_i is the total number of particles of each component, N_i^{I} and N_i^{II} are the particle numbers in each phase, the electric potentials Φ^{I} and Φ^{II} are Lagrange multipliers enforcing the charge neutrality condition in each of the phases[3], and μ_i^{el} is the electrochemical potential [95, 107] which enforces the particle number conservation. In order to find the minimum, we derive the functional 4.49 with respect to the particle numbers in each of the phases and

---

[3]The electric potential Φ in each of the different phases is known as Donnan potential [101, 106] and is usually introduced when discussing the equilibrium conditions for charge carriers (mobile ions) for a system divided in two parts by a semi-permeable membrane which stops large components moving from one side to the other. Equilibrium is achieved when the chemical potential of the mobile charge carriers is equal in both sides, the different partitioning of these mobile ions causes the electric potential difference and the equilibrium conditions are known as Donnan equilibrium

equate to zero, leading to

$$\frac{\partial \mathcal{G}}{\partial N_i^{\mathrm{I}}} = \mu_i^{\mathrm{I}} - \mu_i^{\mathrm{el}} + z_i \Phi^{\mathrm{I}} = 0 \quad , \quad \text{for } i \in \{1, \ldots, M\} \quad , \tag{4.50}$$

$$\frac{\partial \mathcal{G}}{\partial N_i^{\mathrm{II}}} = \mu_i^{\mathrm{II}} - \mu_i^{\mathrm{el}} + z_i \Phi^{\mathrm{II}} = 0 \quad , \quad \text{for } i \in \{1, \ldots, M\} \quad , \tag{4.51}$$

$$\frac{\partial \mathcal{G}}{\partial \Phi^{\mathrm{I}}} = \sum_{i=1}^{M} z_i N_i^{\mathrm{I}} = 0 \quad , \tag{4.52}$$

$$\frac{\partial \mathcal{G}}{\partial \Phi^{\mathrm{II}}} = \sum_{i=1}^{M} z_i N_i^{\mathrm{II}} = 0 \quad , \tag{4.53}$$

where we used $\partial G^{\mathrm{I}}/\partial N_i^{\mathrm{I}} = \mu_i^{\mathrm{I}}$ and $\partial G^{\mathrm{II}}/\partial N_i^{\mathrm{II}} = \mu_i^{\mathrm{II}}$. We did not include the derivatives with respect to the electrochemical potential since it only states the conservation of each particle number. From Eq. (4.50) and Eq. (4.51), we obtain the condition

$$\mu_i^{\mathrm{el}} = \mu_i^{\mathrm{I}} + z_i \Phi^{\mathrm{I}} = \mu_i^{\mathrm{II}} + z_i \Phi^{\mathrm{II}} \quad , \quad \text{for } i \in \{1, \ldots, M\} \quad , \tag{4.54}$$

where we see that the conditions for a minimum of $\mathcal{G}$ amount to the equality of electrochemical potentials in both phases apart from the charge neutrality conditions given in Eqs. (4.52) and (4.53). Furthermore, each phase has to obey charge neutrality as shown in Eqs. (4.52) and (4.53), this means that we can arbitrarily choose one component in each phase and express its particle number as a function of the rest of the particle numbers corresponding to components carrying a charge. This choice is arbitrary so let us choose the k-th component, with $z_k \neq 0$, in each of the phases and express it as

$$N_k^{\mathrm{I}} = -\sum_{i \neq k} \frac{z_i N_i^{\mathrm{I}}}{z_k} \quad , \tag{4.55}$$

$$N_k^{\mathrm{II}} = -\sum_{i \neq k} \frac{z_i N_i^{\mathrm{II}}}{z_k} \quad , \tag{4.56}$$

which explicitly shows that there are only $M-1$ independent particle numbers in each of the phases. If we now use the k-th condition for the equality of electrochemical potentials and express the electric potential difference across coexisting phases $\Delta\Phi = \Phi^{\mathrm{II}} - \Phi^{\mathrm{I}}$ by

$$\Delta\Phi = \frac{\mu_k^{\mathrm{I}} - \mu_k^{\mathrm{II}}}{z_k} \quad , \tag{4.57}$$

we can rewrite the electrochemical potential conditions given in Eq. (4.54) for the rest of the components, which read

$$\mu_i^{\mathrm{I}} - \frac{z_i}{z_k} \mu_k^{\mathrm{I}} = \mu_i^{\mathrm{II}} - \frac{z_i}{z_k} \mu_k^{\mathrm{II}} \quad , \quad \text{for } i \in \{1, \ldots, M\} \text{ and } i \neq k \quad . \tag{4.58}$$

This condition defines the equality of some type of relative chemical potential enforcing the charge neutrality condition when there are exchanges of particles from one phase to the other. Please note that we eliminated the electric potential difference from the description by using Eq. (4.57) and that we also eliminated the particle number N_k from the description, thus there are now $M - 1$ coexistence conditions defined by Eq. (4.58) which can be solved to find the compositions in both phases. After solving these conditions and finding the equilibrium compositions given by the particle numbers N_i^{I} and N_i^{II} with $i \neq k$, these particle numbers can be substituted in the charge neutrality conditions to find the particle number N_k in both phases and also to compute the electric potential difference using Eq. (4.57). These steps show that finding the explicit value of the electric potential is not needed in order to find the compositions of the coexisting phases, nevertheless, it is worth knowing how to compute its value since it might be a relevant quantity determining the partitioning of ions and other charged molecules partition into the different phases [100]. Besides the existence of the electric potential difference, one very interesting result that has come naturally from our derivation, is that for charged components, there is a chemical potential difference between the two phases, which as we show below, for the case of hydrogen ions, amounts to a difference in pH across the phases.

Let us consider the same multicomponent system as in Eq. (4.49), with one of the components being hydronium ions, denoted by H_3O^+. The coexistence conditions for H_3O^+ would be given by

$$\mu_{H_3O^+}^{\mathrm{II}} - \mu_{H_3O^+}^{\mathrm{I}} = -e\Delta\Phi \quad , \tag{4.59}$$

where e denotes the elementary charge. Considering equilibrium of water across the two phases, using the pH definition given in Eq. (4.19) and making slight modifications of the equation, we can reformulate Eq. (4.59) as

$$\mathrm{pH}^{\mathrm{II}} - \mathrm{pH}^{\mathrm{I}} = -\frac{\Delta\Phi \log_{10} e}{k_{\mathrm{B}}T} \quad , \tag{4.60}$$

where we see that there is a direct connection between the difference in pH across the interface and the electric potential difference.

We decided to present the full derivation in order to show how by consistently adding the constraints in the system reflect macroscopic properties of the coexisting phases that might be of relevance. If we think on protein condensates, what we discussed in this section clearly shows that by means of phase separation, the chemical and physical environment of protein condensates could be regulated by charges. Moreover, that biomolecular condensates might have well defined pH values that allows to perform specific tasks, but this remains to be further studied and be backed by experiments. To conclude, more consequences of electric potential and pH differences across interfaces remain to be explored in the future.

4.8 Summary

In this chapter, we have established a thermodynamic framework to study liquid-liquid phase separation in a system where the pH is controlled. We started by introducing chemical reactions controlling the charge states of macromolecules which are in turn determined by the pH value of the mixture. Using conservation laws for a system undergoing chemical reactions, we identified the effective thermodynamic conjugate variables at chemical equilibrium. We then found the relevant thermodynamic variables controlling the pH of the system, namely the chemical potential difference, $\mu_{\mathrm{H_3O^+}} - \mu_{\mathrm{H_2O}}$, and the temperature T. That allowed us to construct the corresponding thermodynamic potential for a system with a fixed pH value by means of a Legendre transform which makes, $\mu_{\mathrm{H_3O^+}} - \mu_{\mathrm{H_2O}}$, a natural variable of the corresponding free energy. Based on this thermodynamic potential we showed how the chemical and phase equilibrium are controlled by the pH of the system and calculated the corresponding phase diagrams.

We found that phase separation typically occurs around pH values corresponding to the isoelectric point pI. This finding could be relevant to processes in living cells since many cytosolic proteins have a similar isoelectric point. For example in yeast cells, many proteins exhibit isoelectric points around pH $\simeq$ 5.3 and pH $\simeq$ 9 [16]. Thus, our finding of phase separation around the pI is consistent with observations in yeast cells [16], where many proteins separate from the cytosol once the pH is lowered to pH $\simeq$ 5. Using typical parameters for cellular proteins (Fig. 4.5), we find that phase separation occurs from concentrations ranging from μM corresponding to a typical saturation concentration of phase separating cellular proteins [28, 30] to mM. Our model is also in agreement with phase separation for a concentration of the order of $1\,\mathrm{mM}$, which is an estimation of the total concentration of proteins in yeast [108]. Note that phase separation is lost, both at low and high concentrations, which can be understood from the analogy with the quasi-binary mixture shown in Fig. (4.2)(a,e). We further predict that upon decreasing pH even more, reentrant behaviour leading to a mixed state will be observed. However, this behaviour may not always be observable in living systems due to several reasons. One of them is that upon decreasing the pH below the isoelectric point, proteins might denature and aggregate irreversibly before reaching conditions for phase separation that we describe in this manuscript. Our approach considers phase separation controlled by pH at thermodynamic equilibrium. Neglecting the effect on phase separation related to the consumption of ATP is reasonable since the reported starvation induced phase separation occurs upon ATP depletion conditions [16]. We concluded by exploring the consequences of considering the charge neutrality condition in the system, and showed that there may exist a pH and electric potential difference across interfaces of different phases. This could be relevant in the context of biomolecular condensates that form via a demixing transition in which charged macromolecules are involved.

Chapter 5

Protein Phase Separation Regulated by RNA Concentration *in vitro* and *in vivo*

Many different membraneless organelles, such as P granules [12], nucleoli [19], Cajal bodies [109] and stress granules [21], are mainly composed of RNA-binding proteins and RNA. It has been shown that RNA plays a prominent role in protein phase separation *in vitro* and *in vivo* [3, 18, 22, 23, 25, 27–29, 36–38, 52], thus having an understanding of the influence of RNA in protein phase separation is highly relevant for understanding the formation of biomolecular condensates. In this chapter, we are concerned with RNA as a regulator of protein phase separation, in particular, we study the role of RNA on the positioning of P granules in the *C. elegans* embryo. For presentation purposes, let us briefly describe again the pioneering work of Brangwynne et al. [12]. In that work, it was shown that P granules are liquid-like droplets that localize to the posterior of the *C. elegans* embryo by means of a dissolution/condensation mechanism. It was further shown that a MEX-5 protein (hereafter MEX-5) gradient is established prior to the first cell division and that this gradient anti-correlates with the positioning of P granules, i.e. higher concentrations of MEX-5 at the anterior side of the embryo are connected to dissolution of P granules, whereas P granules condense at the posterior side which has lower concentrations of MEX-5 [110, 111]. Although this anti-correlation was known, the mechanism driving the positioning remained elusive. One step towards describing a possible mechanism for P granule segregation [112] was based on a minimal reaction diffusion model for two components. There, the authors described effective components which are divided in P granule constituents and non-constituents. In their model, components can diffuse and have spatially dependent production and degradation rates. They showed that there is an accumulation of the constituent material of droplets at the posterior side but they did not discuss the underlying thermodynamic mechanism in charge of regulating the locally dependent phase

separation.

In previous work [28], we and our experimental collaborators proposed a mechanism that can account for the positioning of P granules on a thermodynamic basis. This mechanism is based on an mRNA-binding competition between constituent proteins of P granules and the MEX-5 protein. The hypothesis was that a MEX-5 gradient can regulate the amount of RNA available for binding across the *C. elegans* embryo, which in turn tunes the phase separation tendency of P granule constituents. To test this hypothesis, our collaborators purified PGL-3 protein, a key component of P granules, and studied its ability to form liquid-like drops. In addition, they also performed experiments to study the effect of RNA on PGL-3 phase separation as well as the effect of having both MEX-5 and RNA simultaneously present in a PGL-3 solution. We then used a model to explain the experimental observations and showed that a positioning mechanism based on an mRNA binding competition between MEX-5 and PGL-3 is a feasible explanation for P granule positioning.

The content of this chapter is largely based in our publication [28] and is organized as follows: we first present the experimental evidence showing that the PGL-3 protein forms liquid-like drops. We then show experimental results from our collaborators regarding the RNA-binding affinities of both, PGL-3 and MEX-5. We close this part by showing that mRNA regulates the saturation concentration of PGL-3 and that MEX-5 can decrease the tendency to phase separate of PGL-3 in the presence of mRNA. In order to explain the experimental observations, we propose a model describing PGL-3 phase separation in addition to including mRNA and MEX-5 in the description as regulating components. Using this model we then fit the experiments performed by our collaborators and show that a minimal mean field model can account for the observed phase separation in the *in vitro* experiments. We end by showing solutions to the dynamical equations for the concentrations of all components corresponding to the free energy of the full thermodynamic model, and showed that indeed, the mRNA binding competition between MEX-5 and PGL-3 accounts for the positioning of liquid phase-separated drops *in vivo*.

5.1 P granule constituent protein forms liquid droplets *in vitro*

It had previously been reported that two key components of P granules undergo phase separation *in vitro* [36], namely PGL-1 and PGL-3. Our collaborators, purified the proteins PGL-3 and PGL-1, tagged with a green fluorescent protein (mEGFP) from insect cells and studied solutions of PGL-3 or PGL-1 immersed in buffer (a mixture that resembles the physiological conditions of the cytosol). They showed that PGL-1 did not undergo phase separation for physiological salt conditions in the buffer, whereas PGL-3 forms liquid-like drops, see Fig. 5.1. Liquid phase separation was assessed by fluorescence microscopy of
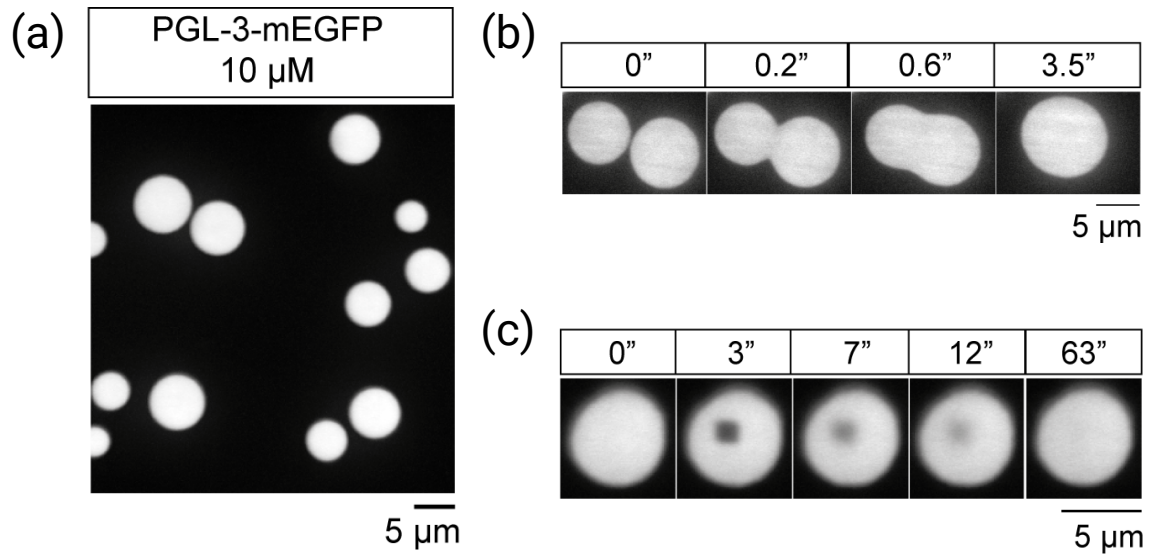
FIGURE 5.1: PGL-3 forms liquid-like drops. (a) Maximum intensity projection of the PGL-3-mEGFP solution. PGL-3 forms spherical droplets with a higher concentration than in solution. (b) Fusion event of two spherical droplets showing the fluid behavior of the phase-separated drops. (c) Fluorescence recovery is observed after photo-bleaching a small region in the interior of the drop, which is evidence of dynamic rearrangement of particles in the interior of the drops. PGL-3-mEGFP refers to the tagged PGL-3 protein with the green fluorescent protein mEGFP. All panels are adapted from [28].

GFP:tagged PGL-3, collecting the fluorescence obtained from the mEGFP tag to PGL-3 proteins. The evidence showing that PGL-3 forms liquid-like drops is the following: first, the condensed phase is constituted by spherical droplets shown in Fig. 5.1(a), second, the spherical drops undergo fusion on a short time scale shown in the sequence of images in Fig. 5.1(b) and third, fluorescence recovers after photo-bleaching a small area in the center of the PGL-3 liquid drops, indicating a constant exchange of matter with the surroundings, shown in Fig. 5.1(c). Showing that PGL-3 forms liquid-drops *in vitro*, similar to P granules *in vivo*, serves as the first step towards understanding the mechanism of P granule positioning in the *C. elegans* embryo.

5.2   Quantification of Protein-RNA binding

Protein binding to RNA plays a prominent role in phase separation processes in the cell, here we discuss how RNA interacts with two key players in P granule positioning, PGL-3 and MEX-5. In order to do so, we discuss the experimental results of the RNA-binding assays of PGL-3 and MEX-5, shown in Fig. 5.2. Experimental measurements show that PGL-3 binds weakly to RNA, characterized by a dissociation constant $K_{PR} \simeq 230\,\text{nM}$, see
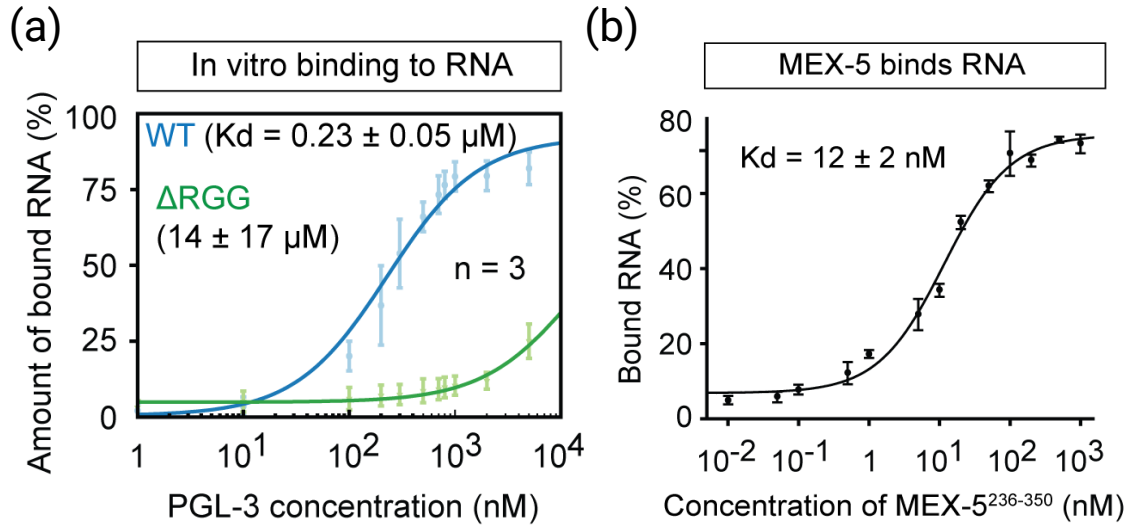
FIGURE 5.2: RNA-binding assays of PGL-3 and MEX-5. (a) RNA-binding curves showing that PGL-3 binds to RNA. For the wild-type PGL-3, the blue line shows a weak binding affinity, which is mediated via six RGG repeats in the amino-acid sequence of PGL-3. The green line shows an extremely reduced binding affinity for a PGL-3 mutant that lacks the six RGG repeats. (b) The MEX-5 binding curve shows a strong binding affinity between MEX-5 and RNA which is mediated by two zinc Fingers. The dissociation constants in each panel are denoted by Kd. Both panels are adapted from [28].

blue line in Fig. 5.2(a). This binding is mediated by the six RGG [1] repeats of PGL-3 [28]. This is concluded after observing that a mutant of PGL-3 lacking the RGG repeats, shows an extremely reduced binding affinity to RNA, see green line in Fig. 5.2(a). Interestingly, experimental measurements show that MEX-5 binds 20 times stronger to RNA than PGL-3, characterized by a dissociation constant $K_{\mathrm{MR}} \simeq 12\,\mathrm{nM}$. The binding occurs via two zinc finger domains in the MEX-5 protein[2]. In what follows, we discuss experiments assessing the regulatory effect of MEX-5 and RNA in PGL-3 phase separation.

5.3   Regulation of protein phase separation *in vitro*

In this section we discuss experimental results showing the regulating effect of RNA and MEX-5 on PGL-3 phase separation. In [28] it was shown RNA promotes phase separation, in particular, that mRNA promotes phase separation, whereas rRNA, due to its structured nature does not (data not shown). Hereafter we only consider mRNA in our discussion.

The experiments we discuss were done *in vitro* using fluorescence microscopy, the images shown in Fig. 5.3(a,b,c) are taken after some time of equilibration. They show maximum

---

[1]The RGG repeat is characterized by the combination arginine-glycine-glycine in the amino acid sequence of a protein. This repeat is known to be important in mediating binding interactions between proteins and RNA. For detailed information about the RGG repeat we refer the reader to [113].

[2]In fact, the full form of MEX-5 could not be purified and instead a shorter sequence which contains two zinc fingers was used. These zinc fingers are known to strongly interact with RNA and DNA [114]

intensity projections of the different image stacks in the z direction (considering the image-plane to be $x - y$), white circles show a high density of PGL-3. In the absence of mRNA, the saturation concentration of PGL-3 is approximately $2\,\mu$M, shown in the third panel in Fig. 5.3(a). At concentrations larger than $2\,\mu$M the protein condensed phase (white circles) grow, as expected from a simple picture of phase separation in binary mixtures. For PGL-3 solutions in the presence of mRNA at a concentration of $100\,$nM (shown in Fig. 5.3(b)), we observe a significant formation of tiny drops for concentrations of PGL-3 of $100\,$nM. This then shows that mRNA regulates PGL-3 phase separation by increasing its tendency to phase separate. Similar experiments to assess the effect of MEX-5 on PGL-3 phase separation were done. A solution with a PGL-3 concentration of $0.6\,\mu$M was analyzed in four different conditions, see Fig. 5.3(c). The PGL-3 solution was analyzed with and without MEX-5, showing no sign of PGL-3 phase separation, shown in the two upper panels in Fig. 5.3(c). In contrast, when mRNA is added to the PGL-3 solution, shown in the two lower panels in Fig. 5.3(c), MEX-5 is shown to inhibit the phase separation promoted by mRNA, this is seen by comparing the lower left panel (no MEX-5) with the lower right panel (with MEX-5). To further study the effects of mRNA and MEX-5 on PGL-3 phase separation, the total GFP fluorescence in drops was measured for different conditions, see Fig. 5.3(d). It is shown that a PGL-3 concentration of $0.6\,\mu$M remains well mixed with buffer when MEX-5 is added (first triangle from left to right). In contrast, if MEX-5 is absent and mRNA is added to the PGL-3 solution, 25% of the fluorescence is concentrated in the PGL-3 condensed phase (second triangle from left to right). If both mRNA and MEX-5 are present in the PGL-3 solution (last three triangles from left to right), the total fluorescence in the PGL-3 condensed phase increases for increasing amounts of mRNA.

In summary, the presence of mRNA increases the tendency of PGL-3 to phase-separate whereas MEX-5 does not seem to have a direct effect on PGL-3 phase behavior in the absence of mRNA. Interestingly, when both MEX-5 and mRNA are present in a PGL-3 solution, MEX-5 reduces the effect of mRNA on promoting PGL-3 phase separation. This is in agreement with MEX-5 having a stronger binding affinity to mRNA than PGL-3. All the previously explained experimental observations set the basis to formulate an mRNA binding competition mechanism as a plausible explanation for P granule positioning in the *C. elegans* embryo. In what follows, we back this hypothesis by building up a model to explain PGL-3 phase separation and its regulation by mRNA and MEX-5 based on the experimental evidence.

5.4   Theory of regulation of protein phase separation

In this section, we introduce a model to describe the ability of PGL-3 to phase separate and discuss the role played by mRNA and MEX-5 in regulating this behavior. We begin our discussion by describing the system using a Flory-Huggins free energy, where the
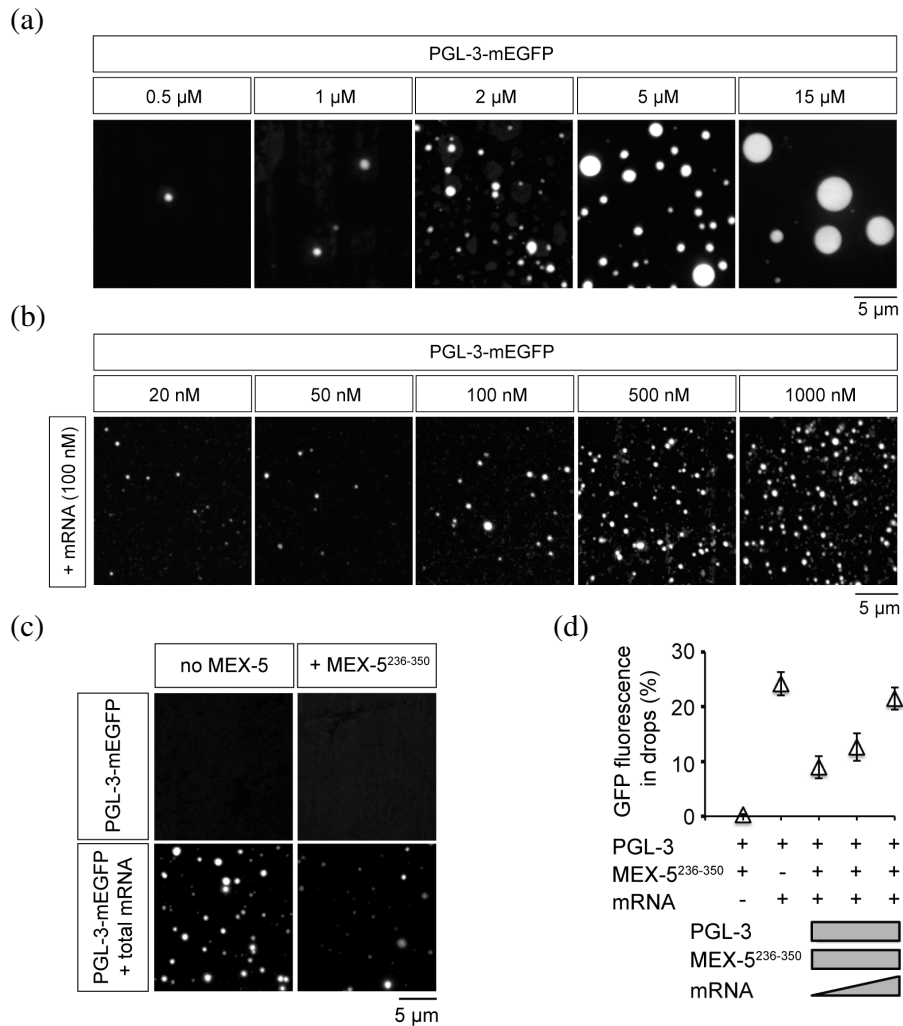
FIGURE 5.3: PGL-3 phase separation regulated by mRNA and MEX-5. (a,b) Maximum intensity projections of confocal images in the z direction of PGL-3 solutions. (a) Different concentrations of PGL-3 tagged with green fluorescent protein exhibiting phase separation. (b) Different concentrations of PGL-3 tagged with green fluorescent protein in the presence of 100 nM mRNA exhibiting phase separation at low PGL-3 concentrations. (c) Assessment of MEX-5 and mRNA effects on a solution with PGL-3 concentration of $0.6\,\mu$M. In the left upper panel, PGL-3 is well-mixed. In the right upper panel, MEX-5 is added and there is no effect on PGL-3 phase separation. In the lower left panel mRNA is added and shown to promote phase separation. In the lower right panel, MEX-5 is shown to reduce the effect of mRNA promoting PGL-3 phase separation in buffer. (d) Total fluorescence inside of the condensed phase of a $0.6\,\mu$M PGL-3 solution with different concentrations of MEX-5 and mRNA. The $+/-$ symbols indicate the presence/absence of specific components in the mixture. The last three symbols from left to right, show the interplay between MEX-5 and mRNA on regulating PGL-3 phase separation. The MEX-5 concentration is kept constant at 150 nM with varying mRNA concentrations (from left to right) 50, 100, or 150 ng/μl. These three points show an increase in the fluorescence in the condensed phase for increasing values of mRNA. All panels are adapted from [28].

components are given by: the phase-separating components which are PGL-3, denoted by P, and the complex PGL-3:mRNA, denoted by PR. The regulating components, MEX-5, denoted by M, and mRNA, denoted by R. Finally, the solvent is given by W and the complex formed by MEX-5 and mRNA is denoted by MR. The free energy density reads

$$f = \frac{k_{\mathrm{B}}T}{v_W} \left[\sum_i \epsilon_i \phi_i \left(\ln \phi_i + w_i/k_{\mathrm{B}}T \right) + \chi_{P,W} \phi_P \phi_W + \chi_{PR,W} \phi_{PR} \phi_W \right] \quad , \qquad (5.1)$$

where ϕ_i stands for the volume fraction, $\epsilon_i = v_W/v_i$ is the molar[3] volume ratio between the solvent W and the component i, w_i is the free energy per particle due to internal degrees of freedom and the Flory-Huggins parameters $\chi_{P,W}$ and $\chi_{PR,W}$ describe the tendency to phase separate of PGL-3 and PGL-3:mRNA, respectively and the index i runs over all components, including the solvent W. Our choice of Flory-Huggins parameters is consistent with the fact that MEX-5 and mRNA do not show any tendency to phase separate. For compactness, we used ϕ_W to denote the solvent component in Eq. (5.1) instead of the explicit form $\phi_W = 1 - \sum_{j \neq W} \phi_j$, in terms of the rest of the components. The volume fractions are connected to the concentrations via, $\phi_i = v_i n_i$, where n_i is the concentration.

We have presented experimental evidence showing that PGL-3 and MEX-5 bind to mRNA. We now propose a set of binding reactions to account for the observed regulation of the phase behavior of the PGL-3 protein. We express these mRNA-binding reactions of PGL-3 and MEX-5 as

$$P + R \ \rightleftharpoons \ PR \quad , \qquad (5.2a)$$

$$M + R \ \rightleftharpoons \ MR \quad . \qquad (5.2b)$$

These reactions describe how PGL-3:mRNA is formed by a PGL-3 protein binding to an mRNA molecule and how the complex MEX-5:mRNA forms when a MEX-5 protein binds to an mRNA molecule.

In order to simplify our discussion of binding equilibrium, we consider that binding is mainly driven by differences in the internal free energies of the molecules and neglect the effect of molecular interactions on the binding constants. More precisely, we only consider this approximation in the case of the binding reaction of PGL-3 to mRNA, because, as we show in Appendix H, the binding equilibrium of MEX-5 and mRNA remains unchanged because they do not have a tendency to phase separate. We then express the binding

---

[3]We use molar volume since the concentrations are given in molar units M = mol/l. One can transform from molar to molecular volume using Avogadro constant.

equilibrium as

$$\frac{n_P \, n_R}{n_{PR}} \simeq K_{PR} \quad , \tag{5.3}$$

$$\frac{n_M \, n_R}{n_{MR}} \simeq K_{MR} \quad , \tag{5.4}$$

where K_{PR} is the dissociation constant of the complex PGL-3:mRNA and K_{MR} is the dissociation constant of MEX-5:mRNA. A thorough discussion of the binding equilibrium conditions is presented in Appendix H.

Since we are interested in studying how PGL-3 and the complex PGL-3:mRNA undergo phase separation, we now use Eqs. (5.3) and (5.4) to find a relation between their concentrations, n_P and n_{PR}. This relation is given by

$$n_{PR}(n_P) \simeq \frac{\sqrt{\left[K_{PR}n_M^T+\left(1+\frac{K_{PR}}{n_P}\right)K_{MR}n_P-K_{PR}n_R^T\right]^2+4K_{MR}K_{PR}n_R^Tn_P\left(1+\frac{K_{PR}}{n_P}\right)}}{2\left(1+\frac{K_{PR}}{n_P}\right)K_{PR}}$$
$$+\frac{K_{PR}n_R^T-K_{PR}n_M^T-\left(1+\frac{K_{PR}}{n_P}\right)K_{MR}n_P}{2\left(1+\frac{K_{PR}}{n_P}\right)K_{PR}} \quad , \tag{5.5}$$

where $n_R^{\mathrm{T}} = n_R + n_{PR} + n_{MR}$ is the total mRNA concentration and $n_M^{\mathrm{T}} = n_M + n_{MR}$ is the total concentration of MEX-5. For fixed values of the total amounts of PGL-3, MEX-5 and mRNA, the binding reactions constrain the possible values of n_{PR} and n_P, thus we will refer to Eq. (5.5) as the constrained path.

Please note that in the absence of mRNA, there is no relation connecting all the components. Furthermore, since we are interested in explaining some experiments where MEX-5 is absent, we also calculate the constrained path for a system where only PGL-3 and mRNA are considered, for which Eq. 5.5 simplifies to

$$n_{PR}(n_P) \simeq \frac{n_R^{\mathrm{T}}}{1 + \frac{K_{PR}}{n_P}} \quad . \tag{5.6}$$

In Fig. 5.4, we show the constrained path given in Eq. (5.6) for the values $K_{PR} = 250\,\mathrm{nM}$ and $n_R^{\mathrm{T}} = 100\,\mathrm{nM}$. We see that in this case for a total concentration of PGL-3, $n_P^{\mathrm{T}} \simeq 1\,\mu\mathrm{M}$, most of the mRNA is already bound and n_{PR} has reached its maximum concentration.

Having discussed the model and the binding equilibrium of the system, we are now in a position to fit the theory to the experimental data.

5.5   Comparison of theory to experiments

In this section we use the free energy density given in Eq. (5.1) to fit the experimental data for two different scenarios. The first scenario is a PGL-3 solution in the absence of mRNA and MEX-5, this scenario corresponds to the experimental situation shown in
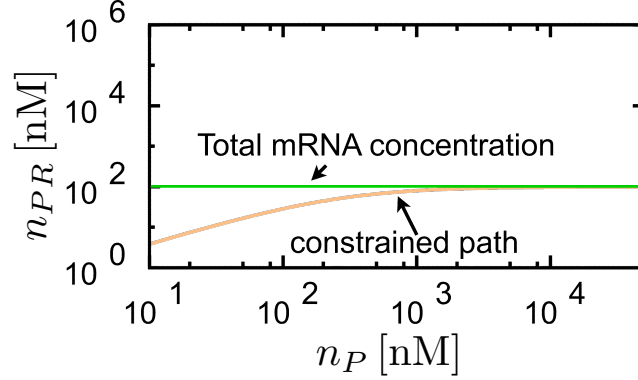
FIGURE 5.4: Constrained path (yellow line) showing the concentration of the complex PGL-3:mRNA, n_{PR} as a function of the free PGL-3 concentration, n_P. The total mRNA concentration (green line) is $n_R^T = 100$ nM and the dissociation constant is $K_{PR} = 230$ nM.

Fig. 5.3(a). The second one, is the solution for which mRNA is added at concentration $n_R = 100$ nM, corresponding to the experimental data acquired from Fig. 5.3(b). From the fits, we extract the molecular volume ratios, $1/\epsilon_P = v_P/v_W$ and $1/\epsilon_{PR} = v_{PR}/v_W$, and the Flory-Huggins parameters $\chi_{P,W}$ and $\chi_{PR,W}$.

We begin by studying the case of PGL-3 phase-separating from the solvent in the absence of mRNA. In this case, the system is described as a binary mixture with components P and W. From the experimental data (see Fig. 5.3(a)) we obtain the fluorescence density in the different phases, we connect these measurements to PGL-3 concentrations by considering a linear relation between fluorescence and concentration. We transform the fluorescence density inside and outside of the drops to concentration values using the following relations:

$$\mathrm{I}^{\mathrm{in}} = \mathrm{I}_0 n_P^{\mathrm{in}}, \tag{5.7}$$

$$\mathrm{I}^{\mathrm{out}} = \mathrm{I}_0 n_P^{\mathrm{out}}. \tag{5.8}$$

where I_0 is a constant relating fluorescence and PGL-3 concentration, n_P^{in} and n_P^{out} are total PGL-3 concentrations in the condensed phase (inside the drops) and in the dilute phase (outside the drops), respectively, and the fluorescence concentrations in the condensed and in the dilute phase are given by I^{in} and $\mathrm{I}^{\mathrm{out}}$, respectively. For this case, $n_{PR} = 0$, therefore the total PGL-3 concentration is given by $n_P^T = n_P$.

The quantities that we fit are the saturation concentration of PGL-3 and the ratio between the fluorescence densities in the drops and outside of the drops. The saturation concentration value obtained from the experiments is given by $n_P^{\mathrm{s}} \simeq 2\,\mu\mathrm{M}$ and the fluorescence ratio is $\mathrm{I}^{\mathrm{in}}/\mathrm{I}^{\mathrm{out}} \simeq 50$. We obtain this fluorescence ratio by averaging over the fluorescence values corresponding to the last five PGL-3 fluorescence concentrations obtained experimentally (see the blue squares in Fig. 5.5(b)). We only considered these

points since the concentrations close to the saturation concentration did not seem to have equilibrated yet at the time of measurement. Moreover, we do not need to consider the temperature dependence of the Flory-Huggins parameters since the temperature of the experiments is kept fixed at room temperature.

A sketch of the phase diagram obtained for the fitting procedure is shown in Fig. 5.5(a). The PGL-3 concentrations used in the experimental setup lie on the blue dotted line, which is shown at a fixed value of $\chi_{P,W}$ because the temperature is fixed, thus $\chi_{P,W}$ remains the same for all experiments in this system. The fitting routine works as follows we vary the values of ϵ_P and $\chi_{P,W}$ and construct the corresponding coexisting phases. We then compare the concentration of the low density phase with the saturation concentration obtained from the experiments and the ratio between the concentrations in both phases with the fluorescence ratio measured experimentally, setting a tolerance of less than one percent. We repeat the procedure until the solution is within this tolerance. From the fit we extract $1/\epsilon_P = 1.48 \times 10^4$ and $\chi_{P,W} = 0.512$. This very large value of $1/\epsilon_P$ is obtained because at very low PGL-3 concentrations the system readily phase separates. The value of the Flory-Huggins parameter $\chi_{P,W}$ is relatively close to the critical value, which for large macromolecules is close to $1/2$. In Fig. 5.5(b) we show the fitted curve (blue line) to the experimental data (blue squares). Since we model the system as a binary mixture, it can only exhibit a step-like behavior which fits reasonably well to the last experimental data points.

Let us now discuss the case in which mRNA is also present in the PGL-3 solution, see Fig. 5.3(b). In this case, the free energy density given in Eq. (5.1) would also include the mixing entropy terms of the complex PGL-3:mRNA and of the free mRNA, in addition to the term proportional to the Flory-Huggins parameter $\chi_{PR,W}$, but in order to simplify our discussion we neglect the mixing entropy of the mRNA. The reasons behind this choice are that mRNA has a very low concentration, a big fraction of it will be bound to PGL-3 and mRNA is a rather large macromolecule so its contribution to the mixing entropy is rather low. Therefore, the main role played by mRNA in our model is to change the effective properties of PGL-3 when creating a complex PGL-3:mRNA.

We now use the parameters $\chi_{P,W}$ and ϵ_P obtained from the previous fit, to discuss the effect of mRNA on PGL-3 phase separation. In order to do so, let us remember that due to the binding reaction between PGL-3 and mRNA, the possible concentrations of PGL-3, n_P, and that of the complex PGL-3:mRNA, n_{PR}, are constrained to the path shown in Fig. 5.4. If we now consider that the total PGL-3 concentration is $n_P^{\mathrm{T}} = n_P + n_{PR}$, then $\Delta \mathrm{I} = \mathrm{I}^{\mathrm{in}} - \mathrm{I}^{\mathrm{out}}$, is given by

$$\Delta \mathrm{I} = \mathrm{I}_0(n_P^{\mathrm{in}} + n_{PR}^{\mathrm{in}} - n_P^{\mathrm{out}} - n_{PR}^{\mathrm{out}}) \quad . \tag{5.9}$$
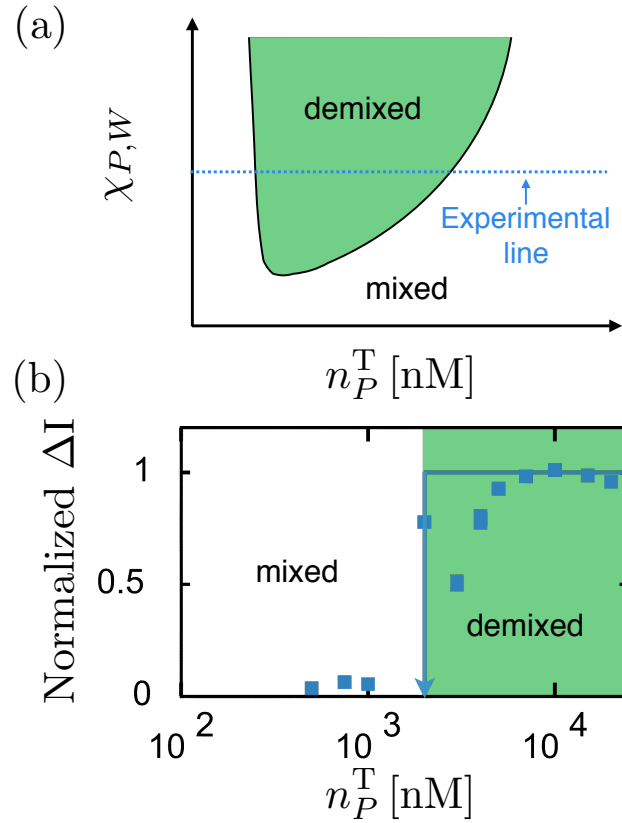
FIGURE 5.5: Binary mixture fit to the experimental data. (a) Sketch of the phase diagram of a binary mixture. The blue line represents the line where the experiments are performed (fixed $\chi_{P,W}$ and varying PGL-3 concentrations), the point where the binodal (black line) and the experimental line meet for the first time is the saturation concentration of PGL-3, this crossing defines the PGL-3 concentration of the dilute phase, the second crossing defines the concentration of the condensed phase. (b) Fitting curve (blue line) to the experimentally measured differences in fluorescence density (blue squares) between the two phases. The normalized ΔI is normalized by the average fluorescence density difference between the condensed and dilute phases (taken over the last five experimental points). The saturation concentration is indicated by an arrow in the blue solid line. Panel (b) is adapted from [28].

Because in the experiments PGL-3 is labeled independently of its binding to mRNA, we also considered the contributions to the fluorescence density by PGL-3:mRNA.

From the experimental data shown in Fig. 5.6(b), we see that the saturation concentration of total PGL-3 lies between 100 nM and 250 nM. In order to fit the data, we vary the parameters $\chi_{PR,W}$ and ϵ_{PR}, and construct binodals for which the first crossing point between the constrained path and the binodal lies within the previously mentioned concentration range. For the set of parameters that obey the previous condition, we calculate the concentration difference between the phases and compare with the experimental data. This is done by finding the tie line, connecting the condensed and dilute phases, which intersects the constrained path at the prescribed PGL-3 total concentration in the experiments, see Fig. 5.6(a).

From the fitting procedure we obtain $1/\epsilon_{PR} = 1.56 \times 10^4$, and $\chi_{PR,W} = 0.518$, which show that PGL-3:mRNA must have a larger tendency for phase separation than PGL-3 alone and also that the complex PGL-3:mRNA has a larger effective volume than PGL-3. One might wonder, why is the effective volume of the PGL-3:mRNA so close to the one of PGL-3 alone, this is because this volume corresponds to an effective interaction volume of the complex with the solvent, which we expect to be close to the one of PGL-3 alone since mRNA seems not to have a tendency for phase separation[4]. The results from the fitting procedure are shown in Fig. 5.6. In Fig. 5.6(a) we use the parameters obtained from fitting the experiments of both scenarios (the PGL-3 solution with and without mRNA) to construct the corresponding phase diagram. The first crossing between the binodal line (purple line) and the constrained path (yellow line) defines the saturation concentration of the total amount of PGL-3, $n_P^{\mathrm{T}} \simeq 200$ nM. We see that for compositions in the constrained path that lie near the binodal, the corresponding condensed phase is mostly composed of the complex PGL-3:mRNA, $n_{PR}^{\mathrm{in}} > n_P^{\mathrm{in}}$, whereas for increasing values of n_P^{T}, i.e. moving from left to right in the constrained path, the corresponding condensed phases start being mostly formed by free PGL-3 alone. We see in Fig. 5.6(b) that the fitted curve (orange line) correctly reproduces the decrease in the fluorescence density difference between phases obtained experimentally (orange circles), this results from a decreasing concentration of total PGL-3 in the condensed phase while moving from left to right along the constrained path.

In this section we have shown that our minimal model correctly accounts for the experimental observations regarding PGL-3 phase separation. Interestingly, we found that in order to reproduce the experimental data, the complex PGL-3:mRNA must have a stronger tendency to phase separate than PGL-3 and that its effective interaction volume must also be larger than that of PGL-3. Thus, we see that mRNA regulates phase separation by

---

[4]In fact, for large macromolecules the mixing entropy term is very small and many times negligible, thus the relevant volume to be considered, is an effective volume that takes into account the regions of the complex PGL-3:mRNA driving phase separation. This might the reason behind having such a similar value of ϵ_{PR} and ϵ_P.
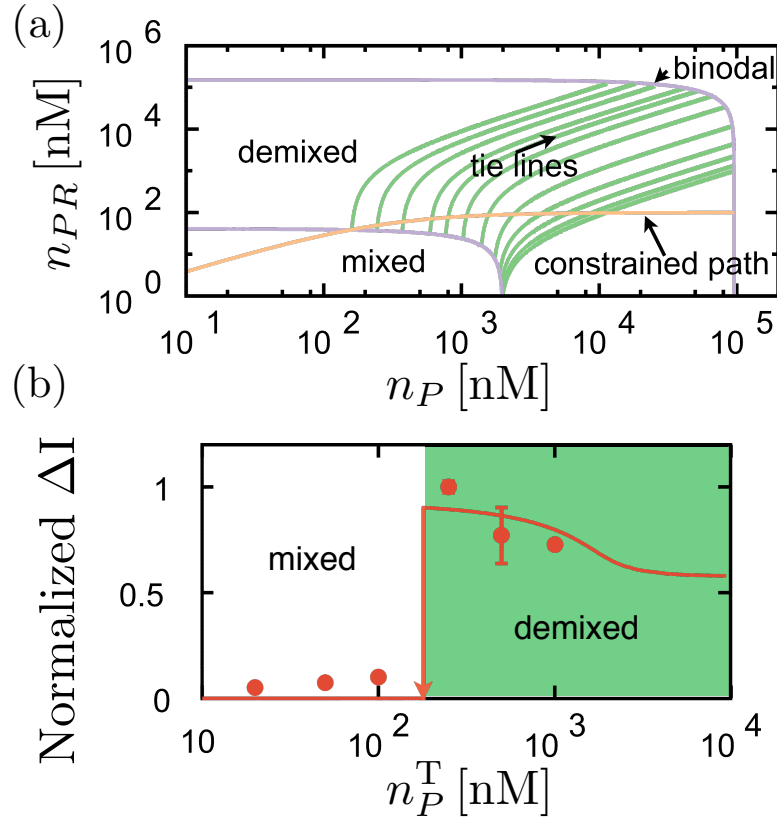
FIGURE 5.6: Fit for PGL-3 phase separation regulated by mRNA. (a) Phase diagram in the $n_P - n_{PR}$ composition plane, obtained using the parameters obtained from fitting the experimental results. The yellow line shows the constrained path given in Eq. (5.6), the purple line is the binodal and tie lines are shown as green lines. The first crossing point from left to right of the constrained path and the binodal defines the total PGL-3 saturation concentration, $n_P T \simeq 200 \, \text{nM}$. (b) Fit of the theory to the experimental data. The data from the experiments is shown as orange circles and the orange line is the curve calculated from the model. The green shaded region shows for which of the concentrations of total PGL-3 there will be phase separation. The parameters used are: $\chi_{P,W} = 0.512$, $\chi_{PR,W} = 0.518$, $1/\epsilon_P = 1.48 \times 10^4$, $1/\epsilon_{PR} = 1.56 \times 10^4$, $K_{PR} = 230 \, \text{nM}$ and $n_R^T = 100 \, \text{nM}$. Panels adapted from [28].

binding and increasing the effective interaction of PGL-3 in the complex. In the following, using the parameters obtained from the fits discussed in this section, we study how does MEX-5 interferes with this mechanism.

5.6   An mRNA-Binding competition regulates protein phase separation *in vitro*

Here, we briefly discuss the effect of MEX-5 in the regulation of phase separation. We do so, by using the parameters obtained from the fit of the theory to the experimental data. We include MEX-5 in our description in an effective way, since we do not consider its mixing entropy or that of the complex MEX-5:mRNA. We do this since MEX-5 does not exhibit phase separation and its volume fraction is rather small, thus it has a negligible mixing entropy.

The presence of MEX-5 modifies the binding equilibrium of the system. In particular, it modifies the constrained path that relates the concentrations of PGL-3 and PGL-3:mRNA is given by Eq. (5.5). In order to assess the effect of MEX-5 on PGL-3 phase separation, we study the different scenarios discussed in Fig. 5.3(d) by fixing the total concentration of PGL-3 protein to be $n_P^{\mathrm{T}} = 0.6\,\mu\mathrm{M}$ and varying the values of the total concentrations of MEX-5 and mRNA (see the legend of Fig. 5.7). We show the different constrained paths in Fig. 5.7(a), where we see how the interplay between MEX-5 and mRNA affects phase separation. The only case in which the composition lies in the mixing region (green circle), is the one in which the concentration of MEX-5 exceeds that of mRNA. The presence of MEX-5 inhibits phase separation by binding to most of the available mRNA, thus decreasing the ability of PGL-3 to form PGL-3:mRNA complexes which have a higher tendency to phase separate. The rest of the compositions lie within the binodal, which is in agreement with the experimental observations shown in Fig. 5.3(d). In Fig. 5.7(b) we show the fraction of the total fluorescence within the condensed phase, $\mathcal{I}$, defined by

$$\mathcal{I} = \frac{\bar{\mathrm{I}} - \mathrm{I}^{\mathrm{out}}}{\mathrm{I}^{\mathrm{in}} - \mathrm{I}^{\mathrm{out}}} = \frac{n_P^{\mathrm{T}} - n_P^{\mathrm{out}}}{n_P^{\mathrm{in}} - n_P^{\mathrm{out}}} \quad , \tag{5.10}$$

where $\bar{\mathrm{I}}$ is the total fluorescence density, given by $\bar{\mathrm{I}} = \mathrm{I}_0 n_P^{\mathrm{T}}$. The predictions from our model are in qualitative agreement with the experimental observation of larger fractions of total fluorescence within the drops corresponding to larger concentrations of total mRNA.

We have shown that PGL-3 phase separation can be mediated by an mRNA-binding competition. If we now remember that in the *C. elegans* embryo, there is a MEX-5 gradient established from anterior to posterior, we may then propose that the mechanism driving P granule positioning is the depletion of free mRNA at the anterior of the embryo. This depletion would lead to dissolution of P granules, whereas at the posterior side the amount
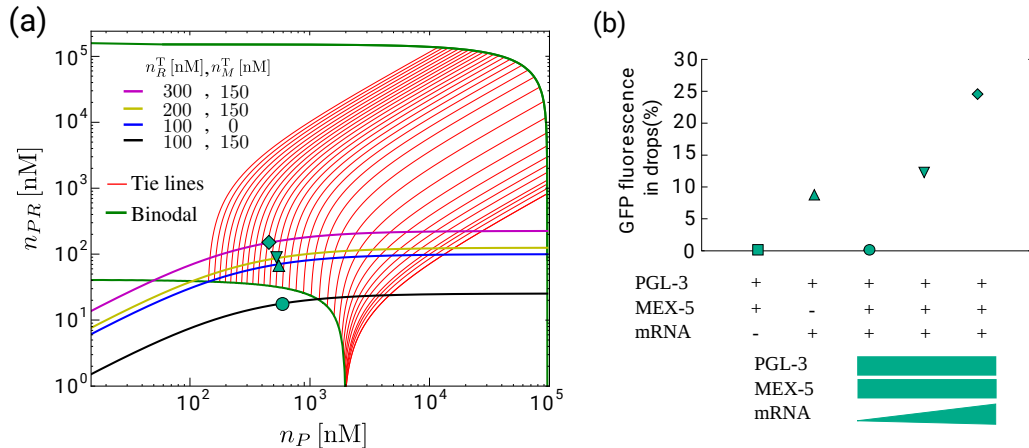
Figure 5.7: Regulation of PGL-3 phase separation by MEX-5 and mRNA. (a) Phase diagram using the parameters extracted from the fits to the experimental data. Constrained paths corresponding to different compositions are shown as colored lines (the compositions are indicated in the legend). All symbols refer to a total concentration of PGL-3, $n_P^T = 0.6\,\mu$M in the different corresponding constrained paths. The binodal is shown as a green line and the red lines correspond to different tie lines. (b) Fraction of the total fluorescence within the condensed phase as predicted by our model. Each symbol denotes a different composition given by: square ($n_R^T = 0$ nM, $n_M^T = 150$ nM), triangle ($n_R^T = 100$ nM, $n_M^T = 150$ nM), circle ($n_R^T = 200$ nM, $n_M^T = 150$ nM), square ($n_R^T = 300$ nM, $n_M^T = 150$ nM). Parameters: $\chi_{P,W} = 0.512$, $\chi_{PR,W} = 0.518$, $1/\epsilon_P = 1.48 \times 10^4$, $1/\epsilon_{PR} = 1.56 \times 10^4$, $K_{PR} = 230$ nM. Figures adapted from [28].

of free mRNA will be higher and PGL-3 and other P granule components such as PGL-1 will be able to bind to it, enabling the condensation of P granules, thus effectively leading to a positioning of the P granules to the posterior of the embryo.

5.7   Positioning of protein droplets in the cell

Here, we assess the possibility of the mRNA-binding competition mechanism to explain the P granule positioning observed in the *C. elegans* embryo. In [28], we also presented solutions to a set of dynamical equations derived from our thermodynamic model. There, we considered dynamic equations based on conservation laws and linear irreversible thermodynamics [115]. Since the dynamical equations describe an inhomogeneous system, a further dependence on gradients of the concentration of phase-separating components is added to the free energy density in the spirit of the Cahn-Hilliard free energy [116]. The derivation and the numerical solutions to the dynamic equations was developed by Dr. Christoph A. Weber and for this reason, we refer the reader to the paper for a full discussion of the equations [28]. Here, for the sake of completeness we only present the numerical results. In Fig. 5.8 we see the process of droplet positioning due to the mRNA-binding competition mechanism. The starting point is a homogeneous distribution of droplets (left upper panel) as well as a homogeneous distribution of MEX-5 and mRNA (right upper

panel). When the MEX-5 gradient starts building up (right middle panel), PGL-3 droplets start dissolving on the left region of the domain (left middle panel). Finally, when the gradient of MEX-5 is fully established (right lower panel), the droplets are positioned to the right side of the domain (lower left panel). What this shows is that the positioning of liquid drops (red circles on the left panels) happens concomitantly with the establishment of a MEX-5 gradient along the x-axis (red line shown in the panels on the right). Thus, the numerical solutions to the dynamic equations support the hypothesis of P granule positioning being driven driven by an mRNA-binding competition mechanism.
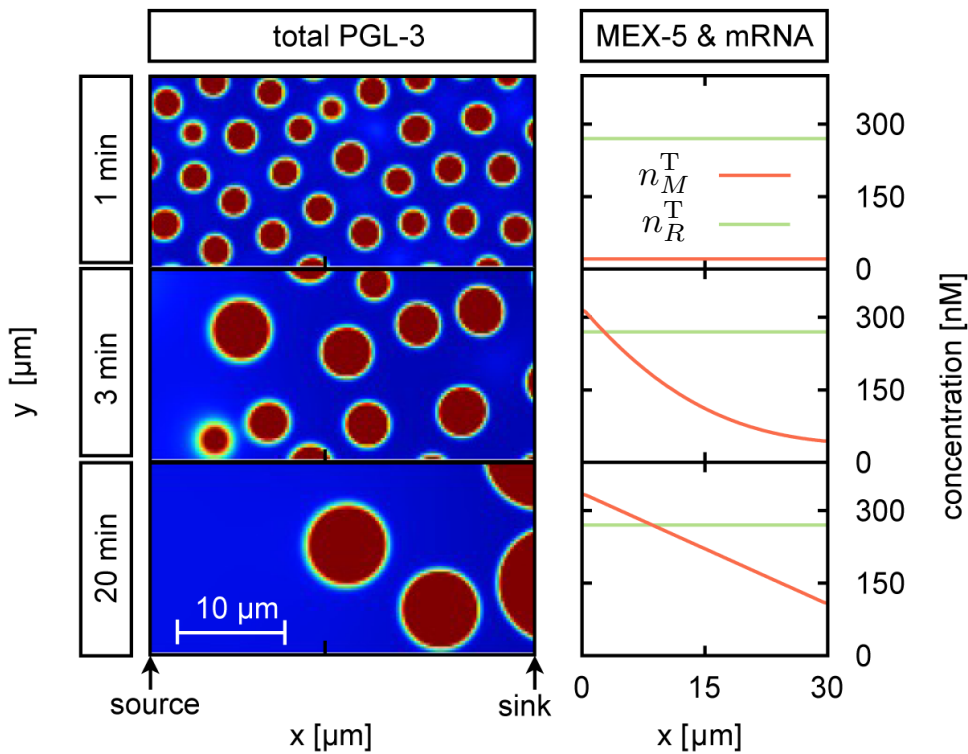

FIGURE 5.8: Positioning of liquid droplets. Left panels show the time evolution of total PGL-3 in solution with mRNA and MEX-5. Going from top to bottom the time increases. The system starts with homogeneous nucleation of drops in the entire domain. Establishment of a MEX-5 gradient leads to dissolution of liquid drops on the left and to condensation on the right. The bottom panel shows localized drops on the right. Right panels show the time evolution of the total concentration of MEX-5 and mRNA averaged over the y-axis. Figure adapted from [28].

5.8 Summary

In this chapter we combined experimental data and theory to explain how an mRNA-binding mechanism regulates PGL-3 phase separation and could account for the positioning of P granules to the posterior side of the *C. elegans*. We first presented experimental data showing that PGL-3 undergoes liquid phase separation and forms liquid-like drops *in vitro*.

We then showed that this behavior is regulated by mRNA, which acts as a promoter of PGL-3 phase separation. The increase in the tendency for phase separation seems to be associated with the ability of PGL-3 to bind mRNA and form complexes, PGL-3:mRNA. We further presented experimental data revealing that MEX-5, the protein whose gradient from posterior to anterior side in the embryo anticorrelates with P granule positioning, also binds to mRNA and it does with a much stronger binding affinity than PGL-3. Further experimental evidence shows that MEX-5 reduces the effect of mRNA on promoting PGL-3 phase separation. All these observations suggested that a subtle interplay between the tendency of PGL-3 to form drops regulated by mRNA-binding and a competition between MEX-5 and PGL-3 to bind the existing mRNA lead to spatial control of dissolution and condensation of PGL-3 drops.

In order to explain the experimental observations, we constructed a minimal model for PGL-3 phase separation. The model is based on a Flory-Huggins type of free energy complemented by a set of binding reactions based on experimental evidence. We discussed the binding equilibrium for a simplified version of our model and showed that the binding reactions constrain the possible concentrations of PGL-3 and the complex PGL-3:mRNA to lie within what we called the constrained path. The constrained path is fully determined by: the total concentration of mRNA, the total concentration of MEX-5 and the dissociation constants of the complexes PGL-3:mRNA and MEX-5:mRNA, for which we used the experimentally measured values. We then fitted first the model of a binary mixture to the PGL-3 solution in the absence of mRNA and MEX-5. From this fit we extracted the molecular volume ratio between PGL-3 and the solvent (buffer), ϵ_P, which shows that PGL-3 has a volume 10^4 larger than the solvent. We also found the value of the Flory-Huggins parameter $\chi_{P,W} = 0.512$, which is relatively close to the critical value, showing that due to the large interacting volume of PGL-3, the interaction needed for phase separation is a rather weak one. We then used these values to fit the experimental data corresponding to the scenario where the presence of mRNA promotes PGL-3 phase separation. For this fit, we used the constrained path (in the absence of MEX-5) and showed that to account for the experimental data, the tendency for phase separation of the complex PGL-3:mRNA must be higher than that of PGL-3 alone, $\chi_{PR,W} > \chi_{P,W}$ with $\chi_{PR,W} = 0.518$. In order to further test our model, we included MEX-5 in the description and studied its effect on PGL-3 phase separation. We showed that our model can also account for the inhibition effect associated with the presence of MEX-5, which only occurs when mRNA is also present in the solution, thus highlighting the mRNA-binding competition effect between PGL-3 and MEX-5. We concluded by showing numerical solutions to a set of dynamic equations based on conservation laws describing changes in time of the concentrations of each component in the system. The numerical solutions showed that an mRNA-binding competition mechanism can fully account for the positioning of liquid droplets to a particular side.

This study set the basis for explaining different phenomena related to protein condensation and its spatial localization. Surprisingly, we showed that a relatively simple model sheds light on a mechanism that is highly relevant for the proper development of an adult *C. elegans.* We believe that this work may encourage further research on *in vitro* protein solutions. These studies will help in characterizing the different protein phase separation properties and how different binding partners of these proteins such as mRNA, may influence protein phase separation. To conclude, the mechanism presented here is generic, meaning that other proteins capable of forming liquid-like drops and having similar mRNA-binding properties as those of PGL-3, would show a similar regulation mechanism.

Chapter 6

Conclusions and Outlook

In this work, we have established a consistent thermodynamic framework to study phase separation at chemical equilibrium and highlighted the importance of considering the conservation laws associated to the chemical reactions in the system. Using this thermodynamic framework, we showed that minimal models of multicomponent mixtures can account for liquid phase separation processes observed in experiments of protein solutions *in vitro* as well as in living organisms. Remarkably, the minimal models presented here, provide a physical interpretation of biological phenomena occurring inside the utterly complex environment of the cell, in particular, we addressed the underlying mechanism leading to P granules position inside the *C. elegans* embryo and the pH-dependent macromolecular assembly observed in yeast cells, which were discussed in Chapter 1. One important factor that enabled us to perform our studies was the development of numerical tools to construct phase diagrams for multicomponent systems. They were based on the idea of the convexity of the free energy (density) as a function of its composition variables, and importantly, the construction of coexisting phases in systems with chemical reactions was performed in a thermodynamic space with a reduced dimension defined by the number of conserved components and not by the total number of components involved in the chemical reactions. We now recapitulate our developments and findings.

In Chapter 2, we described the numerical methods used to construct phase diagrams. Since the convex hull selects the convex region of the free energy density, it allow us to discriminate between stable and unstable regions in the composition space. Using the convex hull we discussed how to construct the phase diagrams in relatively simple systems, namely binary and ternary mixtures. We gave a brief analysis of the dependence of phase diagrams in different choices of the parameters, where we argued that by studying simple binary mixtures composed of proteins and buffer as a function of temperature, we could differentiate the mechanisms behind phase-separating system. We also discussed the already rich behavior of a ternary system, where with the convex hull construction, the determination of the binodals becomes a relatively simple problem without having any prior

knowledge of the phase behavior of the system. We concluded this chapter by showing possible quantitative phase diagrams of proteins that are intrinsically disordered. The goal of this chapter was mainly to show how to use the convex hull in the construction of phase diagrams and discuss different properties arising from large asymmetries in the molecular volumes as well as from the different interaction among components.

In Chapter 3, we developed a generic thermodynamic framework to discuss phase separation in systems undergoing chemical reactions. We showed how starting from the stoichiometry matrix of the system, one can construct a complete set of conserved components. Interestingly, the set of the particle numbers corresponding to the conserved components can be chosen as the set of independent composition variables describing the system at chemical equilibrium. By doing a variable transformation from the initial set of components to one that includes the conserved ones, the chemical equilibrium conditions can be derived together with the thermodynamic conjugate chemical potentials of these conserved components. These conjugate variables are the ones needed to perform correct Legendre transforms in systems with chemical reactions [76]. We also showed that if one wants to find coexisting phases in systems with chemical reactions, by using the conserved particle numbers as natural variables, the methods developed in Chapter 2 can be applied without any further change, of course granted that we evaluate the free energy in compositions satisfying the chemical equilibrium conditions. We concluded by illustrating the methods developed in this chapter applied to a system in which a molecule can transition from a state which does not have a tendency to phase-separate to a state in which it does. We selected the parameters of the system in such a way that the transition from the non-phase-separating state to the phase-separating of the molecule was promoted by heating. We then constructed a full phase diagram in the composition-temperature space showing that this simple system exhibits reentrant behavior as a function of temperature.

In Chapter 4, we presented a minimal model to study the effect of pH on liquid-liquid phase separation. The model considered was a multicomponent mixture composed of water, hydronium ions, hydroxide ions, and macromolecules which could exist in three different charge states. In order to assess the effect of pH on macromolecular phase separation, we coupled a phase-separating system with a set of chemical reactions controlled by pH. This set of chemical reactions described the protonation and deprotonation of the macromolecule as well as the self-ionization of water. Using the conserved particle numbers under chemical changes as independent composition variables, we constructed their thermodynamic conjugate chemical potentials at chemical equilibrium. This then allowed us to perform the correct Legendre transform to fix the pH of the system without using any approximation of infinite dilution in the concentration of hydronium ions. Furthermore, using the chemical equilibrium conditions and the particle numbers corresponding to the conserved components, we reduced the effective dimension of the thermodynamic space to one composition, namely the total macromolecule concentration. We constructed phase

diagrams by means of the convex hull as a function of temperature, pH and total macro-molecular volume fractions. The main features that we found were: Broad regions of phase separation at the isoelectric point of the macromolecules, different topologies depending on whether the dominant interaction driving phase separation is an attraction between oppositely charged macromolecules or an attraction between neutral macromolecules, and reentrant behavior as a function of pH. Additionally, we discussed a more realistic scenario where we constructed a phase diagram using values for the total amount of fixed charges and maximal charge close to those that are found in intrinsically disordered proteins which have a pH-responsive behavior, such as Pab1 [34] and Sup35 [30]. This phase diagram, showed that phase separation for macromolecules with these parameter choices will happen for a very wide range of concentrations, going from μM concentrations, in agreement with protein phase separation *in vitro* [30], to concentrations of mM, close to an estimation of the total protein concentration in yeast cells [108]. Finally, we showed that differences in both electric potential and pH across coexisting phases naturally emerge if the charge neutrality condition is considered in an explicit manner.

In Chapter 5, we proposed a physical mechanism to explain the phenomenon of P granule positioning in the *C. elegans* embryo [12]. This mechanisms consists of a a competition between PGL-3 and MEX-5 to bind to mRNA, spatially modulated by the MEX-5 gradient across the embryo. In order to reach this conclusion, we presented results of our work [28], which combined *in vitro* reconstitution of protein liquid droplets, a careful study of protein binding to RNA, and physical modeling. In our study we focused on two of them, PGL-3, a key component of P granules and MEX-5. After presenting the experimental data showing that: PGL-3 form liquid-like drops *in vitro*, PGL-3 and MEX-5 have an RNA-binding affinity, with MEX-5 binding approximately twenty-fold times stronger to RNA than PGL-3, and the presence of mRNA and MEX-5 can regulate the phase separation tendency of PGL-3. We then presented a model at thermodynamic equilibrium describing phase separation regulated via binding reactions. We first considered a multicomponent Flory-Huggins free energy composed of PGL-3, mRNA, the complexes formed by PGL-3 binding to RNA, PGL-3:mRNA, and solvent (buffer). Using this model we fitted two different sets of experiments, one where PGL-3 formed drops in the absence of mRNA and the second one where mRNA is added and shown to enhance the tendency to phase separate of PGL-3. To fit the first set of experiments we used a simple binary mixture model, from which the Flory-Huggins interaction parameter between PGL-3 and the solvent can be extracted. Using this value, we showed that in order to reproduce the experimental results, the tendency to phase separate of the complex PGL-3:mRNA had to be larger than that of PGL-3 alone. We then included MEX-5 in the description by considering its binding reaction to mRNA, this led to a reduced tendency of PGL-3 to phase separate in the presence of mRNA, due to the fact that MEX-5 has a higher binding affinity to RNA than PGL-3. Finally, in order to make sure that the binding competition was enough to

segregate protein droplets to a specific side, we presented the solutions of the dynamical equations for the concentrations of all components corresponding to the free energy of the full thermodynamic model, and showed that indeed, the mRNA binding competition between MEX-5 and PGL-3 can for the positioning of liquid phase separated drops. This mechanism is generic in the sense that any other protein with similar binding properties as PGL-3 such as MEG-3 [39] could also be positioned via the mRNA-binding competition mechanism.

Regarding future work, maybe the most interesting aspect that could be further explored is whether or not protein condensates develop a difference in pH and electric potential across their interfaces. An electric potential difference could serve as a selective mechanism to enrich these condensates specifically in certain charged molecules and a pH difference could play a relevant role for changes in the reaction rates inside these condensates. Exploring the equilibrium thermodynamics could set bounds on the pH difference that could be achieved via passive phase separation. Further extensions of the theory to out-of-equilibrium systems could explore whether or not pH differences across the interfaces of biomolecular condensates can be actively controlled. Another relevant aspect that was not addressed in this thesis is the multicomponent nature of biomolecular condensates. For example, P granules alone colocalize over thirty different proteins [117], one question that arises is, what is the precise selection mechanism that allows a particular group of proteins to phase separate together and form these protein condensates, this question could be explored by trying to establish what defines such groups of proteins in terms of the interactions among them, one could try to address this question using the mean field free energy models discussed in this thesis. If we consider the fact that although there is a huge number of different macromolecules in the cell, there is a rather limited number of simultaneous coexisting phases, the question of how does the cell resolves this conundrum might be worth studying. A first step to address this question could be to combine a multicomponent theory and study how does a system of chemical reactions constrain the thermodynamic space that can be visited by such system in order to select specific regions in which only a low number of simultaneous coexisting phases may exist. We mention this only as a first step since the active nature of the cell could be the one responsible of limiting the number of coexisting phases, however, this remains to be explored. Finally, we believe that our work can simply be taken as a starting point in exploring phase diagrams of intrinsically disordered proteins which undergo different reactions that affect their tendency to phase-separate.

Let us conclude by posing the same question that was asked in Chapter 1 but in a different tense: Was it relevant to study systems at thermodynamic equilibrium in order to understand what happens inside of a living cell? The answer seems to be yes and this work tried to give a small glimpse of what can be done by considering equilibrium systems to understand recurrent behaviors in cells.

Appendix A

Flory-Huggins Parameters

Here we show how to describe the free energy given in Eq. (1.2) in terms of suitably defined effective Flory-Huggins (FH) interaction parameters. In order to find the relation between these interaction parameters we will use the condition $\sum_i \phi_i = 1$. Using the free energy (1.2) we define the interaction term as

$$U_{\text{int}} = G - k_{\text{B}}T \sum_{i=0}^{M} N_i \ln \phi_i - \sum_{i=0}^{M} w_i N_i - p\,V = \frac{V k_{\text{B}}T}{2v_0} \sum_{i,j=0}^{M} \chi_{ij} \phi_i \phi_j \quad . \tag{A.1}$$

We add and subtract $\frac{V k_{\text{B}}T}{2v_0} \sum_{k=0}^{M} \chi_{kk} \phi_k$ to U_{int} and rearrange terms to get

$$U_{\text{int}} = \frac{V k_{\text{B}}T}{2v_0} \left(\sum_{k=0}^{M} \chi_{kk} \phi_k \left(\phi_k - 1 \right) + \sum_{k \neq l} \phi_k \phi_l - \sum_{k=0}^{M} \chi_{kk} \phi_k \right) \quad ,$$

where the sum $\sum_{k \neq l}$, considers that k and l take all possible values from 0 to M in which they are different from each other. Finally, we add and subtract $\frac{V k_{\text{B}}T}{2v_0} \sum_{k \neq l} \chi_{kk} \phi_k \phi_l$ and obtain

$$
\begin{aligned}
U_{\text{int}} &= \frac{V k_{\text{B}}T}{2v_0} \sum_{k=0}^{M} \chi_{kk} \phi_k \left(\phi_k + \sum_{l \neq k} \phi_l - 1 \right) + \frac{V k_{\text{B}}T}{2v_0} \sum_{k \neq l} \left(\chi_{kl} - \chi_{kk} \right) \phi_k \phi_l \\
&\quad - \frac{V k_{\text{B}}T}{2v_0} \sum_{k=0}^{M} \chi_{kk} \phi_k \quad .
\end{aligned}
\tag{A.2}
$$

Because $\sum_{i=0}^{M} \phi_i = 1$, the first term in Eq. (A.2) vanishes. It is simple now to relate the Flory-Huggins parameter to the interaction parameters by

$$\chi_{kl}^{\text{FH}} = \frac{2\chi_{kl} - \chi_{kk} - \chi_{ll}}{2} \quad . \tag{A.3}$$

Using Eq. (A.3), U_{int} can then be expressed as

$$U_{\text{int}} = \frac{V k_{\text{B}} T}{2 v_0} \sum_{k \neq l} \chi_{kl}^{FH} \phi_k \phi_l - \frac{V k_{\text{B}} T}{2 v_0} \sum_{k} \chi_{kk} \phi_k \quad , \tag{A.4}$$

Defining $\omega_k = \left(w_k - \frac{1}{2 v_k} \chi_{kk} \right)$ and using Eq. (A.4), we can rewrite the free energy given in Eq. (1.2) as

$$G(T, p, \{N_i\}) = k_{\text{B}} T \left(\sum_{i=0}^{M} N_i \ln \phi_i + \frac{V}{2 v_0} \sum_{i \neq j} \chi_{ij}^{\text{FH}} \phi_i \phi_j \right) + \sum_{i} \omega_i N_i + p V \quad . \tag{A.5}$$

We have showed here that we can also express the free energy Eq. (1.2) using the Flory-Huggins parameter by redefining the linear term in the Gibbs free energy.

Appendix B

Stability Conditions for Multicomponent Mixtures

Here we present the general stability conditions for an incompressible multicomponent mixture composed of M different chemical components and described by a free energy density $f(T, \{n_i\})$. The free energy density is locally stable if the free energy density f obeys the following relations [118]:

$$\frac{\partial^2 f}{\partial n_i^2} \geq 0 \quad \text{for all } i \quad , \tag{B.1}$$

$$\begin{vmatrix} \dfrac{\partial^2 f}{\partial n_1^2} & \dfrac{\partial^2 f}{\partial n_1 \partial n_2} & \cdots & \dfrac{\partial^2 f}{\partial n_1 \partial n_M} \\ \dfrac{\partial^2 f}{\partial n_2 \partial n_1} & \dfrac{\partial^2 f}{\partial n_2^2} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial n_M \partial n_1} & \cdots & \cdots & \dfrac{\partial^2 f}{\partial n_M^2} \end{vmatrix} \geq 0 \quad . \tag{B.2}$$

These local stability conditions are weaker conditions than the global convexity of the free energy but can be used to find the boundary between stable and unstable states. The locus of points defining the boundary between stable and unstable regions of the free energy density in composition space is called the spinodal [63, 118]. The spinodal is defined by compositions in which the left hand side of Eq. B.2 is equal to zero

$$\begin{vmatrix} \dfrac{\partial^2 f}{\partial n_1^2} & \dfrac{\partial^2 f}{\partial n_1 \partial n_2} & \cdots & \dfrac{\partial^2 f}{\partial n_1 \partial n_M} \\ \dfrac{\partial^2 f}{\partial n_2 \partial n_1} & \dfrac{\partial^2 f}{\partial n_2^2} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial n_M \partial n_1} & \cdots & \cdots & \dfrac{\partial^2 f}{\partial n_M^2} \end{vmatrix} = 0 \quad . \tag{B.3}$$

The critical points of the system are precisely located on the spinodal and can be found by substituting any of the horizontal lines in the determinant in Eq. (B.3) for

$$
\frac{\partial}{\partial n_1}\left(\frac{\partial^2 f}{\partial n_1 \partial n_k}\right), \frac{\partial}{\partial n_2}\left(\frac{\partial^2 f}{\partial n_2 \partial n_k}\right), \ldots, \frac{\partial}{\partial n_M}\left(\frac{\partial^2 f}{\partial n_M \partial n_k}\right) \quad , \tag{B.4}
$$

where we have arbitrarily chosen the k-th horizontal line in the determinant. If we choose $k = M$, without any loss of generality, the second critical condition is then given by

$$
\begin{vmatrix}
\dfrac{\partial^2 f}{\partial n_1^2} & \dfrac{\partial^2 f}{\partial n_1 \partial n_2} & \cdots & \dfrac{\partial^2 f}{\partial n_1 \partial n_M} \\
\dfrac{\partial^2 f}{\partial n_2 \partial n_1} & \dfrac{\partial^2 f}{\partial n_2^2} & \cdots & \vdots \\
\vdots & \vdots & \ddots & \vdots \\
\dfrac{\partial^3 f}{\partial n_M \partial n_1^2} & \dfrac{\partial^3 f}{\partial n_M \partial n_2^2} & \cdots & \dfrac{\partial^3 f}{\partial n_M^3}
\end{vmatrix} = 0 \quad , \tag{B.5}
$$

Thus, the critical points are given by simultaneous solutions of Eqs. (B.3) and (B.5) and they are characterized by a critical temperature T_c and critical concentrations n_i^c. For a binary mixture with only one independent concentration n, Eqs. (B.3) and (B.5) reduce to the usual critical conditions presented in every thermodynamics textbook, given by

$$
\left.\frac{\partial^2 f}{\partial n^2}\right|_{T_c, n_c} = 0 \quad ,
$$
$$
\left.\frac{\partial^3 f}{\partial n^3}\right|_{T_c, n_c = 0} \quad .
$$

where T_c and n_c are the critical temperature and the critical concentration, respectively.

Appendix C

Convex Hull Construction

Here we give a brief and simple introduction to what is a convex hull and how does the *Quickhull* algorithm works [83]. To this end, let us first imagine a set of points $\mathcal{X}$ in the $x-y$ plane with each of the points defined by a pair of coordinates, (x_i, y_i), see Fig. C.1(a). The convex hull $\mathcal{C}$ is defined as the minimum set of points within $\mathcal{X}$ that encloses all of the points in $\mathcal{X}$ (with the boundary included). The points of the set which are part of the convex hull are called the vertices of the convex hull.

If we now ask ourselves, how do we construct the convex hull $\mathcal{C}$, one intuitive and relatively fast way of doing it is by means of the *Quickhull* algorithm. In what follows we describe the steps of the construction for the set of points $\mathcal{X}$.

1. Find the two extreme points in any direction (x or y) of the set and join them with a line, in Fig. C.1(b) we show the two extreme points along the x-axis (1,2). These points are at the extreme values of x so they are vertices of the convex hull (vertices are shown as blue points).

2. We then find the point lying at maximum positive distance from the line joining the two extreme points and connect it with a line, this point is also a vertex of the convex hull. We consider the distance to a point positive if the point lies on the half-plane to the right of the corresponding line. In Fig. C.1 we show the triangle that is formed by joining the points 1-2-3.

3. We check if there are points inside of the triangle formed by the points from the previous point. We see in Fig. C.1(c) that the points 4 and 5 lie within the area enclosed by the triangle 1-2-3, thus they are not part of the convex hull (yellow points).

4. We repeat the same procedure as in step 2 but now measuring the distance from the segment given by the directed line from 2, which is point 6. We join these points to form the triangle 1-2-6 shown in Fig. C.1(d) and repeat step 3.

5. We repeat steps 2-3 from the new lines formed by joining the vertices 1-3, 3-2, 6-1 and 2-6.

6. We repeat the steps of joining furthest distance to each of the lines formed by the new vertices of the convex hull until there are no further points lying outside of the convex hull.
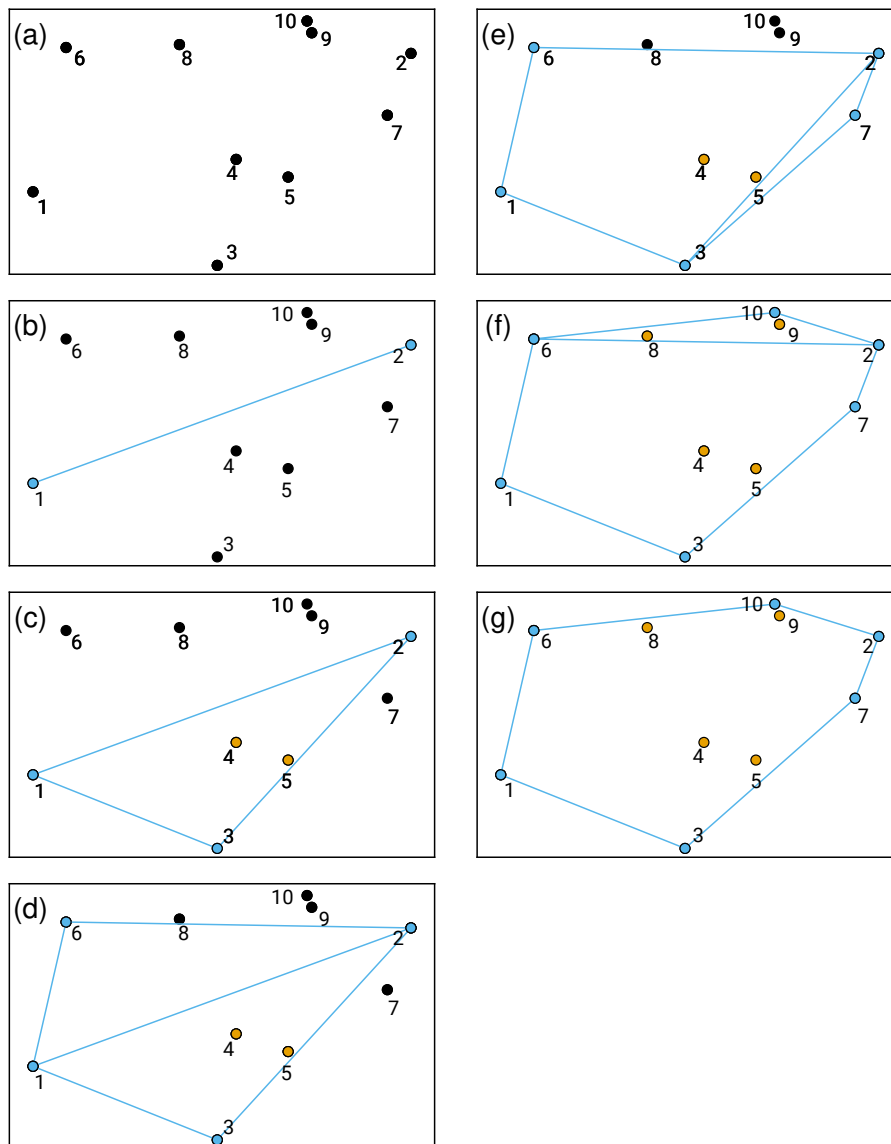

FIGURE C.1: Steps describing the construction of the convex hull of a set of points using the *Quickhull* algorithm.

The *Quickhull* algorithm is very simple to visualize in two dimensions but it can be generalized to any number of dimensions. Instead of lines, one has planes and one measures distances from points to planes and check for points lying inside of the volumes instead of

the area. Interestingly the convergence of the method improves with the number of points lying inside the convex hull.

To conclude this Appendix, if we calculate the convex hull of a free energy density, we obtain the set of compositions for which the free energy density is convex, note that this is not a local convexity property but a global one. The convex hull rules out metastable states that are locally stable under perturbations, thus selecting only the minimum convex function, therefore the convex hull allows us to select between states that are globally stable and those who are not.

Appendix D

Construction of a Set of Independent Conserved Components and Their Corresponding Particle Numbers

Here we show how to construct a set of independent conserved components for a specific example. Let us consider a system with two reactions and five chemical components. One of the reactions is the self-ionization of water (3.1), involving three of the five chemical components, water H_2O, hydronium H_3O^+ and hydroxide OH^-. The other is a reaction where one macromolecule, denoted by M, may become positively charged by absorbing one proton and turn into an ionized macromolecule denoted by HM^+. These two reactions can be expressed as follows

$$2\,H_2O \rightleftharpoons H_3O^+ + OH^- \quad , \tag{D.1}$$

$$HM^+ + H_2O \rightleftharpoons M + H_3O^+ \quad . \tag{D.2}$$

The stoichiometric matrix ν_{ij} is given by

$$\nu_{ij} = \begin{pmatrix} 1 & 1 & -2 & 0 & 0 \\ 1 & 0 & -1 & 1 & -1 \end{pmatrix} \quad , \tag{D.3}$$

where each row represents a different chemical reaction and each column the number of molecules of each of the components involved in the chemical reaction in the following order: H_3O^+, OH^-, H_2O, M and HM^+. In order to find the null space of ν_{ij}, we need to solve

$$\sum_{j=0}^{s} \nu_{ij} C_j = 0 \quad ,$$

for all i, i.e. for every reaction. The system of equations reads

$$C_1 + C_2 - 2C_3 \;=\; 0 \quad, \tag{D.4}$$

$$C_1 - C_3 + C_4 - C_5 \;=\; 0 \quad. \tag{D.5}$$

Solving for C_1 and C_2 we obtain

$$C_1 \;=\; C_3 - C_4 + C_5 \quad, \tag{D.6}$$

$$C_2 \;=\; C_3 + C_4 - C_5 \quad. \tag{D.7}$$

This allow us to write the solution vector $\bar{C}$ as

$$\bar{C} = C_3 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + C_4 \begin{pmatrix} -1 \\ 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} + C_5 \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad, \tag{D.8}$$

where C_3, C_4 and C_5 are arbitrary numbers. The three vectors on the right hand side of Eq. (D.8) form a basis of the null space of the stoichiometry matrix. We name this three basis vectors as

$$\bar{C}^1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad, \quad \bar{C}^2 = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad, \quad \bar{C}^3 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad. \tag{D.9}$$

The conservation matrix associated to this basis is

$$\mathcal{C}_{ij} = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{D.10}$$

We can then proceed to construct a set of particle numbers for the independent conserved components, we do this by means of Eq. (3.5) to obtain three independent particle numbers,

which are given by

$$N_1^c = N_{\text{H}_3\text{O}^+} + N_{\text{OH}^-} + N_{\text{H}_2\text{O}} \quad , \tag{D.11a}$$

$$N_2^c = -N_{\text{H}_3\text{O}^+} + N_{\text{OH}^-} + N_{\text{M}} \quad , \tag{D.11b}$$

$$N_3^c = N_{\text{H}_3\text{O}^+} - N_{\text{OH}^-} + N_{\text{HM}^+} \quad . \tag{D.11c}$$

The first conserved particle number has a very clear meaning, it is the number of oxygen atoms in the system, whereas the remaining two do not seem to have an intuitive meaning. In order to give Eq. (D.11b) and Eq. (D.11c) a meaning, let us exploit the fact that we have a rather simple system and guess three different conserved components. The first one is the number of oxygen atoms N_{O}, the second is the number of hydrogen atoms N_{H} and the third one is the total number of macromolecules, N_{M_T}. The particle number of these conserved components are defined as

$$N_{\text{O}} = N_{\text{H}_3\text{O}^+} + N_{\text{OH}^-} + N_{\text{H}_2\text{O}} \quad , \tag{D.12a}$$

$$N_{\text{H}} = 3N_{\text{H}_3\text{O}^+} + N_{\text{OH}^-} + 2N_{\text{H}_2\text{O}} + N_{\text{HM}^+} \quad , \tag{D.12b}$$

$$N_{\text{M}_\text{T}} = N_{\text{M}} + N_{\text{HM}^+} \quad . \tag{D.12c}$$

We now show that these two different sets of conserved quantities are related, we call them quantities in order to be precise since N_{c_2} and N_{c_3} can be negative depending on the initial state of the system. The set of conserved quantities defined in Eqs. (D.11) can be expressed as a function of N_{O}, N_{H} and N_{M_T} as follows

$$N_1^c = N_{\text{O}} \quad , \tag{D.13a}$$

$$N_2^c = N_{\text{M}_\text{T}} + 2N_{\text{O}} - N_{\text{H}} \quad , \tag{D.13b}$$

$$N_3^c = N_{\text{H}} - 2N_{\text{O}} \quad , \tag{D.13c}$$

where we see that N_2^c and N_3^c, are a non-obvious combination of the more intuitive conserved quantities.

Appendix E

pH in Diluted Systems

We now show that our definition of pH (eq. (4.19)) is equivalent to the most commonly used definition $\text{pH} = -\log_{10}(n_{\text{H}_3\text{O}^+}/n^0_{\text{H}_3\text{O}^+})$ which only applies for ideal solutions of H_3O^+ and OH^- in water. In the absence of macromolecules and considering ideal solution conditions, the chemical potentials of water $\mu_{\text{H}_2\text{O}}$ and of hydronium ions $\mu_{\text{H}_3\text{O}^+}$ can be expressed as

$$\mu_{\text{H}_2\text{O}} = v_0 P + w_{\text{H}_2\text{O}} \quad , \tag{E.1}$$

$$\mu_{\text{H}_3\text{O}^+} = k_\text{B} T \ln(v_0 n_{\text{H}_3\text{O}^+}) + v_{\text{H}_3\text{O}^+} P + w_{\text{H}_3\text{O}^+} \quad , \tag{E.2}$$

where we assumed that the contribution of the ions to the volume V is negligible, i.e. taking the volume fraction of water $v_0 n_{\text{H}_2\text{O}} = 1$. The standard chemical potentials $\mu^0_{\text{H}_2\text{O}}$ and $\mu^0_{\text{H}_3\text{O}^+}$ are given by the following expressions

$$\mu^0_{\text{H}_2\text{O}} = v_0 P + w_{\text{H}_2\text{O}} \quad , \tag{E.3}$$

$$\mu^0_{\text{H}_3\text{O}^+} = k_\text{B} T \ln(v_0 n^0_{\text{H}_3\text{O}^+}) + v_{\text{H}_3\text{O}^+} P + w_{\text{H}_3\text{O}^+} \quad . \tag{E.4}$$

Substituting Eqs. (E.1)-(E.4) in the definition given in Eq. (4.19) we obtain

$$\text{pH} = -\log_{10}\left(\frac{n_{\text{H}_3\text{O}^+}}{n^0_{\text{H}_3\text{O}^+}}\right) \quad . \tag{E.5}$$

Appendix F

Critical Points at the Isoelectric Point

In this section we calculate some limiting critical values at the isoelectric point, where $\psi = 0$ and $h_\psi = 0$.

F.1 Critical point at $\bar{n} = 1$

Here, we calculate the critical value which emerges at $\bar{n} = 1$ using the free energy density (4.42), which at the isoelectric point is

$$v\bar{f} \quad = \quad k_{\mathrm{B}}T\left(2\phi\ln\phi + (1-2\phi)\ln(1-2\phi)\right) + \chi\left(\phi^2 - 2\lambda\phi + \lambda/2\right)/\epsilon - h_\phi\phi + w_{\mathrm{M}} \quad (\mathrm{F.1})$$

We first differentiate $\bar{f}$ with respect to ϕ which gives

$$v\frac{\partial\bar{f}}{\partial\phi} = k_{\mathrm{B}}T\left(2\ln\phi - 2\ln(1-2\phi)\right) + 2\chi\left(\phi - \lambda\right)/\epsilon - h_\phi \quad . \qquad (\mathrm{F.2})$$

The conditions for finding a critical point in this case are $\partial^2\bar{f}/\partial\phi^2 = 0$ and $\partial^3\bar{f}/\partial\phi^3 = 0$, these conditions read

$$k_{\mathrm{B}}T_c\left(\frac{2}{\phi_c} + \frac{4}{1-2\phi_c}\right) + 2\chi/\epsilon = 0 \quad , \qquad (\mathrm{F.3})$$

$$k_{\mathrm{B}}T_c\left(-\frac{2}{\phi_c^2} + \frac{8}{(1-2\phi_c)^2}\right) = 0 \quad . \qquad (\mathrm{F.4})$$

Solving Eqs. (F.3) and (F.4) we find

$$\phi_c = \frac{1}{4} \quad , \tag{F.5}$$

$$k_{\mathrm{B}}T_c = -\frac{\chi}{8\epsilon} \quad , \tag{F.6}$$

$$h_{\phi,c} = \chi \left(\frac{\ln 2 + 2 - 8\lambda}{4\epsilon} \right) \quad . \tag{F.7}$$

F.2   Effective binary critical point

In order to calculate the critical points describing the effective binary mixture, we first discuss the limits of h_ϕ for $\phi \to 1/2$ and $\phi \to 0$, which correspond to situations where macromolecules are only charged or only neutral respectively.

If we consider the limit of h_ϕ with $\phi \to 1/2$,

$$\lim_{\phi \to \frac{1}{2}} h_\phi = \lim_{\phi \to \frac{1}{2}} \left[2k_{\mathrm{B}}T \ln \frac{\phi}{1-2\phi} + \chi\bar{n}(\phi - \lambda)/\epsilon \right] \quad ,$$

$$\lim_{\phi \to \frac{1}{2}} h_\phi = -2\ln 2 - 2k_{\mathrm{B}}T \lim_{\phi \to \frac{1}{2}} \ln(1-2\phi) + \frac{\chi\bar{n}}{\epsilon} \left(\frac{1}{2} - \lambda \right) \quad ,$$

$$\lim_{\phi \to \frac{1}{2}} h_\phi = \infty \quad ,$$

this shows on one hand that large positive values of h_ϕ are obtained for values of ϕ approaching $1/2$, on the other hand, large negative values of h_ϕ are obtained for ϕ approaching 0.

$$\lim_{\phi \to 0} h_\phi = -\infty \quad .$$

We can study both cases by direct substitution in the free energy density (4.42). We illustrate the case for $\phi = 1/2$, in this case the free energy density reads

$$v\bar{f} = k_B T \left(\bar{n} \ln \left(\frac{\bar{n}}{2} \right) + \frac{(1-\bar{n})}{\epsilon} \ln(1-\bar{n}) \right) + \frac{\chi_{\mathrm{e}}\bar{n}^2}{4\epsilon} + \mathcal{O}(\bar{n}) , \tag{F.8}$$

the linear terms $\mathcal{O}(\bar{n})$ do not affect the stability of the system, therefore we are safe to ignore them in our calculation. From the free energy evaluated at $\phi = 1/2$ (F.8), we calculate the chemical potential up to a constant

$$v\frac{d\bar{f}}{d\bar{n}} = k_B T \left[\ln \left(\frac{\bar{n}}{2} \right) + 1 - \frac{\ln(1-\bar{n}) + 1}{\epsilon} + \frac{\chi_{\mathrm{e}}\bar{n}}{2\epsilon} \right] \tag{F.9}$$

using Eq. (F.9) and conditions $d^2 f/d\bar{n}^2 = 0$ and $d^3 f/d\bar{n}^3 = 0$ we find

$$\bar{n}_c^b = \frac{\sqrt{\epsilon}}{1 + \sqrt{\epsilon}} \quad , \tag{F.10}$$

$$k_{\mathrm{B}} T_c = -\frac{\chi_{\mathrm{e}}}{2(1 + \sqrt{\epsilon})^2} \quad . \tag{F.11}$$

Following the same procedure for a system with $\phi = 0$ gives

$$\bar{n}_c^b = \frac{\sqrt{\epsilon}}{1 + \sqrt{\epsilon}} \quad , \tag{F.12}$$

$$k_{\mathrm{B}} T_c = -2\frac{\chi_{\mathrm{n}}}{(1 + \sqrt{\epsilon})^2} \quad . \tag{F.13}$$

Appendix G

Phase Diagrams as a Function of pH

Here we show the phase diagrams which appear in Fig. 4.4 showing the dependence in ψ.
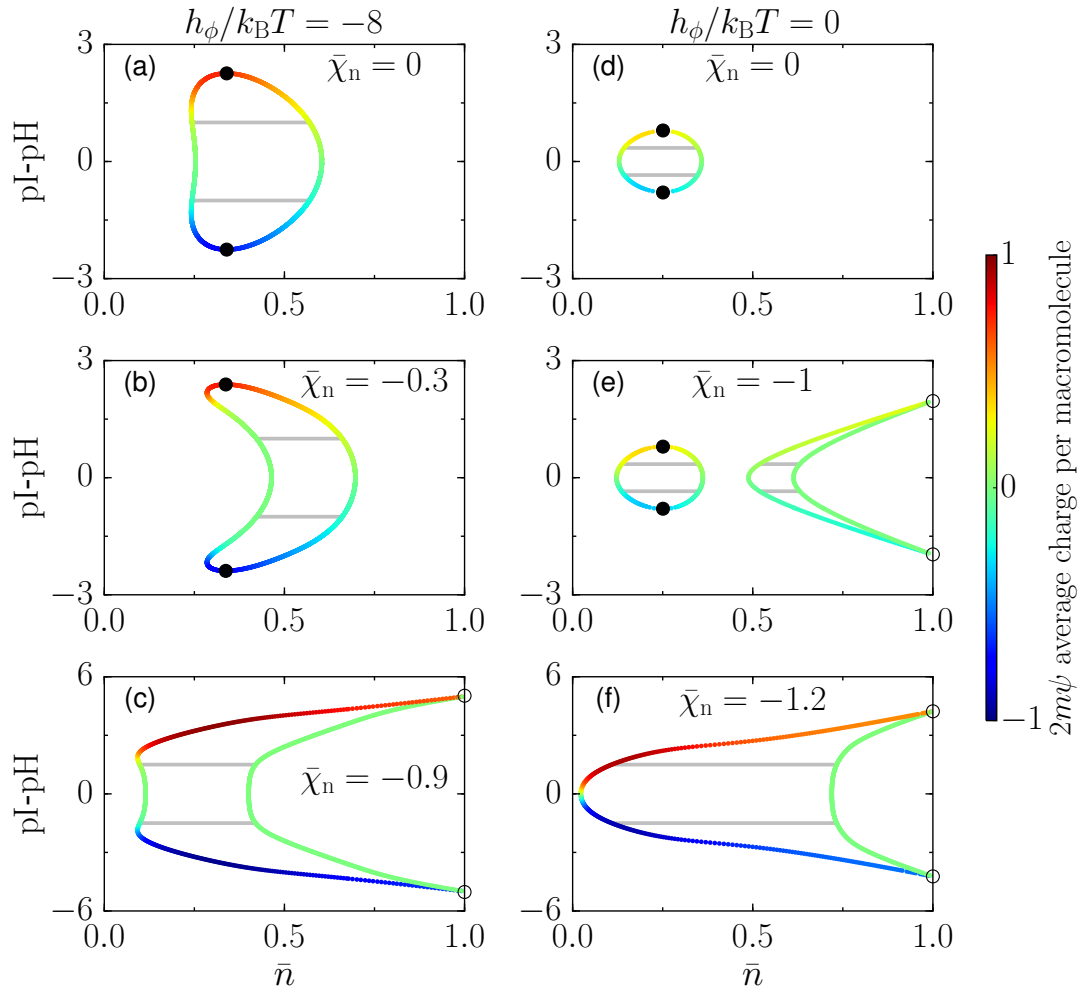

FIGURE G.1: Phase behaviour as a function of pH, the colorbar indicates the value $2m\psi = (n_{M^+} - n_{M^-})/2n$. Parameters $\epsilon = 0.1$ and $\chi_e/k_BT = -3.5$, apply to all panels.

Appendix H

Binding Equilibrium

Let us take the free energy density Eq. 5.1 as starting point and calculate the exchange chemical potentials, $\bar{\mu}_i = \mu_i - \mu_W/\epsilon_i$, leading to

$$\mu_P = \frac{k_{\mathrm{B}}T}{\epsilon_P}\Big(\epsilon_P(\ln\phi_P + 1) - \ln\phi_W - 1 + \epsilon_P w_P - w_W$$
$$+ \chi_{P,W}(\phi_W - \phi_P) - \chi_{PR,W}\phi_{PR}\Big) \quad, \tag{H.1}$$

$$\mu_{PR} = \frac{k_{\mathrm{B}}T}{\epsilon_{PR}}\Big(\epsilon_{PR}(\ln\phi_{PR} + 1) - \ln\phi_W - 1 + \epsilon_{PR}w_{PR} - w_W$$
$$+ \chi_{PR,W}(\phi_W - \phi_{PR}) - \chi_{P,W}\phi_P\Big) \quad, \tag{H.2}$$

$$\mu_R = \frac{k_{\mathrm{B}}T}{\epsilon_R}\Big(\epsilon_R(\ln\phi_R + 1) - \ln\phi_W - 1 + \epsilon_R w_R - w_W$$
$$- \chi_{P,W}\phi_P - \chi_{PR,W}\phi_{PR}\Big) \quad, \tag{H.3}$$

$$\mu_M = \frac{k_{\mathrm{B}}T}{\epsilon_M}\Big(\epsilon_M(\ln\phi_M + 1) - \ln\phi_W - 1 + \epsilon_M w_M - w_W$$
$$- \chi_{P,W}\phi_P - \chi_{PR,W}\phi_{PR}\Big) \quad, \tag{H.4}$$

$$\mu_{MR} = \frac{k_{\mathrm{B}}T}{\epsilon_{MR}}\Big(\epsilon_{MR}(\ln\phi_{MR} + 1) - \ln\phi_W - 1 + \epsilon_{MR}w_{MR} - w_W$$
$$- \chi_{P,W}\phi_P - \chi_{PR,W}\phi_{PR}\Big) \quad, \tag{H.5}$$

where we used $\phi_W = 1 - \sum_{i\neq W}\phi_W$. These chemical potentials will be used to determine the binding equilibrium conditions, which for reactions (5.2) are given by

$$\mu_P + \mu_R = \mu_{PR} \quad, \tag{H.6}$$

$$\mu_M + \mu_R = \mu_{MR} \quad. \tag{H.7}$$

$$\tag{H.8}$$

115

We now substitute the chemical potentials given in Eqs. (H.1)-(H.5) in the binding conditions, leading to

$$\frac{n_P n_R}{n_{PR}} = \frac{v_{PR}}{v_P v_R} \left[\frac{w_{PR} - w_P - w_R}{k_\mathrm{B}T} + \phi_W \left(\frac{\chi_{PR}}{\epsilon_{PR}} - \frac{\chi_P}{\epsilon_P} \right) \right] \quad , \tag{H.9}$$

$$\frac{n_M n_R}{n_{MR}} = \frac{v_{MR}}{v_M v_R} \left[\frac{w_{MR} - w_M - w_R}{k_\mathrm{B}T} \right] \quad . \tag{H.10}$$

If we now define the dissociation constants K_{PR} and K_{MR}, as

$$K_{PR} = \frac{v_{PR}}{v_P v_R} \left(\frac{w_{PR} - w_P - w_R}{k_\mathrm{B}T} \right) \quad , \tag{H.11}$$

$$K_{MR} = \frac{v_{MR}}{v_M v_R} \left(\frac{w_{MR} - w_M - w_R}{k_\mathrm{B}T} \right) \quad , \tag{H.12}$$

and neglect the contributions coming from the interactions in Eq. (H.9) as stated in the main text, we obtain

$$\frac{n_P n_R}{n_{PR}} \simeq K_{PR} \quad , \tag{H.13}$$

$$\frac{n_M n_R}{n_{MR}} = K_{MR} \quad , \tag{H.14}$$

which are the corresponding equations presented in the main text.

Bibliography

[1] B. Alberts, *Molecular biology of the cell* (Garland science, New York, 2017).

[2] J. R. Casey, S. Grinstein, and J. Orlowski, Nature Reviews Molecular Cell Biology **11** (2009).

[3] S. F. Banani, H. O. Lee, A. A. Hyman, and M. K. Rosen, Nature reviews Molecular cell biology **18**, 285 (2017).

[4] Mediran, *Eukaryotic cell (animal)*, (2012) `https://commons.wikimedia.org/wiki/File:Eukaryotic_Cell_(animal).jpg`.

[5] A. A. Hyman, C. A. Weber, and F. Jülicher, Annual Review of Cell and Developmental Biology **30**, 39 (2014).

[6] C. P. Brangwynne, P. Tompa, and R. V. Pappu, Nature Physics **11**, 899 (2015).

[7] Y. Shin and C. P. Brangwynne, Science **357** (2017).

[8] H. Walter and D. E. Brooks, FEBS Letters **361**, 135 (1995).

[9] M. J. Smalley, N. Signoret, D. Robertson, A. Tilley, A. Hann, K. Ewan, Y. Ding, H. Paterson, and T. C. Dale, Journal of Cell Science **118**, 5279 (2005).

[10] T. Schwarz-Romond, C. Merrifield, B. J. Nichols, and M. Bienz, Journal of Cell Science **118**, 5269 (2005).

[11] R. P. Sear, Soft Matter **3**, 680 (2007).

[12] C. P. Brangwynne, C. R. Eckmann, D. S. Courson, A. Rybarska, C. Hoege, J. Gharakhani, F. Jülicher, and A. A. Hyman, Science **324**, 1729 (2009).

[13] A. K. Corsi, B. Wightman, and M. Chalfie, Genetics **200**, 387 (2015).

[14] T. Trcek and R. Lehmann, eLife **6**, e24106 (2017).

[15] D. Ohshima, K. Arimoto-Matsuzaki, T. Tomida, M. Takekawa, and K. Ichikawa, PLOS Computational Biology **11**, 1 (2015).

[16] M. C. Munder, D. Midtvedt, T. Franzmann, E. Nuske, O. Otto, M. Herbig, E. Ulbricht, P. Müller, A. Taubenberger, S. Maharana, L. Malinovska, D. Richter, J. Guck, V. Zaburdaev, and S. Alberti, eLife **5**, e09347 (2016).

[17] K. M. Neugebauer, RNA Biology **14**, 669 (2017).

[18] M. Feric, N. Vaidya, T. S. Harmon, D. M. Mitrea, L. Zhu, T. M. Richardson, R. W. Kriwacki, R. V. Pappu, and C. P. Brangwynne, Cell **165**, 1686 (2016).

[19] C. P. Brangwynne, T. J. Mitchison, and A. A. Hyman, Proceedings of the National Academy of Sciences **108**, 4334 (2011).

[20] P. Li, S. Banjade, H.-C. Cheng, S. Kim, B. Chen, L. Guo, M. Llaguno, J. V. Hollingsworth, D. S. King, S. F. Banani, et al., Nature **483**, 336 (2012).

[21] F. Wippich, B. Bodenmiller, M. G. Trajkovska, S. Wanka, R. Aebersold, and L. Pelkmans, Cell **152**, 791 (2013).

[22] J. Berry, S. C. Weber, N. Vaidya, M. Haataja, and C. P. Brangwynne, Proceedings of the National Academy of Sciences **112**, E5237 (2015).

[23] S. Elbaum-Garfinkle, Y. Kim, K. Szczepaniak, C. C.-H. Chen, C. R. Eckmann, S. Myong, and C. P. Brangwynne, Proceedings of the National Academy of Sciences **112**, 7189 (2015).

[24] D. Zwicker, M. Decker, S. Jaensch, A. A. Hyman, and F. Jülicher, Proceedings of the National Academy of Sciences **111**, E2636 (2014).

[25] A. Molliex, J. Temirov, J. Lee, M. Coughlin, A. Kanagaraj, H. Kim, T. Mittag, and J. Taylor, Cell **163**, 123 (2015).

[26] S. Elbaum-Garfinkle, Y. Kim, K. Szczepaniak, C. C.-H. Chen, C. R. Eckmann, S. Myong, and C. P. Brangwynne, Proceedings of the National Academy of Sciences **112**, 7189 (2015).

[27] H. Zhang, S. Elbaum-Garfinkle, E. M. Langdon, N. Taylor, P. Occhipinti, A. Bridges, C. Brangwynne, and A. Gladfelter, Molecular Cell **60**, 220 (2015).

[28] S. Saha, C. A. Weber, M. Nousch, O. Adame-Arana, C. Hoege, M. Y. Hein, E. Osborne-Nishimura, J. Mahamid, M. Jahnel, L. Jawerth, A. Pozniakovski, C. R. Eckmann, F. Jülicher, and A. A. Hyman, Cell **166**, 1572 (2016).

[29] D. S. W. Protter, B. S. Rao, B. Van Treeck, Y. Lin, L. Mizoue, M. K. Rosen, and R. Parker, Cell reports **22**, 1401 (2018).

[30] T. M. Franzmann, M. Jahnel, A. Pozniakovsky, J. Mahamid, A. S. Holehouse, E. Nüske, D. Richter, W. Baumeister, S. W. Grill, R. V. Pappu, A. A. Hyman, and S. Alberti, Science **359**, eaao5654 (2018).

[31] B. A. Gibson, L. K. Doolittle, M. W. Schneider, L. E. Jensen, N. Gamarra, L. Henry, D. W. Gerlich, S. Redding, and M. K. Rosen, Cell **179**, 470 (2019).

[32] S. Kroschwald, S. Maharana, D. Mateju, L. Malinovska, E. Nüske, I. Poser, D. Richter, and S. Alberti, eLife **4**, e06807 (2015).

[33] C. Rabouille and S. Alberti, Current Opinion in Cell Biology **47**, Cell Organelles, 34 (2017).

[34] J. A. Riback, C. D. Katanski, J. L. Kear-Scott, E. V. Pilipenko, A. E. Rojek, T. R. Sosnick, and D. A. Drummond, Cell **168**, 1028 (2017).

[35] V. Sirri, S. Urcuqui-Inchima, P. Roussel, and D. Hernandez-Verdun, Histochemistry and Cell Biology **129**, 13 (2008).

[36] M. Hanazawa, M. Yonetani, and A. Sugimoto, The Journal of Cell Biology **192**, 929 (2011).

[37] K. A. Burke, A. M. Janke, C. L. Rhine, and N. L. Fawzi, Molecular cell **60**, 231 (2015).

[38] S. Maharana, J. Wang, D. K. Papadopoulos, D. Richter, A. Pozniakovsky, I. Poser, M. Bickle, S. Rizk, J. Guillén-Boixet, T. M. Franzmann, M. Jahnel, L. Marrone, Y.-T. Chang, J. Sterneckert, P. Tomancak, A. A. Hyman, and S. Alberti, Science **360**, 918 (2018).

[39] J. Smith, D. Calidas, H. Schmidt, T. Lu, D. Rasoloson, and G. Seydoux, eLife **5**, e21337 (2016).

[40] E. Griffin, D. Odde, and G. Seydoux, Cell **146**, 955 (2011).

[41] J. V. Gray, G. A. Petsko, G. C. Johnston, D. Ringe, R. A. Singer, and M. Werner-Washburne, Microbiology and Molecular Biology Reviews **68**, 187 (2004).

[42] P. Anderson and N. Kedersha, The Journal of Cell Biology **172**, 803 (2006).

[43] P. Anderson and N. Kedersha, Journal of Cell Science **115**, 3227 (2002).

[44] S. M. Burgess, T. Powers, and J. C. Mell, *Budding yeast saccharomyces cerevisiae as a model genetic organism* (American Cancer Society, 2017), pp. 1–12.

[45] K. L. Shaw, G. R. Grimsley, G. I. Yakovlev, A. A. Makarov, and C. N. Pace, Protein Science **10**, 1206 (2001).

[46] E. K. Khalid, E. E. Babiker, and A. H. EL Tinay, Food Chemistry **82**, 361 (2003).

[47] D. H. G. Pelegrine and C. A. Gasparetto, LWT - Food Science and Technology **38**, 77 (2005).

[48] G. F. Weiller, G. Caraux, and N. Sylvester, Proteomics **4**, 943 (2004).

[49] Y. Lin, D. Protter, M. Rosen, and R. Parker, Molecular Cell **60**, 208 (2015).

[50] T. J. Nott, E. Petsalaki, P. Farber, D. Jervis, E. Fussner, A. Plochowietz, T. D. Craggs, D. P. Bazett-Jones, T. Pawson, J. D. Forman-Kay, and A. J. Baldwin, Molecular Cell **57**, 936 (2015).

[51] S. F. Banani, A. M. Rice, W. B. Peeples, Y. Lin, S. Jain, R. Parker, and M. K. Rosen, Cell **166**, 651 (2016).

[52] C. P. Brangwynne, The Journal of Cell Biology **203**, 875 (2013).

[53] P. Tompa, Trends in Biochemical Sciences **37**, 509 (2012).

[54] P. Tompa, Trends in Biochemical Sciences **27**, 527 (2002).

[55] H.-X. Zhou, V. Nguemaha, K. Mazarakos, and S. Qin, Trends in Biochemical Sciences **43**, 499 (2018).

[56] H.-X. Zhou and X. Pang, Chemical Reviews **118**, 1691 (2018).

[57] T Tanaka, C Ishimoto, and L. Chylack, Science **197**, 1010 (1977).

[58] J. A. Thomson, P Schurtenberger, G. M. Thurston, and G. B. Benedek, Proceedings of the National Academy of Sciences **84**, 7079 (1987).

[59] M. L. Huggins, The Journal of Physical Chemistry **9**, 441 (1941).

[60] P. J. Flory, The Journal of Chemical Physics **10**, 51 (1942).

[61] D. S. Goodsell, Trends in Biochemical Sciences **16**, 203 (1991).

[62] R. Milo and R. Phillips, *Cell Biology by the Numbers* (Garland Science, 2015).

[63] R Koningsveld, H. Chermin, and M Gordon, Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences **319**, 331 (1970).

[64] H. B. Callen, *Thermodynamics and an introduction to thermostatistics* (John Wiley and Sons, New York, 1985).

[65] M. Rubinstein and R. H. Colby, *Polymer physics* (Oxford University Press, Oxford, 2003).

[66] P. Atkins and J. de Paula, *Physical chemistry, eigth edition* (W.H. Freeman and Company, New York, 2006).

[67] R. P. Sear and J. A. Cuesta, Phys. Rev. Lett. **91**, 245701 (2003).

[68] S. A. Safran, *Statistical thermodynamics of surfaces, interfaces, and membranes* (Addison-Wesley, Reading, 1994).

[69] R. A. Alberty, *Thermodynamics of Biochemical Reactions* (John Wiley and Sons, New Jersey, 2003).

[70] Y. Albeck and M. Gitterman, Philosophical Magazine B **56**, 881 (1987).

[71] L. R. Corrales and J. C. Wheeler, The Journal of Chemical Physics **91**, 7097 (1989).

[72] M. Gitterman, Journal of Statistical Physics **58**, 707 (1990).

[73] A. Matsuyama and F. Tanaka, Phys. Rev. Lett. **65**, 341 (1990).

[74] A. P. Minton, Journal of Cell Science **119**, 2863 (2006).

[75] G. Rivas and A. P. Minton, Biophysical Reviews **10**, 241 (2018).

[76] R. A. Alberty, Chemical reviews **94**, 1457 (1994).

[77] R. A. Alberty, The Journal of Chemical Thermodynamics **29**, 501 (1997).

[78] A. Hernández-Vega, M. Braun, L. Scharrel, M. Jahnel, S. Wegmann, B. T. Hyman, S. Alberti, S. Diez, and A. A. Hyman, Cell reports **20**, 2304 (2017).

[79] J. Nitsche, G. Teletzke, L. Scriven, and H. Davis, Fluid Phase Equilibria **17**, 243 (1984).

[80] Y. S. Soh, J. H. Kim, and C. C. Gryte, Polymer **36**, 3711 (1995).

[81] D. Hildebrandt and D. Glasser, The Chemical Engineering Journal and the Biochemical Engineering Journal **54**, 187 (1994).

[82] A. Joseph, C. M. Sands, P. D. Hicks, and H. W. Chandler, Fluid Phase Equilibria **431**, 34 (2017).

[83] C. B. Barber, D. P. Dobkin, D. P. Dobkin, and H. Huhdanpaa, ACM Trans. Math. Softw. **22**, 469 (1996).

[84] A. Patel, H. O. Lee, L. Jawerth, S. Maharana, M. Jahnel, M. Y. Hein, S. Stoynov, J. Mahamid, S. Saha, T. M. Franzmann, A. Pozniakovski, I. Poser, N. Maghelli, L. A. Royer, M. Weigert, E. W. Myers, S. Grill, D. Drechsel, A. A. Hyman, and S. Alberti, Cell **162**, 1066 (2015).

[85] V. Talanquer, The Journal of Chemical Physics **96**, 5408 (1992).

[86] S. Ung and M. F. Doherty, Chemical Engineering Science **50**, 3201 (1995).

[87] J. M. Powers and S. Paolucci, American Journal of Physics **76**, 848 (2008).

[88] S. H. Friedberg, A. J. Insel, and L. E. Spence, *Linear Algebra*, 2nd (Prentice Hall, New Jersey, 1989).

[89] J. Rasper and W. Kauzmann, Journal of the American Chemical Society **84**, 1771 (1962).

[90] Y. Okada and F. Tanaka, Macromolecules **38**, 4465 (2005).

[91] B. Garcia-Moreno, Journal of Biology **8** (2009).

[92] P. Mitchell, Nature **191**, 144 (1961).

[93] O. Adame-Arana, C. A. Weber, V. Zaburdaev, J. Prost, and F. Jülicher, arXiv1910.06733 (2019).

[94] R. Buck, S. Rondinini, A. Covington, F. Baucke, C. Brett, M. Camões, M. Milton, T. Mussini, R. Naumann, K. Pratt, R. Spitzer, and G. Wilson, Pure and Applied Chemistry **74**, 2169 (2002).

[95] I. Mills, T. Cvitaš, K. Homann, N. Kallay, and K. Kuchitsu, *Quantities, Units and Symbols in Physical Chemistry* (Blackwell Science, Oxford, 1993).

[96] D. Marx, M. E. Tuckerman, J. Hutter, and M. Parrinello, Nature **397**, 601 (1999).

[97] J. M. Headrick, E. G. Diken, R. S. Walters, N. I. Hammer, R. A. Christie, J. Cui, E. M. Myshakin, M. A. Duncan, M. A. Johnson, and K. D. Jordan, Science **308**, 1765 (2005).

[98] R. Scheu, B. M. Rankin, Y. Chen, K. C. Jena, D. Ben-Amotz, and S. Roke, Angewandte Chemie International Edition **53**, 9560 (2014).

[99] D. Harries, S. May, and A. Ben-Shaul, Soft Matter **9**, 9268 (2013).

[100] A. Pfennig and A. Schwerin, Fluid Phase Equilibria **108**, 305 (1995).

[101] M. Vis, V. F. Peters, R. H. Tromp, and B. H. Erne, Langmuir **30**, 5755 (2014).

[102] J. Llopis, J. M. McCaffery, A. Miyawaki, M. G. Farquhar, and R. Y. Tsien, Proceedings of the National Academy of Sciences **95**, 6803 (1998).

[103] M. Grabe and G. Oster, The Journal of General Physiology **117**, 329 (2001).

[104] N. Sperelakis, *Cell physiology sourcebook: a molecular approach* (Elsevier, 2001).

[105] J.-P. Hansen and I. R. McDonald, *Theory of simple liquids* (Elsevier, 1990).

[106] M. Doi, *Soft matter physics* (Oxford University Press, 2013).

[107] H. DeVoe, *Thermodynamics and chemistry, Second Edition* (Prentice Hall Upper Saddle River, NJ, New Jersey, 2001).

[108] R. Milo, BioEssays **35**, 1050 (2013).

[109] M. Strzelecka, S. Trowitzsch, G. Weber, R. Lührmann, A. C. Oates, and K. M. Neugebauer, Nature structural & molecular biology **17**, 403 (2010).

[110] C. DeRenzo, K. J. Reese, and G. Seydoux, Nature **424**, 685 (2003).

[111] R. J. Cheeks, J. C. Canman, W. N. Gabriel, N. Meyer, S. Strome, and B. Goldstein, Current Biology **14**, 851 (2004).

[112] C. F. Lee, C. P. Brangwynne, J. Gharakhani, A. A. Hyman, and F. Jülicher, Phys. Rev. Lett. **111**, 088101 (2013).

[113] P. Thandapani, T. R. O'Connor, T. Bailey, and S. Richard, Molecular Cell **50**, 613 (2013).

[114] M. Fu and P. J. Blackshear, Nature Reviews Immunology **17**, 130 (2017).

[115] S. R. De Groot and P. Mazur, *Non-equilibrium thermodynamics* (Dover, New Jersey, 1984).

[116] J. W. Cahn and J. E. Hilliard, The Journal of Chemical Physics **28**, 258 (1958).

[117] D. Updike and S. Strome, Journal of Andrology **31**, 53 (2010).

[118] J. W. Gibbs, *The collected works of J. Willard Gibbs.* (Yale Univ. Press, 1948).

Versicherung

Hiermit versichere ich, Omar Adame Arana, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt. Diese Arbeit wurde unter der wissenschaftlichen Betreuung von Prof. Dr. Frank Jülicher am Max-Planck Institut für Physik komplexer Systeme in Dresden angefertigt. Ich erkläre hiermit, dass keine früheren erfolglosen Promotionsverfahren stattgefunden haben. Ich erkenne die Promotionsordnung der Fakultät für Mathematik und Naturwissenschaften der Technische Universität Dresden an.

_____

Unterschrift
Dresden, November 2019