OPEN ACCESS
OPUS AUGSBURG

Universität Augsburg
Universitätsbibliothek

Emergent behavior in non-dilute reacting mixtures away from equilibrium: From oscillations and excitability to selection and replication of sequences

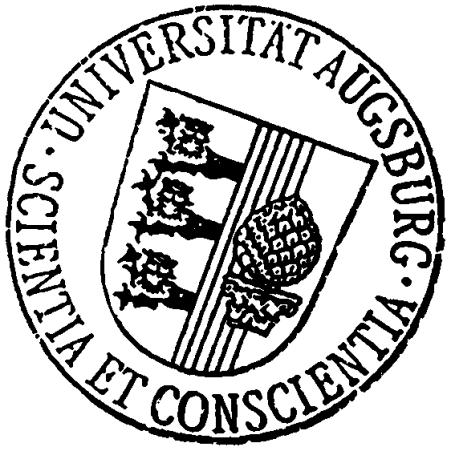
Ivar Svalheim Haugerud

Angaben zur Veröffentlichung / Publication details:

Haugerud, Ivar Svalheim. 2026. "Emergent behavior in non-dilute reacting mixtures away from equilibrium: From oscillations and excitability to selection and replication of sequences." Augsburg: Universität Augsburg.

Nutzungsbedingungen / Terms of use: **CC BY-NC 4.0**

Dieses Dokument wird unter folgenden Bedingungen zur Verfügung gestellt: / This document is made available under these conditions:
CC-BY-NC 4.0: Creative Commons: Namensnennung - Nicht kommerziell
Weitere Informationen finden Sie unter: / For more information see:
https://creativecommons.org/licenses/by-nc/4.0/deed.de



Emergent behavior in non-dilute reacting mixtures away from equilibrium

From oscillations and excitability to selection and replication of sequences

von

Ivar Svalheim Haugerud

Dissertation zur Erlangung des akademischen Grades

Doctor rerum naturalium

eingereicht an der
Mathematisch-Naturwissenschaftlich-Technischen
Fakultät der Universität Augsburg

Universität
Augsburg
University

October 2025

Erstgutachter: Prof. Dr. Christoph A. Weber
Zweitgutachter: Prof. Dr. Markus Heyl
Drittgutachter: Prof. Dr. Erwin Frey
Tag der Mündlichen Prüfung: 19. Dezember 2025

Abstract

Creating life in the laboratory would be a revolutionary event in science, providing insight into universal principles of life and enabling novel applications through synthetic biological systems. A central challenge in creating life anew is to understand how complex behaviors can emerge from the reactions and interactions of much simpler constituent components.

The emergent behaviors must tackle several fundamental obstacles, including accelerating otherwise unfavorable reactions, regulating chemistry, selecting and stabilizing sequences, and maintaining their information through accurate replication. Understanding how such behaviors can emerge requires identifying the physical mechanisms and non-equilibrium conditions that support their emergence. In this thesis, we argue that the interplay between chemistry and non-dilute mixtures away from equilibrium, along with their ability to phase separate and form condensates, can allow for rich emergent behavior, paving the path towards creating life.

To investigate this idea, we employ concepts from non-equilibrium thermodynamics and statistical mechanics to develop models of non-dilute, interacting mixtures driven away from equilibrium. These models allow us to explore how interactions and condensed phases alter reaction kinetics and transport, and whether they can give rise to new behavior.

We show that wet–dry cycles can accelerate thermodynamically hindered reaction pathways through intermolecular interactions, which can be optimized by varying the cycling frequency. Orthogonally, when active droplets are driven far from equilibrium, the system itself can create self-sustaining oscillations, behaving as an excitable medium that supports both transient and persistent chemical activity. Such behavior is absent in a homogeneous mixture, making excitability and self-sustained oscillations an emergent property of active droplets.

Furthermore, we investigate systems composed of many interacting and reacting components by developing a thermodynamic description of oligomerization coupled to phase separation. We demonstrate that condensed phases impose sequence-specific selection pressures, favoring distinct motifs under different non-equilibrium conditions. Building on this framework, we show that condensates can enhance the fitness and copying fidelity of templated ligation, allowing information copying within Eigen's error threshold. The coupling between sequence distributions and condensate properties provides a minimal genotype–phenotype link, reminiscent of living systems.

Taken together, our results suggest that non-dilute interacting mixtures, leading to phase separation, can give rise to emergent behaviors relevant to the emergence of life.

Zusammenfassung

Die Erschaffung von Leben im Labor wäre ein bahnbrechendes Ereignis in der Wissenschaft, das Einblicke in universelle Prinzipien des Lebens liefern und neue Anwendungen durch synthetische biologische Systeme ermöglichen würde. Eine zentrale Herausforderung bei der Entstehung neuen Lebens besteht darin, zu verstehen, wie aus den Reaktionen und Wechselwirkungen viel einfacherer Bestandteile komplexe Verhaltensweisen entstehen können.

Damit Leben entstehen kann, müssen chemische Systeme mehrere grundlegende Eigenschaften vorweisen, darunter die Beschleunigung ansonsten ungünstiger Reaktionen, die Regulierung ihrer Dynamik, die Auswahl und Stabilisierung von informationshaltigen Molekülen wie Oligomeren und die Aufrechterhaltung ihrer Informationen durch ausreichend genaue Replikation. Um zu verstehen, wie solche Eigenschaften entstehen können, müssen die physikalischen Mechanismen und Nichtgleichgewichtsbedingungen identifiziert werden, die sie unterstützen. In dieser Arbeit argumentieren wir, dass das Zusammenspiel zwischen Chemie und nicht verdünnten Gemischen außerhalb des Gleichgewichts emergentes Verhalten ermöglichen und den Weg zur Schaffung von Leben ebnen kann.

Um diese Situation zu studieren, verwenden wir Konzepte aus der Nichtgleichgewichtsthermodynamik und der statistischen Mechanik, um Modelle für nicht verdünnte, interagierende Gemische zu entwickeln, die vom Gleichgewicht abgelenkt werden. Diese Modelle ermöglichen es uns zu untersuchen, wie kondensierte Phasen die Reaktionskinetik und den Transport verändern und ob sie zu neuem emergentem Verhalten führen können.

Wir zeigen, dass Nass-Trocken-Zyklen thermodynamisch gehemmte Reaktionswege durch intermolekulare Wechselwirkungen beschleunigen können, die durch Veränderung der Zyklusfrequenz optimiert werden können. Im Gegensatz dazu kann das System selbst, wenn aktive Tröpfchen weit vom Gleichgewicht entfernt sind, selbsttragende Schwingungen erzeugen und sich wie ein erregbares Medium verhalten, das sowohl vorübergehende als auch anhaltende Aktivität unterstützt. Ein solches Verhalten tritt in einem homogenen Gemisch nicht auf, was zeigt, dass Erregbarkeit eine emergente Eigenschaft der Phasentrennung ist.

Wir diskutieren dann Systemen mit vielen interagierenden und reagierenden Komponenten über, für die wir eine thermodynamische Beschreibung der reversiblen Oligomerisierung in Verbindung mit Phasentrennung entwickeln. Wir zeigen, dass kondensierte Phasen sequenzspezifische Selektionsdrücke ausüben, die unter verschiedenen Nichtgleichgewichtsbedingungen bestimmte Motive begünstigen. Aufbauend auf dieser thermodynamischen Beschreibung finden wir, dass Kondensate die Fitness und Kopiergenauigkeit der templatgesteuerten Ligation verbessern können,

wodurch die Informationserhaltung jenseits des von Manfred Eigen gefundenen Schwellwertes ermöglicht wird. Die Kopplung zwischen Sequenzverteilungen und Kondensateigenschaften stellt eine minimale Verbindung zwischen Genotyp und Phänotyp her.

Zusammengenommen deuten die Ergebnisse darauf hin, dass nicht verdünnte, interagierende Gemische, die zu einer Phasentrennung führen, ein physikalisches Prinzip darstellen, das emergente Verhaltensweisen unterstützt, die für die Entstehung von Leben relevant sind.

Acknowledgments

It takes a village to write a PhD thesis. But nothing in this village would have worked without the endless support, knowledge, patience, and infectious enthusiasm of Christoph. No matter what I brought to our meetings in terms of variable progress and mood, I would leave with a head full of ideas and an enthusiasm for our projects, without which nothing would have progressed. I can not overstate how much I have learned from you over the last four years.

Whenever I've had the impulse to give up and depart from the village, it has always been the people who made me stay. This thesis would never have been written without Pranay, Nick, Kafa, Susanne, Advait, Hidde, Gaetano, Samuel, and Leonardo. I could not have asked for better colleagues. I am also very grateful for the advice from our senior citizens: Giacomo, Sudarshana, Jonathan, and Jacqueline, and for everything outside of work: Jonah, Jean, Akhil, Alexei, Nan, Timo, Marianna, Mareike, Bhavesh, Filipo, Atharva, Susanne Graf, Elisabeth, FCA, Riegele, the welcome service, and many, many more. You've all made the four years in Augsburg wonderful.

A village, of course, requires paperwork. Given the number of documents I've signed over the years, it is only a miracle worker like Christine who could ensure that I am still employed and allowed to attend and (eventually) get reimbursed for conferences.

Jeg vil også takke landsbyen utenfor landsbyen i kronologisk rekkefølge: Oscar, Marlon, Taran, Håkon, Sveinung, Astrid, Ylva, Eline, Erlend, Torbjørn, Metin, Cecilie, Eivind, Anna, Hilde, Oslo sportslag, osv.

Selve grunnmuren i landsbyen er selvfølgelig mamma, pappa og søster. Takk for tålmodigheten, støtten og hjelpen, både her nede og hjemme i Oslo

List of publications

The research presented in this thesis has resulted in the following publications and manuscripts:

[1] *"Nonequilibrium wet–dry cycling acts as a catalyst for chemical reactions".*
Ivar S. Haugerud, Pranay Jaiswal, and Christoph A. Weber.
The Journal of Physical Chemistry B 128.7 (2024): 1724-1736 .

[2] *"The interplay between biomolecular assembly and phase separation.".*
Giacomo Bartolucci, Ivar S. Haugerud, Thomas C. T. Michaels, and Christoph A. Weber.
eLife 13 (2024).

[3] *"A roadmap toward the synthesis of life.".*
Christine M.E. Kriebisch, Olga Bantysh, Lorena B. Pellejero, Andrea Belluati, Eva Bertosin, Kun Dai, ..., Ivar S. Haugerud, ..., Kerstin Göpfrich, Christoph A. Weber, and Job Boekhoven.
Chem 11.3 (2025)

[4] *"Theory for sequence selection via phase separation and oligomerization.".*
Ivar S. Haugerud, Giacomo Bartolucci, Dieter Braun and Christoph A. Weber.
arXiv preprint:2410.08778 (Submitted on 11th of October 2024).

[5] *"Excitability and oscillations of active droplets.".*
Ivar S. Haugerud, Hidde D. Vuijk, Job Boekhoven, and Christoph A. Weber.
arXiv preprint:2503.11604 (submitted 15th of March 2025).

[6] *"Harvesting chemical power from cyclic environments"*
Pranay Jaiswal, Ivar S. Haugerud, Hidde D. Vuijk, Christoph A. Weber.
arXiv preprint:2510.10266 (Submitted on 11th October 2025)

The corresponding references are listed in the full bibliography, which begins on page 159.

Contents

1 An introduction to the emergence of life **5**

1.1 The early Earth, the emergence of life, and the last universal common ancestor . . . 5

1.2 The RNA hypothesis and the molecular origins of life . . . 6

1.3 De novo life . . . 9

1.4 Compartments in extant life and the protocell . . . 11

1.5 Outline and scope of thesis . . . 12

2 Theory of chemical reactions in non-dilute inhomogeneous systems **15**

2.1 Physics of interacting mixtures . . . 15

2.1.1 Thermodynamics of mixtures in the Gibbs ensemble . . . 15

2.1.2 From the Gibbs to the Helmholtz ensemble . . . 16

2.1.3 Thermodynamic stability and the free energy density . . . 17

2.1.4 Relaxation to thermodynamic equilibrium . . . 19

2.1.5 Conditions of phase coexistence . . . 22

2.2 Thermodynamics of chemical reactions in homogeneous mixtures . . . 24

2.2.1 Chemical reaction networks . . . 25

2.2.2 Chemical reactions towards equilibrium . . . 26

2.2.3 Retrieving the law of mass action . . . 28

2.2.4 Breaking detailed balance of the rates . . . 29

2.3 Dynamics of chemically active mixtures . . . 30

2.3.1 Reaction-diffusion in dense systems . . . 30

2.3.2 Chemical kinetics at phase coexistence . . . 32

2.3.3 Non-equilibrium steady states . . . 34

3 Accelerating chemical reactions through external reservoirs **37**

3.1 Introduction . . . 37

3.2 Theory for wet-dry cycles with chemical reactions . . . 39

3.3 Thermodynamics of evaporation/condensation . . . 43

3.4 Evaporation/condensation of chemically reacting mixtures . . . 44

3.5 Accelerating chemical reactions through evaporation and condensation 45

3.6 Chemical reactions subject to wet-dry cycles 46
3.7 Comparison to reactant cycles 49
3.8 Discussion and conclusion 51

4 Excitability and self sustaining oscillations of active droplets 55
4.1 Introduction 55
4.2 Theory of active droplets 57
4.3 Saddle-node bifurcation to a bistable droplet 59
4.3.1 Bistability of a ternary system 59
4.3.2 Bistability of active phase separating polymers 62
4.4 An excitable active droplet 66
4.5 An oscillating active droplet 67
4.6 Entropy production of an excitable droplet 71
4.7 Emulsions of excitable and oscillating droplets 73
4.8 Discussion and conclusion 79

5 Selecting sequence patterns through phase separation 83
5.1 Introduction 83
5.2 Theory for oligomerization at phase equilibrium 86
5.3 Sequence evolution toward thermodynamic equilibrium 89
5.4 Sequence distribution at thermodynamic equilibrium 92
5.5 Sequence selection by non-equilibrium fragmentation 99
5.6 Discussion and conclusion 106

6 The emergence of a high fidelity replicator 109
6.1 Introduction 109
6.2 Modeling templated ligation in phase-separated systems 112
6.3 bistability of sequence patterns 116
6.4 Surviving the error threshold 118
6.5 Discussion and conclusion 122

7 Outlook and conclusion 125
7.1 Sequence selection through wet-dry cycles 125
7.2 Excitable media without membranes 126
7.3 Advancing phase-separated sequence kinetics 127
7.4 Further theoretical steps towards de novo life 129
7.5 Conclusion 130

8 Appendices 135
A Flory-Huggins free energy 135
B Deriving the exchange flux in the fast diffusion limit 137

C Reaction scheme with an explicit fuel component 138
D Deriving the normal for of bistable droplets 139
E Deriving the normal form of excited droplets 141
F Deriving the the critical fuel strength 142
G Full phase diagrams of excitable and excited droplets 145
H Determining sequence-sequence interaction strengths 146
I Details on cooperativity and rank distributions 148
J Effects of ligation on sequence distribution away from equilibrium . . 150
K Net oligomerization rate with non-equilibrium fragmentation 151
L Gillespie-algorithm for sequence kinetics 152
M Parameter tables for figures: 155
M.1 Parameters for chapter 3 155
M.2 Parameters for chapter 4 156
M.3 Parameters for chapter 5 157
M.4 Parameters for chapter 6 158

Bibliography **159**

Chapter 1

An introduction to the emergence of life

Dirty underwear and grains of wheat, sealed together in a box for 21 days, should, according to the recipe [7], produce two fully grown mice. In practice, however, following this or similar recipes [8] for turning *dead* matter *alive* has proven far more difficult than the instructions suggest [3]. Nevertheless, such a transition must at some point have occurred.

In this chapter, we summarize the current knowledge about the emergence of life on Earth and in the laboratory. We further discuss leading theories on how life could have emerged and might emerge again, the role of condensates in biological and prebiotic systems, and the key outstanding challenges to our understanding. Lastly, the structure and scope of the thesis are given.

1.1 The early Earth, the emergence of life, and the last universal common ancestor

The emergence of life on Earth occurred relatively shortly after the formation of the planet. Geochemical evidence indicates its formation approximately 4.54 ± 0.05 billion years ago [9]. Earth's formation occurred by accretion of mass from the solar system, resulting in high temperatures and a dilute atmosphere [10]. The last 10% of the Earth's mass is believed to have been gained through a collision with a planet-sized body, forming the moon in the process [11]. Over the time span of 200 million years, the Earth cooled, transitioning to wet and clement conditions, which stabilized the crust and led to early tectonic activity [12].

The resulting conditions on early Earth were highly different from what we observe today. Temperature estimates span from near freezing in certain freshwater environments [13] to in excess of 300°C in hydrothermal vents [14]. The temperature

gradients were accompanied by acidic to strongly alkaline pH conditions, with alkaline environments appearing most common [15]. Elevated levels of carbon dioxide likely dominated the atmosphere, with nitrogen and methane also present in high quantities [15, 16]. It remains disputed whether the surface of the Earth was covered in water or not [17]. The combination of a much weaker ozone layer and strong solar activity resulted in much higher levels of biologically harmful UV radiation and ionizing radiation [18].

Despite, or perhaps because, of such a hostile environment, life emerged. The earliest evidence of life has been dated through microfossils in hydrothermal vent deposits as early as 4.28 billion years ago [19]. Measurements of isotopically light carbon in 4.1-billion-year-old minerals further indicate early biological processes [20]. The most widely accepted fossil evidence points to its emergence by at least 3.7–3.8 billion years ago [16]. An illustration of the timeline is shown in Fig. 1.1. The earliest hypothesized life-form is known as LUCA.

The last universal common ancestor, known as LUCA, is a hypothesized common ancestral cell to all extant life. The exact time frame of LUCA is disputed, with the earliest estimates as early as 4.2 billion years ago [21]. LUCA was a complex organism. It was an anaerobic, thermophilic prokaryote [22], with a genome of approximately 2.5 megabases, encoding around 2,600 proteins [21]. LUCA used molecular hydrogen as an electron donor, had a sophisticated metabolism [23], and used the CO_2-rich atmosphere as a carbon source [22]. At the prebiotic, hostile early Earth, the complexity of LUCA emerged in around a couple of hundred million years. The leading theory for how complex DNA-based life, like LUCA, could have emerged is that it evolved from an RNA-based life form.

1.2 The RNA hypothesis and the molecular origins of life

In extant life, genetic information is stored in the DNA and transcribed into RNA, which is translated into functional proteins. This is known as the central dogma of cell biology [25, 26]. Proteins need DNA to make RNA, and RNA is needed to make proteins. But since proteins and DNA rely on each other for their existence, what came first? DNA (information) or proteins (function)? This poses a chicken-and-egg problem, as it seems impossible for one to emerge without the other, and the simultaneous emergence of these different and complex macromolecules appears highly unlikely.

The RNA hypothesis solves this dilemma by proposing that life relied on self-replicating RNA for both function and information storage [27, 28]. The chicken

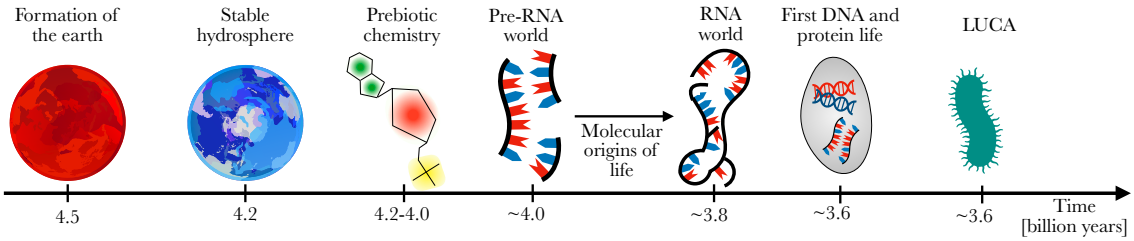


Figure 1.1: **Time-line of the origin of life:** From the stabilization of the Earth's hydrosphere, it took around 600 million years or less for complex life in the form of LUCA to emerge. The transition from a world of prebiotic chemistry consisting of nucleotides and non-functional RNA to the RNA world, where RNA can self-replicate, is called the molecular origins of life. In the RNA world, RNA is used for both information storage and function in the form of simple ribozymes. The exact time estimates for each period have changed over time and vary with the methods used. The figure is recreated based on Ref. [24].

is the egg. RNA is made up of four different nucleotides (A, U, G, and C), and the order of these nucleotides forms a linear code, allowing for information storage [29, 30]. This code determines the possible folding structures of an RNA, and hence its functions [31, 28]. One such possible function is to catalyze chemical reactions, meaning RNA can speed up specific chemical processes just like protein enzymes, without being consumed in the process [32]. Not only can RNA perform functions of both DNA and proteins, but it can form from simple molecules present at early Earth [33, 34] under alkaline conditions [35, 36, 37]. Although the RNA world hypothesis is an accepted theory for the origin of life, the transition from a world of unorganized and non-functioning RNA to one where RNA is complex enough to self-replicate is far from trivial. This transition is called the molecular origins of life.

Despite the properties of RNA to both store information and achieve function being known, there are significant challenges to understanding its emergence and existence without biological machinery [27]. RNA is chemically more prone to degradation compared to proteins and DNA [38]. In water, RNA can self-cleave its backbone, fragmenting the RNA strand, making it inherently unstable [39]. Furthermore, fragmentation is typically faster than ligation of oligomers [40, 41], resulting in long oligomers being heavily suppressed.

If an oligomer becomes long and stable, its information must be copied accurately enough to overcome information loss through mutations. The longer the oligomer is, the more accurate must the copying machinery be. In extant life the problem of mutations is solved by highly specialized enzymes that read and repair errors, allowing for information storage over millions of base-pairs [42]. A major challenge discussed by M. Eigen [43, 44], is that mutations put an upper bound on RNA lengths where information can be maintained. For realistic mutation rates, this upper bound is much shorter than the shortest length of error-correcting enzymes [44].

An additional fundamental problem is the exponentially large number of different possible sequences even for the minimal sequence length needed for functionalities [45]. For four nucleotides, the number of possible sequences exceeds the number of atoms in the universe at a length of only ~ 135 nucleotides [46]. The vast number of configurations rules out the possibility of a uniform exploration of sequence space. This problem renders the need for a strong selection mechanism at the molecular origin of life, which would allow a much smaller dynamic subspace to be explored at a time.

Overcoming these obstacles in a prebiotic setting requires mechanisms that can stabilize RNA against degradation, enhance ligation relative to fragmentation, and introduce rudimentary forms of accurate information-copying and selection. Such processes would need to emerge from the underlying chemistry and environment itself, providing a bridge between fragile prebiotic polymers and the robust function and information copying of modern biology [47, 48, 49, 50, 51]. These considerations raise the central question of how such barriers could have been crossed without enzymes. This has driven experimental efforts to find purely chemical and physical pathways to stabilize, replicate, and enrich functional molecules.

A landmark step in this direction was the seminal paper by S.L. Miller [52]. Miller found that the building blocks of proteins, amino acids, can spontaneously form from simple inorganic compounds under non-equilibrium conditions thought to resemble those of early Earth. Since then, various mechanisms based on the non-equilibrium conditions presumably acting on early Earth, such as salt and temperature gradients [53, 54], wet-dry [55, 56] or freeze-thaw cycles [57], shear-driven replication [58, 59], accumulation at liquid-vapor interfaces [60] or mineral surfaces [61], have been individually found to facilitate the self-assembly of prebiotic chemistry. It still remains unclear how prebiotic macromolecules like RNA could have formed with sufficient length, sequence specificity, and concentration to perform functional roles in prebiotic environments.

The field of origin of life research has been criticized for being unscientific, arguing that theories like the RNA-world are unfalsifiable [62], and that no agreed definition of life makes the goal of the research vague at best [63]. We might never uncover exactly how life emerged on Earth approximately four billion years ago, but understanding how life-like functions can emerge from simple chemistry proposes not just a fundamental scientific question, but a novel avenue of potential applications. With this in mind, we should perhaps follow Feynman's advice of creating a thing in order to understand it [64].

1.3 De novo life

The two main approaches to creating minimalistic systems with life-like functions in the laboratory are top-down and bottom-up [3, 65]. The top-down approach starts with a living system and breaks it down into smaller, simpler, and more manageable parts to explore how far extant life can be reduced while still sustaining essential functions. The bottom-up approach, on the other hand, aims to assemble complex systems with life-like functionalities from well-defined, simpler components. This can involve synthetic, non-biological components, or the canonical building blocks of extant life. While top-down approaches work by reduction, bottom-up efforts attempt to construct life de novo.

Being able to synthesize life in the laboratory might have consequences similar to synthesizing organic molecules [66], turning the origin of life from an academic question into an avenue of potential applications. Though still speculative, parallels can be drawn between the consequences of synthesizing organic molecules and the possibilities of synthesizing life. It was initially thought that only living organisms could produce organic molecules, until they were obtained synthetically. This breakthrough led to the field of organic chemistry, with applications ranging from medicine to materials [66]. The possible potential applications of synthesizing life are difficult to predict. But it is not hard to imagine that the evolution of an initially simple system to survive in a designed environment could have widespread and far-reaching consequences. For example, could the breakdown of microplastics [67] or efficient carbon capture and recycling [68] be selected for, creating a potentially revolutionary approach to challenging questions. Furthermore, the ability to synthesize life will help us find generalizing features of life. A more general understanding of what life is will aid in identifying bio-signatures elsewhere in the universe.

Despite not agreeing on a definition of *life*, the bottom up approach can create out-of-equilibrium systems with properties that are regarded as necessities for *life* [3]. S. Otto argues that under the right conditions, the integration of metabolism, self-replication, and compartmentalization is sufficient for life to emerge [69, 66]. If these three ingredients are maintained away from equilibrium in an environment suitable for open-ended evolution, then life will emerge, as depicted in Fig. 1.2. The mechanisms to achieve the three aforementioned functions have to be simpler than extant life, but can increase in complexity as the system evolves.

These ideas were inspired by key experiments performed in the lab of S. Otto, where they have designed a self-replication system of oligomers that drive the formation of molecules that assemble themselves [66]. The oligomers are reversibly formed from peptide-based building blocks, assembling into fibers that promote the oligomerization process, thereby enhancing their own self-replication. Through ex-

ternal mechanical strain, the fibers can break. The number of ends increases with each braking event, which means the number of growth-sites also increases, leading to exponential growth. Significant challenges remain. Two main ones are creating stable compartments and achieving open-ended evolution with continuous variation and selection. These ongoing efforts aim to bridge the gap between simple chemical replicators and fully autonomous synthetic life.

For open-ended evolution to occur, the system must have a high enough fidelity to avoid rapid information loss through destructive mutations, yet be adaptive enough to explore a wide chemical landscape [66]. Open-ended evolution refers to a system that does not reach a non-equilibrium steady state, but instead evolves new molecular structures, functions, or interactions over time, allowing the system to increase its complexity and adapt to environmental changes [70]. In a prebiotic context, genotype and phenotype are intimately linked, since the sequence of a polymer not only stores information but also directly dictates its chemical and structural functions. [44]. In a prebiotic setting, the tight coupling between genotype and phenotype means that sequence changes immediately alter molecular function, leaving little room for neutral variation. This makes evolution both highly efficient in selecting functional sequences and simultaneously fragile, as many mutations are strongly deleterious [43, 71, 24]. Achieving a balance between high fidelity while remaining adaptive requires highly accurate replication below the error threshold, especially as the system's informational complexity grows [43].

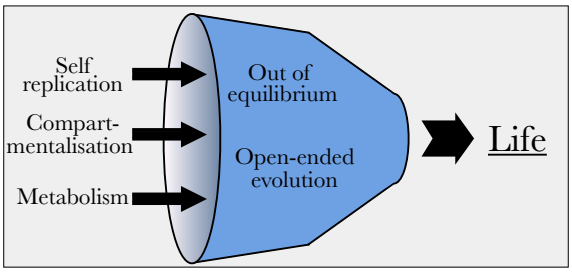


Figure 1.2: **Minimal ingredients for life:** S. Otto argues that a system capable of self-replication, compartmentalization, and metabolism will, under the right conditions (out of equilibrium, open-ended evolution), lead to de novo life. Figure recreated based on Ref. [66].

Compartmentalization has been proposed as a method for stabilizing evolutionary dynamics to achieve a balance between high fidelity and adaptability [72]. Compartments allow for a more stable environment and protection against parasites. Parasites do not produce their own metabolic products or replication machinery, instead, they rely on what their hosts provide. Without a compartment, the products of catalytic activity will aid all replicators, not just the ones that generated the products. In the presence of a compartment, only the replicator that produced the products will benefit from them. Compartments, therefore, can protect against "kleptoparasites", which have been shown to emerge in synthetic systems [73]. The protective role of compartments has been further observed in RNA-based systems [74, 75] and

has been shown to enhance the rate of adaptation during evolution [76]. By examining how living cells compartmentalize, we may uncover organizational principles relevant to abiotic compartments.

1.4 Compartments in extant life and the protocell

A macromolecule typically requires only about 0.01 seconds to diffuse across the 1 µm span of a cell [77]. Without mechanisms for spatial organization, cellular components would quickly mix uniformly. For many years, such an organization was thought to depend dominantly on membrane-bound compartments, where lipid bilayers, with hydrophilic head groups facing outward and hydrophobic tails facing inward, create a separation between the interior and exterior environments.

This view shifted after 2009, when Brangwynne, Jülicher, Hyman, and coworkers observed objects in C. elegans exhibiting behavior of liquid-like droplets, including fusion, deformation, and wetting on surfaces [78]. The objects were RNA-rich protein condensates called P granules. These structures, now commonly referred to as biomolecular condensates, are primarily protein-based assemblies that create localized intracellular environments without a membrane, and obey local thermal equilibria [79]. By concentrating or excluding specific molecules, condensates can regulate activity, buffer fluctuations, generate mechanical forces, and modulate reaction kinetics [80, 81, 82]. Sixteen years after the first observation of P granules [78], condensates are now recognized as central to cellular function and linked to a wide range of diseases [82, 83, 84].

The physical principles underlying the thermodynamic forces towards phase separation have been known long before their discovery within cells. Still, their biological role and properties in non-equilibrium conditions are being unraveled. One way of maintaining condensates away from equilibrium is through chemical reactions that break detailed balance of the rates, making the droplets active [85, 86, 87, 88]. Active droplets can behave fundamentally different from their equilibrium counterpart, exhibiting properties such as accelerated ripening [89], self propulsion [90], size control [91, 92, 93], arrested ripening [94], surface capillary waves [95], shell-formation [96], and various other spatial morphologies and patterns [97, 98]. Could such emergent behavior also play a role for the emergence of life?

Zwicker et al. [92] provided a compelling example: active droplets that divide, sustaining cycles of growth and division much like living cells. The division of active droplets occurs when the droplet material is produced with the aid of a fuel in the dilute phase and degraded back inside the dense droplet phase. The droplet's shape instability occurs due to a degradation mechanism within the droplet that scales with its volume, while the diffusion of material to the droplet scales with its surface

area. Recently, the theoretical prediction was observed experimentally [99].

Long before their discovery in extant life, membrane-less compartments were proposed by A. Oparin [100] as a protocell for the origins of life. The protocell, formed by the aggregation of organic molecules, could provide a confined environment where essential biochemical reactions occur, potentially leading to the emergence of life. Due to the thermodynamic force towards aggregation, once the molecules exceed the saturation concentration, forming these compartments is simpler than forming membrane-bound compartments, enhancing their prebiotic relevance [101]. The recent discovery of membrane-less organelles' existence and range of functions in extant life might indicate their relevance also for the emergence of life.

1.5 Outline and scope of thesis

The general challenge of creating de novo life is to achieve new collective behavior from the reactions and interactions of simpler constituent components. In physics, this phenomenon is known as emergence [102]. Through emergence, the system must overcome different obstacles in its path towards de novo life. Several key questions remain as to how and which physical mechanisms and non-equilibrium conditions can allow for such emergent behavior. In this thesis, we aim to tackle some of them:

- How can a system speed up specific chemical reactions through its non-equilibrium cyclic environment?
- How can chemistry be regulated to allow for robust responses and stable self-sustaining oscillations?
- How can oligomers be enriched and localized to be sufficiently long, abundant, and with specific sequence patterns?
- How can the sequence information of oligomers be copied with a high enough fidelity for the information to be maintained?

In order to answer these questions, we argue that the interplay between chemistry in non-dilute mixtures away from equilibrium, together with their ability to phase separate and form condensates, allows for complex emergent behavior. The collective phenomena occurring in such a system suggest a promising path to understanding how complex systems can give rise to de novo life.

To this end, we employ concepts from non-equilibrium thermodynamics and statistical mechanics to develop models of dense, interacting mixtures away from equilibrium. These models allow us to investigate how condensed phases alter reaction kinetics, transport, and whether new behavior can emerge. We aim to understand

how physical processes inherent to dense mixtures can facilitate behavior that may be essential for the transition toward de novo life.

Our investigation can be divided into two main sections. In the first section, chapter 3 and 4, we investigate minimalistic ternary systems to study the role of oscillations. In chapter 3, the oscillations externally drive the system, while in chapter 4 they emerge from the system itself. In the second section, chapter 5 and 6, we study the interplay between oligomers and the condensate they form. In chapter 5, we see how condensed phases can direct oligomerization to select specific sequence patterns. The framework is continued in chapter 6 to study the replication of sequence information. A reading guide with a more detailed outline of this thesis is given as follows:

Chapter 2 **Theory of chemical reactions in non-dilute inhomogeneous systems:** Modeling the interplay between chemical reactions and condensates relies on a thermodynamic understanding of both processes. In Chapter 2, the necessary theoretical foundation is introduced. The models developed in the following chapters build directly on these concepts.

Chapter 3 **Accelerating chemical reactions through external reservoirs:** Complex chemical products can be thermodynamically disfavored by large energetic barriers. In chapter 3, we develop a thermodynamic model for chemical reactions in dense mixtures subject to material exchange with a reservoir. We show how specific chemical reaction pathways can be accelerated through repeated cycles of wet and dry conditions. The speed-up of specific chemical reaction pathways emerges due to interactions between the components.

Chapter 4 **Excitability and self-sustaining oscillations of active droplets:** Chemistry can be modulated through both clock-like timers, and by amplifying weak inputs into robust outputs. In chapter 4, we find that chemical reactions in phase-separated systems driven sufficiently far away from equilibrium behave as an excitable medium. The system exhibits both excited trajectories and self-sustaining oscillations. The same reaction network does not oscillate in a homogeneous system, making the excitability a direct consequence of phase separation.

Chapter 5 **Selecting sequence patterns through phase separation:** Sequence space grows exponentially with length, rendering a strong selection mechanism necessary to avoid dilution as sequence space is explored. By developing a thermodynamic description of reversible oligomerization coupled to phase separation, we show in chapter 5 that condensed phases can impose

selection pressures favoring distinct sequence motifs. Different motifs are selected under different non-equilibrium conditions.

Chapter 6 **The emergence of a high fidelity replicator:**
Accurate copying of information is necessary for information to be maintained in a system. In chapter 6, we extend our model of interacting oligomers to show how condensates enhance the copying fidelity of templated ligation, allowing it to survive Eigen's error threshold. Furthermore, we show how selecting different sequence motifs leads to different condensate properties, thereby providing a minimalistic phenotype-genotype relationship.

Chapter 7 **Outlook and conclusion:**
Lastly, we conclude on the findings from the previous chapters, and discuss future avenues of research. The findings of the thesis are then summarized.

Chapter 2

Theory of chemical reactions in non-dilute inhomogeneous systems

In this chapter, we first introduce the diffusive relaxation of non-dilute mixtures towards thermodynamic equilibrium, and their properties at equilibrium, before presenting the thermodynamics of chemical reactions in homogeneous mixtures. Lastly, we combine the two previous sections in order to introduce chemical reactions in phase-separated systems.

2.1 Physics of interacting mixtures

2.1.1 Thermodynamics of mixtures in the Gibbs ensemble

A thermodynamic system externally controlled by temperature T, pressure p, and particle numbers $\{N_i\}$, is described in the Gibbs ensemble. In the Gibbs ensemble, it is the Gibbs free energy $G(\{N_i\}, p, T)$, which is minimized at thermodynamic equilibrium [103]. From derivatives of the Gibbs free energy with respect to the control parameters, we can calculate their conjugate variables:

$$\mu_i = \left(\frac{\partial G}{\partial N_i}\right)_{T,p,N_{j\neq i}}, \qquad V = \left(\frac{\partial G}{\partial p}\right)_{T,N_{j\neq i}}, \qquad S = -\left(\frac{\partial G}{\partial T}\right)_{p,N_{j\neq i}}, \tag{2.1}$$

where μ_i are the chemical potentials, V is the system volume V, and S is the entropy. From the conjugate variables, we find the total differential of the free energy

$$\mathrm{d}G = -S\mathrm{d}T + V\mathrm{d}p + \sum_{i=0}^{M} \mu_i \mathrm{d}N_i \,. \tag{2.2}$$

Throughout our investigations, we will consider incompressible systems, where the molecular volumes $\nu_i = \partial V/\partial N_i\big|_{T,p,N_{j\neq i}}$ are constants under isothermal and isobaric

conditions, making them independent of composition. Within the experimentally relevant ranges of temperature and pressure, the volume of aqueous solutions remains essentially constant [104], which justifies the incompressibility approximation and considerably simplifies the thermodynamic description. From the system's volume $V = \sum_{i=0}^{M} \nu_i N_i$ we can define the Gibbs free energy density

$$g(\{n_i\}, p, T) = G(\{N_i\}, p, T)/V \,, \tag{2.3}$$

where n_i are the concentrations $n_i = N_i/V$.

2.1.2 From the Gibbs to the Helmholtz ensemble

Due to incompressibility, the system volume $V = \sum_{i=0}^{M} N_i \nu_i$ is independent of pressure, temperature, and composition. This allows us to efficiently describe the system through the non-solvent particle numbers (N_i for $i = 1, \ldots, M$) and the system volume V, as the number of solvent particles can be determined by

$$N_0 = \frac{V - \sum_{i=1}^{M} \nu_i N_i}{\nu_0} \,. \tag{2.4}$$

We can perform a Legendre transformation [103] to go from the Gibbs to the Helmholtz ensemble, finding the Helmholtz free energy F, to be

$$F(\{N_i\}, T, V) = G(\{N_i\}, T, p) - pV \,. \tag{2.5}$$

Since the system volume V is now an independent variable, we have reduced the number of variables by one. By treating the system this way, the thermodynamic conjugate variables of particle number and volume are altered compared to Eq. (2.1). We see this by calculating the differential of the Helmholtz free energy $\mathrm{d}F$, using the above expression for the number of solvent particles N_0, finding

$$\mathrm{d}F = -S\mathrm{d}T - \left(p - \frac{\mu_0}{\nu_0}\right) \mathrm{d}V + \sum_{i=1}^{M} \left(\mu_i - \mu_0 \frac{\nu_i}{\nu_0}\right) \mathrm{d}N_i \,. \tag{2.6}$$

The conjugate variable of the particle number becomes the exchange chemical potential $\bar{\mu}_i$ (instead of a chemical potential μ_i in the Gibbs ensemble), defined as

$$\bar{\mu}_i = \mu_i - \frac{\nu_i}{\nu_0} \mu_0 \,. \tag{2.7}$$

The exchange chemical potential captures the energy cost of exchanging a particle with the equivalent volume of solvent particles. Furthermore, the conjugate variable to volume becomes the osmotic pressure Π (instead of pressure p in the Gibbs

ensemble), defined as

$$\Pi = p - \frac{\mu_0}{\nu_0} \,. \tag{2.8}$$

The osmotic pressure measures the amount of pressure needed to prevent the solvent from flowing through the system boundary. We can further express the concentrations n_i in terms of the proportion of volume they occupy, their volume fractions,

$$\phi_i = \nu_i n_i \,. \tag{2.9}$$

The $M+1$ volume fractions are not independent, as the solvent volume fraction ϕ_0 can be defined in terms of the non-solvent volume fractions

$$\phi_0 = 1 - \sum_{i=1}^{M} \phi_i \,. \tag{2.10}$$

Converting back to the Gibbs ensemble can be done by

$$G = F + p \sum_{i=0}^{M} \nu_i N_i. \tag{2.11}$$

We see that for an incompressible mixture, the pressure adds a term proportional to $\phi_i p$ to the free energy density. Importantly, as we will see in the following sections, contributions to the free energy density that are linear in the volume fractions do not affect phase equilibria, making the properties of coexisting phases equal in the Gibbs and Helmholtz ensembles.

2.1.3 Thermodynamic stability and the free energy density

By calculating the double derivatives of the free energy, we get its curvature, which determines whether a homogeneous thermodynamic phase is locally stable or unstable [105]. For a phase to be stable, the curvature with respect to intensive variables (pressure p, temperature T) must be convex, and concave with respect to extensive variables (particle number N_i, volume V). If a homogeneous mixture is unstable, it can separate into distinct regions in space with different compositions and properties. Upon phase separation, the order parameter, which quantifies the composition difference between the phases, discontinuously transitions from a value of zero when homogeneous to a finite value, except when transitioning through the critical point.

A change in the sign of the free energy's curvature can occur due to competing contributions to the free energy, such as between maximizing entropy and minimizing enthalpy. One free energy density f which captures this competition, is the Flory-Huggins free energy density [106, 107]. Derived from coarse-graining a lattice model

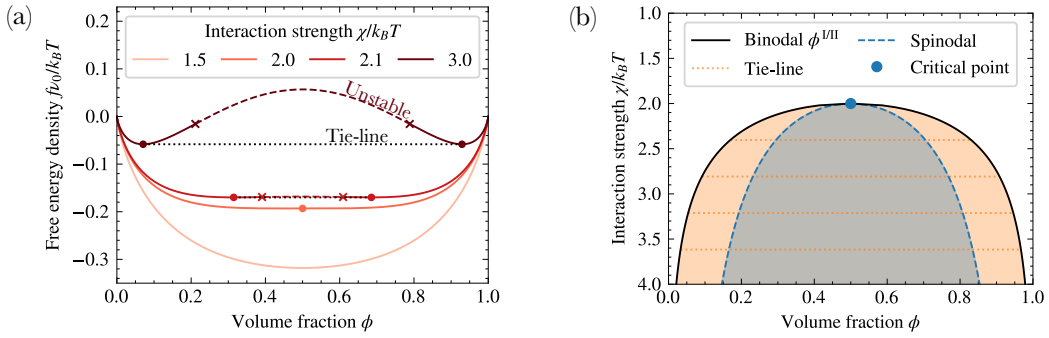


Figure 2.1: **Binary free energy density and phase diagram:** (a) The Flory-Huggins free energy density in Eq. (2.12) for a binary mixture below and above the critical interaction strength of $\chi = 2k_BT$. Above the critical interaction strength, the free energy density curvature changes from positive (stable, solid) to negative (unstable, dashed) at the spinodal (cross). When phase separated, the conditions of phase coexistence (Eq. (3.11)) are satisfied at the binodal points $\phi^{\mathrm{I/II}}$ (dots). (b) As the interaction strength χ is increased, the compositional difference between the dense and dilute phase increases, together with the spinodal domain. The curves are generated for the symmetric case $\nu_0 = \nu_1$.

of interacting polymers, as derived in Appendix A, the Flory-Huggins free energy density reads

$$f = \frac{k_BT}{\nu_0}\left[\sum_{i=0}^{M}\left(\frac{\phi_i}{\rho_i}\log\left(\frac{\phi_i}{\rho_i}\right) + \frac{\omega_i\phi_i}{k_BT}\right) + \frac{1}{2k_BT}\sum_{i=0}^{M}\sum_{j=0}^{M}\chi_{ij}\phi_i\phi_j\right], \tag{2.12}$$

where we have introduced ρ_i as the relative molecular volume of component i with respect to the solvent $\rho_i = \nu_i/\nu_0$. The logarithms quantify the mixing entropy, while the first-order virial coefficients ω_i are found from expanding the free energy to first order. The ω_i are often referred to as internal energies, but can also contain entropic contributions. The quadratic coefficients are referred to as Flory-Huggins parameters, defined as

$$\chi_{ij} = \frac{z}{2}\left(e_{ij} - \frac{e_{ii} + e_{jj}}{2}\right), \tag{2.13}$$

where e_{ij} are the interaction energy between component i and j, and z is the number of nearest neighbors on the lattice. The Flory-Huggins parameters quantify how well two components interact with each other, compared to themselves. In this description, molecular details are not explicitly accounted for, but reduced to their virial coefficients and Flory-Huggins parameters. In Eq. (2.12), $M+1$ is the number of components, where the solvent is typically index zero. A factor of thermal energy per solvent molecular volume is factored out, giving us our reference volume and energy scales. Rather than accounting for the microscopic spatial fluctuations and correlations between individual molecules, the Flory-Huggins free energy treats interactions between molecules on the level of average concentrations over the volume, making it a mean-field description.

In Fig. 2.1(a), the Flory-Huggins free energy density f (Eq. (2.12)) for a binary system with equal molecular volumes is shown above and below the critical interaction strength. For $\chi < 2k_BT$, entropy dominates over enthalpy, making the free energy density convex at all compositions. The system is therefore homogeneous. Above the critical interaction strength of $\chi > 2k_BT$, two phases of different composition can coexist. The phases are connected by a tie-line. In between the two compositions, there exists a range of compositional values that are unstable, characterized by the concave free energy density $f''(\phi) < 0$. The conditions of phase coexistence, which will be introduced in section 2.1.5, corresponds to the free energy density having equal slope ($f'|_{\phi^{\mathrm{I}}} = f'|_{\phi^{\mathrm{II}}}$) and value ($f(\phi^{\mathrm{I}}) = f(\phi^{\mathrm{II}})$) in the two phases, which we observe in the figure.

The free energy density in Eq. (2.12) does not account for energy costs associated with gradients in the volume fractions ϕ_i. The energy cost of compositional gradients gives rise to surface tension, making it an important contribution to the free energy of liquids. This energy cost is incorporated through the interfacial energy terms κ_{ij}, making the total free energy in three dimensions

$$F(\{\phi_i\}, \{\nabla\phi_i\}) = \int_V \left(f(\{\phi_i\}, T) + \frac{1}{2} \sum_{i=0}^{M} \sum_{j=0}^{M} \kappa_{ij} \left(\nabla\phi_i\right) \cdot \left(\nabla\phi_j\right) \right) \mathrm{d}^3\boldsymbol{x} \,. \tag{2.14}$$

We will only consider diagonal interfacial-energy matrices $\kappa_{ij} = \kappa_i\delta_{ij}$, which do not result in qualitatively different behavior. The exchange chemical potentials $\bar{\mu}_i$ can be calculated from the functional derivative of the free energy

$$\bar{\mu}_i = \nu_i \left(\frac{\delta F}{\delta \phi_i} \right)_{T,V,\phi_{j\neq i}} \,. \tag{2.15}$$

Independently of the chosen free energy density, the chemical potentials μ_i, which relate to the exchange chemical potentials through Eq. (2.7), can be decomposed into a composition-independent term, the reference chemical potential μ_i^0, and a composition-dependent term γ_i known as the activity coefficient, defined such that

$$\mu_i = \mu_i^0(p, T) + k_BT \log\left(\phi_i\gamma_i\right) + \nu_i\kappa_i \left(\nabla\phi_i\right)^2 \,. \tag{2.16}$$

2.1.4 Relaxation to thermodynamic equilibrium

In response to thermodynamic forces, the volume fractions vary in space $\boldsymbol{x}$ and time t through conserved fluxes $\boldsymbol{j}_i$, according to [108]

$$\frac{\partial\phi_i(\boldsymbol{x}, t)}{\partial t} = -\nabla \cdot \boldsymbol{j}_i \,. \tag{2.17}$$

The fluxes are in the direction of increasing the total entropy of the system and the heat bath, that is, decreasing the Helmholtz free energy F. We can determine the magnitude of the fluxes by calculating the free energy change

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \int_V \mathrm{d}^3\boldsymbol{x} \sum_{i=1}^{M} \frac{\partial \phi_i}{\partial t} \left(\frac{\partial f}{\partial \phi_i} - \frac{\partial f}{\partial \phi_0} \right) , \tag{2.18}$$

where we have assumed that the system volume is constant. Note here that we have $M+1$ components, but they are not independent due to Eq. (2.10), such that $\partial_t \phi_0 = -\sum_{i=1}^{M} \partial_t \phi_i$. Employing that the volume fractions are conserved (Eq. (2.17)), the definition of the chemical potentials (Eq. (2.15)), and applying integration by parts for a closed system, we find

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \int_V \mathrm{d}^3\boldsymbol{x} \sum_{i=1}^{M} \frac{1}{\nu_i} \nabla(\mu_i - \frac{\nu_i}{\nu_0}\mu_0) \cdot \boldsymbol{j}_i \, , \tag{2.19}$$

where we have assumed that the molecular volumes are constant in space. To ensure that the free energy change is negative or zero, the diffusive flux $\boldsymbol{j}_i$ must be proportional to the gradients in exchange chemical potentials $\bar{\mu}_i$ from Eq. (2.7), such that

$$\boldsymbol{j}_i = -\sum_{k=1}^{M} \Lambda_{ik} \nabla \bar{\mu}_k \, , \tag{2.20}$$

where Λ_{ik} are the mobility coefficients. The mobility coefficients Λ_{ik} quantify how strongly the volume fractions respond to thermodynamic forces. For the fluxes to reduce the free energy, the eigenvalues of the mobility matrix Λ_{ik} must be positive.

Throughout this thesis, we will choose a mobility matrix that retrieves Fick's law in the dilute limit (2.20). By dilute limit, we refer to cases where the activity coefficients γ_i from Eq. (2.16) are constant. This occurs when all non-solvent components i are dilute ($\phi_i \to 0$ for $i = 1, \dots M$), which is equivalent to solvent rich systems ($\phi_0 \to 1$). In this limit, the dynamic Eq. (2.17) becomes

$$\frac{\partial \phi_i}{\partial t} = k_B T \nabla \cdot \left(\sum_{k=1}^{M} \Lambda_{ik} \left(\frac{1}{\phi_k} \nabla \phi_k - \frac{\rho_k}{\phi_0} \nabla \phi_0 + \frac{\kappa_k \nu_k}{k_B T} \nabla \left(\nabla \phi_k \right)^2 - \frac{\kappa_0 \nu_k}{k_B T} \nabla \left(\nabla \phi_0 \right)^2 \right) \right) , \tag{2.21}$$

where we have used the relative molecular volumes $\rho_i = \nu_i / \nu_0$. The diffusion equation is retrieved in the dilute limit with a mobility matrix

$$\Lambda_{ik} = \frac{D_i \delta_{ik}}{k_B T} \phi_i \phi_0 \, , \tag{2.22}$$

which is a symmetric mobility matrix that obeys Onsager reciprocal relation [109].

A mobility on this form represents the case where a molecule of species i must be exchanged with the solvent to diffuse. With this mobility matrix, we find

$$\frac{\partial \phi_i}{\partial t} \simeq D_i \phi_0 \nabla^2 \phi_i - D_i \rho_i \phi_i \nabla^2 \phi_0 + D_i \left(1 - \rho_i\right) \nabla \phi_0 \cdot \nabla \phi_i \,, \tag{2.23}$$

where we have ignored higher order gradients $\mathcal{O}(\nabla^4)$. By employing the constraint that all ϕ_i must sum to 1 (Eq. (2.10)), we see that the second and third terms are of the order $\mathcal{O}(\phi_i \nabla^2 \phi_i)$ and $\mathcal{O}((\nabla \phi_i)^2)$ respectively. Thus, by ignoring higher orders in ϕ_i, we find

$$\frac{\partial \phi_i}{\partial t} \simeq D_i \nabla^2 \phi_i \,. \tag{2.24}$$

In the dilute limit, the system is diffusing according to Fick's law.

An alternative approach to relate the diffusive flux in Eq. (2.20) to Fick's law is through the chain rule:

$$\nabla \bar{\mu}_k = \sum_{j=1}^{M} \frac{\partial \bar{\mu}_k}{\partial \phi_j} \nabla \phi_j. \tag{2.25}$$

From this form, we can define the diffusion coefficient D_{ij} such that the diffusive fluxes are written as

$$\boldsymbol{j}_i = -\sum_{j=1}^{M} D_{ij} \nabla \phi_j \,, \qquad\qquad D_{ij} = \sum_{k=1}^{M} H_{jk} \Lambda_{ik} \,, \tag{2.26}$$

where we have defined $H_{jk} = \partial \bar{\mu}_k / (\partial \phi_j)$. In Eq. (2.22), we chose a diagonal mobility matrix Λ_{ik}, a choice which can in general not be made for the diffusion matrix D_{ij}. Choosing a diagonal diffusion matrix $D_{ij} = D_i \delta_{ij}$ can break with microscopic reversibility, as it implies a mobility matrix $\Lambda_{ik} = \sum_{j=1}^{M} D_{ij} H_{kj}^{-1}$, which in general does not obey the Onsager reciprocal relation of $\Lambda_{ik} = \Lambda_{ki}$ [109]. A diagonal diffusion coefficient is valid in the dilute limit (constant γ_i's), as in this limit H_{jk} is a diagonal matrix. Note that since the conjugate flux of the exchange chemical potential is not the volume fractions, but the concentrations n_i, Λ_{ik} is not Onsager's response coefficients. Instead, it is the matrix $\tilde{\Lambda}_{ik} = \Lambda_{ik}/\nu_k$ which must obey the reciprocal relation $\tilde{\Lambda}_{ik} = \tilde{\Lambda}_{ki}$.

While the mobility matrix Λ_{ik} must have positive Eigenvalues for the free energy to reduce with time (Eq. (2.19)), the diffusion matrix D_{ij} does not have to obey the same restrictions. This is due to the matrix $H_{jk} = \partial \bar{\mu}_k / \partial \phi_j$, which can have Eigenvalues of either sign. In fact this matrix H_{jk} is proportional to the double derivative of the free energy density $H_{jk} = \nu_k \tilde{H}_{jk}$, where $\tilde{H}_{jk} = \partial^2 f / (\partial \phi_j \partial \phi_k)$ is the Hessian of the free energy density. If an Eigenvalue of the Hessian is negative, the mixture is unstable [110]. Due to Λ_{ik} being a positive definite symmetric matrix, a negative eigenvalue of the Hessian creates a negative Eigenvalue of D_{ij} [111].

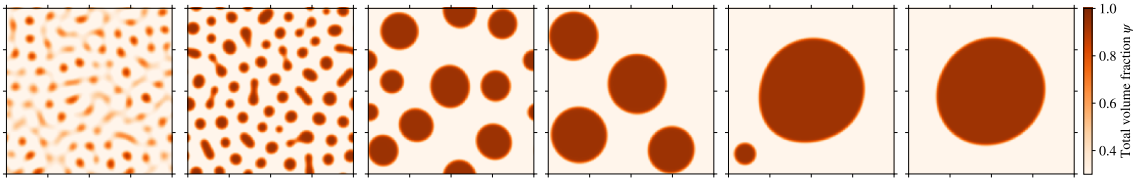


Figure 2.2: **An emulsion's relaxation to thermodynamic equilibrium:** After spinodal decomposition, droplets form, coalesce, and ripen, with the characteristic length scale growing as $t^{1/3}$ until a single dense phase coexists with the dilute environment at equilibrium.

A negative Eigenvalue of D_{ij} creates an effectively negative diffusion coefficient, resulting in any small perturbation growing in time. The mixture's homogeneous composition is unstable.

Compositions leading to one or more negative eigenvalues of the Hessian are in a region of the phase diagram called the spinodal, where any fluctuation will grow with time. In Fig. 2.1(b), we see how this unstable region (spinodal) increases with the interaction strength χ. In figure 2.2, the evolution of an emulsion of droplets right after spinodal decomposition, and towards thermodynamic equilibrium, is shown. Spinodal decomposition leads to an emulsion of droplets of different sizes that coalesce and ripen over time to a single droplet. The ripening process increases the characteristic length scale as $t^{1/3}$ [112] before a single dense phase coexists with a dilute environment as thermodynamic equilibrium is reached [87]. At thermodynamic equilibrium, two phases of different composition coexist with one another.

2.1.5 Conditions of phase coexistence

We start by discussing a system consisting of two homogeneous coexisting phases I and II, where phase I is solvent-rich and phase II is solvent-poor. The total free energy F of such a system can be decomposed into the bulk energy of each phase and the energy associated with the interface between the two phases

$$F = V^{\mathrm{I}} f^{\mathrm{I}} + V^{\mathrm{II}} f^{\mathrm{II}} + \sigma A \,. \tag{2.27}$$

Here, V^α denotes the phase volume with associated free energy density f^α, where α is the phase index $\alpha \in \{\mathrm{I}, \mathrm{II}\}$. The term σA represents the free energy of the interface with area A and associated surface tension σ. The surface tension can be related to κ_i in Eq. (2.14), for a binary mixture of equal molecular volumes as $\sigma \simeq \frac{(\chi - 2k_BT)^{3/2}}{2k_BT}\sqrt{\frac{\kappa}{\nu}}$ [87].

The minimization of the free energy F defines thermodynamic equilibrium. However, the minimization must obey certain constraints. The volume of each phase must combine to the system volume $V = V^{\mathrm{I}} + V^{\mathrm{II}}$, and the number of particles for

each component i is conserved throughout the system, meaning the phase average volume fractions are conserved upon particle exchange between the phases,

$$\bar{\phi}_i = \frac{V^{\mathrm{I}}\phi_i^{\mathrm{I}} + V^{\mathrm{II}}\phi_i^{\mathrm{II}}}{V}\,. \tag{2.28}$$

By assuming that the dense phase, phase-II, is spherical with radius $R = (3V^{\mathrm{II}}/4\pi)^{1/3}$, the free energy F is minimized under volume exchange and composition exchange between the phases when the exchange chemical potentials $\bar{\mu}_i$ and the osmotic pressure Π, including Laplace pressure, is equal between the two phases

$$\bar{\mu}_i^{\mathrm{I}} = \bar{\mu}_i^{\mathrm{II}}\,, \qquad \text{and} \qquad \Pi^{\mathrm{I}} = \Pi^{\mathrm{II}} + \frac{2\sigma}{R}\,. \tag{2.29}$$

The exchange chemical potential and osmotic pressure are defined in Eq. (2.7) and Eq. (2.8), respectively. The compositions ϕ_i^{I} and ϕ_i^{II}, which satisfy the conditions of phase coexistence, create what is known as the binodal. In Fig. 2.1a, the binodal points are depicted together with the free energy density, where we see that the coexisting compositions have equal derivatives. The binodal separates the region in phase space where it is thermodynamically favored to be homogeneous or phase-separated. It corresponds to the yellow shaded domain in Fig. 2.1b, and the blue shaded domain in Fig. 2.3. The binodal does not overlap with the spinodal, meaning a homogeneous thermodynamic state can be locally stable, even though the thermodynamic equilibrium is phase separated.

In the Gibbs ensemble, the equivalent conditions of phase coexistence are $M+1$ chemical potentials being equal in both phases, $\mu_i^{\mathrm{I}} = \mu_i^{\mathrm{II}}\ \forall i$, which we see from the definition of the exchange chemical potentials (2.7) and osmotic pressure (2.8) are equivalent conditions. Thermodynamic equilibrium is therefore the same in both ensembles. Note that $\mu_i^{\mathrm{I}} = \mu_i^{\mathrm{II}}$ is also satisfied at phase coexistence in the Helmholtz ensemble.

At thermodynamic equilibrium, one can find the general number of degrees of freedom $\mathcal{F}$ by counting the number of constraints acting on the $M+1$ components in the $\mathcal{P}$ phases. Subtracting the number of constraints from phase coexistence ($M+1$ constraints between $P-1$ phases) from the number of unknowns (M concentrations in P phases), one finds[1]

$$\mathcal{F} = 1 + M - \mathcal{S} - \mathcal{P} + \mathcal{T}\,. \tag{2.30}$$

Here $\mathcal{S}$ gives the number of additional constraints not yet accounted for, originating from, for example, chemical equilibrium, and $\mathcal{T}$ is the number of thermodynamic variables like pressure and temperature.

[1]Note the $+1$ instead of the more common $+0$, which originates from M counting the non-solvent components.

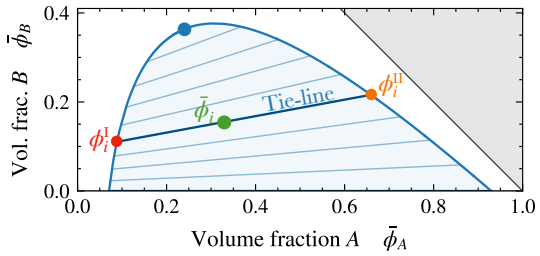


Figure 2.3: **Ternary phase diagram:** The binodal for a ternary system (A, B, C), where A separates from C ($\chi_{AC} > 2k_BT$). The phase averaged volume fraction $\bar{\phi}_i$ separates into a dilute phase I and a dense phase II. The dense and dilute phases obey the conditions of phase coexistence (Eq. (2.29)). Their composition and volumes are uniquely set for a given $\bar{\phi}_i$. The critical point is the blue circle at the top of the binodal.

The number of degrees of freedom must be greater than or equal to zero, setting an upper bound on the number of possible phases $\mathcal{P}$,

$$\mathcal{P} \leq M - \mathcal{S} + \mathcal{T} + 1\,, \tag{2.31}$$

known as Gibbs' phase rule [105]. For a fixed $\bar{\phi}_i$ for all $M+1$ components, we want to determine both the compositions in all phases and the volumes of all phases. Compared to the counting used to derive Gibbs' phase rule, the number of unknowns increases by $\mathcal{P}-1$, as we need to determine the phase volumes. Due to $\bar{\phi}_i$ being known, $\mathcal{M}$ constraints are included in $\mathcal{S}$. For the case of a isothermal and isobaric system ($\mathcal{T}=0$), Eq. (2.30) yields zero degrees of freedom ($\mathcal{F}=0$), meaning all unknowns can be uniquely determined.

For the case of two-phase coexistence ($\mathcal{P}=2$) with a specific temperature and pressure ($\mathcal{T}=0$), the $2M+2$ unknowns ($\phi_i^{\mathrm{I/II}}$ and $V^{\mathrm{I/II}}$) can be uniquely determined for M given phase averaged volume fractions (Eq. (2.28)). In figure 2.3, the unique selection of a dense and dilute phase composition is depicted in a ternary phase diagram. The two phases satisfy the conditions of phase coexistence (Eq. (2.29)), and are connected by a tie-line.

2.2 Thermodynamics of chemical reactions in homogeneous mixtures

As phase separation lowers the free energy through localization of particles in space, chemical reactions can lower the free energy by converting particles into one another. Each particle type is no longer conserved, as in Eq. (2.17). Instead, the reaction network leads to new composite concentrations that are conserved.

2.2.1 Chemical reaction networks

The conversion of components C_i through a chemical reaction α is written as a sum over the $M+1$ components,

$$\sum_{i=0}^{M} C_i \sigma_{i\alpha}^{\rightharpoonup} \rightleftharpoons \sum_{i=0}^{M} C_i \sigma_{i\alpha}^{\leftharpoondown} , \tag{2.32}$$

where the stoichiometric coefficients $\sigma_{i\alpha}^{\rightleftharpoons}$ gives the number of particles for each component i that are converted under the reaction α [113]. If the reaction α occurs from left to right in Eq. (2.32), $\sigma_{i\alpha}^{\rightharpoonup}$ denote the number of particles converted to $\sigma_{i\alpha}^{\leftharpoondown}$. The difference between the two defines the stoichiometric matrix

$$\sigma_{i\alpha} = \sigma_{i\alpha}^{\leftharpoondown} - \sigma_{i\alpha}^{\rightharpoonup} . \tag{2.33}$$

In aqueous mixtures, the volume change due to chemical reactions is small [114]. Throughout the thesis, we will assume that volume is exactly conserved during chemical reactions, restricting the stoichiometric matrix and molecular volumes to satisfy

$$\sum_{i=0}^{M} \sigma_{i\alpha} \nu_i = 0 , \tag{2.34}$$

for each reaction $\alpha = 1, \dots R$. This approximation allows us to reduce the number of independent variables by one, as pressure becomes a constant throughout the chemical kinetics.

Chemical reactions allow volume fractions, which would otherwise be conserved, to change in time, creating conserved composite volume fractions. In a system of $M+1$ components, the number of conserved variables is reduced by the number of linearly independent reactions R, granting $C = M - R$ conserved volume fractions, not counting the conserved volume. The basis vectors q_j spanning the nullspace of $\nu_i \sigma_{i\alpha}$ make up the rows of the matrix Q, such that $\sum_{i=0}^{M} Q_{ji} \nu_i \sigma_{i\alpha} = 0$. From the matrix Q_{ji}, of size $M \times (M-R)$, the conserved quantities ψ_j are found from

$$\psi_j = \sum_{i=0}^{M} Q_{ji} \phi_i . \tag{2.35}$$

Note that the set of conserved volume fractions ψ_j is not unique, as any linear combination of them is also conserved. Accordingly, there exists $R = M - C$ non-conserved quantities, called reaction extents

$$\xi_\alpha = \sum_{i=0}^{M} P_{\alpha i} \phi_i . \tag{2.36}$$

The reaction extents quantify the total number of reaction events which have occurred for each reaction α, where $P_{\alpha i}$ is calculated from the pseudo-inverse of $\nu_i \sigma_{i\alpha}$ [115]. For a thorough discussion, see e.g. Ref. [116]. In a homogeneous mixture, all ψ_j are constant in time, while the ξ_α are the reduced set of dynamical variables. For an inhomogeneous system, it is the system average $\bar{\psi}_j$, defined as

$$\bar{\psi}_j = \int_V \mathrm{d}^3\boldsymbol{x}\, \psi_j(\boldsymbol{x})\,, \tag{2.37}$$

which are conserved.

Let us look at a simple example of one uni-molecular reaction $r_1 : A \rightleftharpoons B$ and one bi-molecular $r_2 : A + B \rightleftharpoons C$. The corresponding stoichiometric coefficients are $\sigma_{1,i} = \delta_{i,A} - \delta_{i,B}$ and $\sigma_{2,i} = \delta_{i,A} + \delta_{i,B} - \delta_{i,C}$. For volume to be conserved, the molecular volumes must satisfy $\nu_A = \nu_B$ and $\nu_C = 2\nu_{A/B}$. The reaction extents for the two reactions become $\xi_1 = \phi_B - \phi_A$ and $\xi_2 = \phi_C - \phi_A - \phi_B$, with conserved quantities $\psi_1 = \phi_A + \phi_B$ and $\psi_2 = \phi_A + \phi_B + \phi_C$. As both reactions are occurring concomitantly, the conserved variable for both reactions is $\phi_A + \phi_B + 2\phi_C$.

2.2.2 Chemical reactions towards equilibrium

The volume fractions of a chemical reacting component i, will evolve according to

$$\frac{\partial \phi_i}{\partial t} = r_i\,, \qquad r_i = \sum_{\alpha=1}^{R} \nu_i \sigma_{i\alpha} r_\alpha\,, \tag{2.38}$$

where r_i is the net source/sink term of component i and $\alpha = 1, \ldots R$ are all reactions occuring in the system. Chemical reactions towards chemical equilibrium are in the direction of lower free energy F. We can derive an expression for r_α which ensures that the total free energy change,

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \int_V \mathrm{d}^3\boldsymbol{x} \sum_{i=0}^{M} \frac{\mu_i}{\nu_i} \frac{\partial \phi_i}{\partial t}\,, \tag{2.39}$$

is negative. Using the expression for the time derivative of ϕ_i in Eq. (2.38), we find

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \int_V \mathrm{d}^3\boldsymbol{x} \sum_{\alpha=1}^{R} r_\alpha \left(\mu_\alpha^{\rightharpoonup} - \mu_\alpha^{\leftharpoondown}\right)\,, \tag{2.40}$$

where we have defined the forward and backward chemical reaction free energy $\mu_\alpha^{\rightleftharpoons}$ as

$$\mu_\alpha^{\rightleftharpoons} = \sum_{i=0}^{M} \sigma_{i\alpha}^{\rightleftharpoons} \mu_i. \tag{2.41}$$

The difference between the forward and backward chemical reaction free energy $\Delta\mu_\alpha$ is the thermodynamic force of the reaction, which quantifies the free energy released upon a reaction event $\Delta\mu_\alpha = \mu_\alpha^{\rightharpoonup} - \mu_\alpha^{\leftharpoondown}$. If $\Delta\mu_\alpha = 0$, there is no dissipation, and chemical equilibrium is reached. We see from Eq. (2.40) that if the reaction rate r_α is negatively proportional to the difference between the forward and backward chemical reaction free energy, the free energy change is always negative or zero. One choice of reaction rates ensuring a reduction in free energy is the linear choice:

$$r_\alpha = \tilde{k}_\alpha \left(\mu_\alpha^{\leftharpoondown} - \mu_\alpha^{\rightharpoonup}\right) . \tag{2.42}$$

An alternative choice, which we will use throughout the thesis, is a non-linear reaction flux on the form

$$r_\alpha = k_\alpha \left(\exp\left\{ \frac{\mu_\alpha^{\leftharpoondown}}{k_B T} \right\} - \exp\left\{ \frac{\mu_\alpha^{\rightharpoonup}}{k_B T} \right\} \right) . \tag{2.43}$$

There are a few reasons for choosing the non-linear reaction flux. One is that the exponential scaling can be derived from Liouville's theorem and microscopic reversibility [117]. Furthermore, with a Flory-Huggins free energy density (2.12), the law of mass action is retrieved without any Taylor expansion, as we will see in section 2.2.3. Lastly, the conditions of detailed balance of the rates between the forward $r_\alpha^{\rightharpoonup}$ and backward $r_\alpha^{\leftharpoondown}$ rates [118],

$$\frac{r_\alpha^{\rightharpoonup}}{r_\alpha^{\leftharpoondown}} = \exp\left\{ \frac{\Delta\mu_\alpha}{k_B T} \right\} , \tag{2.44}$$

is fulfilled. Detailed balance of the rates only constrains the ratio of the forward and backward fluxes, but does not determine their individual magnitude. There are many choices of the forward and backward rates that satisfy this ratio, where we have chosen the symmetric case, where

$$r_\alpha^{\rightleftharpoons} = k_\alpha \exp\left\{ \frac{\mu_\alpha^{\rightleftharpoons}}{k_B T} \right\} , \tag{2.45}$$

and k_α are the positive reaction rate-coefficients associated with each reaction α.

Note that for the detailed balance of the rates to be fulfilled, the rate coefficient must be equal for the forward and backward pathways, but can depend on composition as long as they remain positive.

A common reaction example is the case of a reaction between a set of reactants A and products B with a transition state T in between, such that $A \rightleftharpoons T \rightleftharpoons B$. As the transition state acts as a barrier ($\mu_T - \mu_A \gg k_B T$, $\mu_T - \mu_B \gg k_B T$), it relaxes fast, and thermodynamic properties can be absorbed in the rate coefficient of the

net transition rate from A to B. The dynamics can therefore be represented by a single effective reaction between A and B, with rates determined by their chemical potentials and an effective coefficient $k = \frac{1}{2} r_T e^{-\mu_T/(k_B T)}$, where r_T denotes the decay rate of the transition state. Thermodynamics does not specify rate coefficients. Kinetic rate coefficients must instead be determined from experiments or microscopic theories, meaning we can treat the full rate coefficient k as an unknown kinetic coefficient [86].

2.2.3 Retrieving the law of mass action

As we saw in Eq. (2.16), the chemical potentials μ_i can be decomposed into a composition-independent term, μ_i^0, and a composition-dependent term γ_i known as the activity coefficient, such that $\mu_i = \mu_i^0 + k_B T \log(\phi_i \gamma_i)$ in the absence of gradients. The argument of the logarithm, $\gamma_i \phi_i$, is often referred to as the chemical activity. With this decomposition of the chemical potential, we can express the ratio of volume fractions at chemical equilibrium ($\Delta\mu_\alpha = 0$) as

$$\prod_{i=0}^{M} \phi^{\sigma_{i\alpha}} = \prod_{i=0}^{M} \gamma_i^{-\sigma_{i\alpha}} \exp\left\{ \frac{-\mu_i^0 \sigma_{i\alpha}}{k_B T} \right\}. \tag{2.46}$$

As the right-hand side determines the chemical equilibrium volume fractions, it becomes useful to define the equilibrium coefficient for each reaction α

$$K_\alpha = \prod_{i=0}^{M} \phi^{\sigma_{i\alpha}}. \tag{2.47}$$

The equilibrium coefficients can be decomposed as $K_\alpha = K_\alpha^{\rightharpoonup} / K_\alpha^{\leftharpoondown}$, where

$$K_\alpha^{\rightleftharpoons} = \prod_{i=0}^{M} \gamma_i^{\sigma_{i\alpha}^{\rightleftharpoons}} \exp\left\{ \frac{\mu_i^0 \sigma_{i\alpha}^{\rightleftharpoons}}{k_B T} \right\}. \tag{2.48}$$

The rate equation in Eq. (2.38) can now be expressed in terms of the forward and backward rate coefficients, retrieving the law of mass action

$$\frac{\mathrm{d}\phi_i}{\mathrm{d}t} = \sum_{\alpha}^{R} \nu_i \sigma_{i\alpha} k_\alpha \left[K_\alpha^{\rightharpoonup} \prod_{i=0}^{M} \phi_i^{\sigma_{i\alpha}^{\rightharpoonup}} - K_\alpha^{\leftharpoondown} \prod_{i=0}^{M} \phi_i^{\sigma_{i\alpha}^{\leftharpoondown}} \right]. \tag{2.49}$$

In the dilute limit, or equivalently, the limit of no interactions, the activity coefficients γ_i become constant, resulting in the equilibrium coefficients $K_\alpha^{\rightleftharpoons}$ being constants. Thus, the dilute form of the law of mass action is found.

The dilute form of the law of mass actions entwines the thermodynamic equilib-

rium coefficients $K_\alpha^{\rightleftharpoons}$ with the kinetic rate coefficients k_α, such that

$$\frac{\mathrm{d}\phi_i}{\mathrm{d}t} = \sum_\alpha^R \nu_i \sigma_{i\alpha} \left[k_\alpha^{\rightharpoonup} \prod_{i=0}^M \phi_i^{\sigma_{i\alpha}^{\rightharpoonup}} - k_\alpha^{\leftharpoondown} \prod_{i=0}^M \phi_i^{\sigma_{i\alpha}^{\leftharpoondown}} \right] , \tag{2.50}$$

where we have defined $k_\alpha^{\rightleftharpoons} = k_\alpha K_\alpha^{\rightleftharpoons}$. Through a free-energy framework, where the kinetic rate coefficient k_α is independent of the thermodynamic forces μ_i, it is clear whether a steady state is at thermodynamic equilibrium or not. In section 2.3 we will see that distinguishing between kinetic and thermodynamic origins for reaction rates becomes important for inhomogeneous systems.

2.2.4 Breaking detailed balance of the rates

Detailed balance of the rates is broken by the forward and backward rate coefficient not satisfying Eq. (2.44). One way of achieving this is by including an additional fueled reaction β to the one in Eq. (2.32), which is identical to the reaction α, but includes the turnover of a fuel F to a waste W,

$$\sum_{i=0}^M C_i \sigma_{i\alpha}^{\rightharpoonup} + F \rightleftharpoons \sum_{i=0}^M C_i \sigma_{i\alpha}^{\leftharpoondown} + W \, , \tag{2.51}$$

such that $\sigma_{i\beta} = \sigma_{i\alpha} + \delta_{i,F} - \delta_{i,W}$. The chemical potential of the fuel μ_F and the waste μ_W can be chemostated from the outside, maintaining the system away from chemical equilibrium. Detailed balance of the rates for the reactions α and β becomes

$$\frac{r_\alpha^{\rightharpoonup} + r_\beta^{\rightharpoonup}}{r_\alpha^{\leftharpoondown} + r_\beta^{\leftharpoondown}} = \exp\left\{ \frac{\Delta\mu_\alpha}{k_B T} \right\} \left(\frac{1 + \frac{k_\beta}{k_\alpha} \exp\left\{ \frac{\mu_F}{k_B T} \right\}}{1 + \frac{k_\beta}{k_\alpha} \exp\left\{ \frac{\mu_W}{k_B T} \right\}} \right) . \tag{2.52}$$

The degree to which the system is maintained away from chemical equilibrium is thus captured in the parentheses on the right-hand side, quantified by the parameter

$$\Delta\mu_{\alpha,\beta} = k_B T \log \left(\frac{1 + \frac{k_\beta}{k_\alpha} \exp\left\{ \frac{\mu_F}{k_B T} \right\}}{1 + \frac{k_\beta}{k_\alpha} \exp\left\{ \frac{\mu_W}{k_B T} \right\}} \right) . \tag{2.53}$$

For the case of either $k_\beta = 0$ or $\mu_F = \mu_W$, the detailed-balance obeying rate is retrieved, as $\Delta\mu_{\alpha,\beta} = 0$ [115, 119].

2.3 Dynamics of chemically active mixtures

In this section, we look at the combined dynamics of interacting mixtures leading to phase separation and chemical reactions. We will see that the two processes are intrinsically linked by the same thermodynamic forces.

2.3.1 Reaction-diffusion in dense systems

In a system where components can diffuse (Eq. (2.17)) and chemically react (Eq. (2.38)), the evolution of volume fractions $\phi_i(\boldsymbol{x}, t)$ in space $\boldsymbol{x}$ and time t follow [88, 119, 87]:

$$\partial_t \phi_i(\boldsymbol{x}, t) = -\nabla \cdot \boldsymbol{j}_i + r_i \,. \tag{2.54}$$

With the rates of diffusion and reactions from Eq. (2.20) and Eq. (2.38), the volume fractions follow

$$\frac{\partial \phi_i}{\partial t} = \nabla \left(\cdot \Lambda_i \nabla \bar{\mu}_i \right) + \sum_{\alpha=1}^{R} k_\alpha \nu_i \sigma_{i\alpha} \left(\exp\left\{ \frac{\mu_\alpha^{\leftarrow}}{k_B T} \right\} - \exp\left\{ \frac{\mu_\alpha^{\rightarrow}}{k_B T} \right\} \right) . \tag{2.55}$$

Importantly, the chemical potentials determine both the diffusive fluxes and the reaction fluxes, making the two processes intrinsically linked. The combination of diffusion and reaction introduces reaction–diffusion lengths

$$\ell_{i\alpha} = \sqrt{\frac{\Lambda_i}{k_B T \, k_\alpha}} \,, \tag{2.56}$$

which quantifies how far component i can diffuse during the time scale of the reaction α. Depending on how the characteristic length scale of the system L compares to the reaction-diffusion length scale, either diffusion or reactions can be the dominant timescale of the dynamics. Diffusive equilibrium is reached when the chemical potentials of all components are equal in space ($\nabla \mu_i = 0$), and chemical equilibrium is reached when the chemical potentials on each side of all reactions are equal ($\Delta \mu_\alpha = 0$).

For sufficiently strong interactions between the components, the system can phase-separate, creating stable inhomogeneous volume-fraction profiles at thermodynamic equilibrium, as discussed in section 2.1. The conditions for phase coexistence between two phases I and II is set by the chemical potentials being equal $\mu_i^{\mathrm{I}} = \mu_i^{\mathrm{II}}$. From the decomposition of the chemical potentials in Eq. (2.29), we can express the partitioning coefficient $P_i = \phi_i^{\mathrm{II}} / \phi_i^{\mathrm{I}}$, as

$$P_i = \frac{\gamma_i^{\mathrm{I}}}{\gamma_i^{\mathrm{II}}} \,. \tag{2.57}$$

The ratio of the activity coefficients can equivalently be expressed as the ratio of the chemical equilibrium coefficient K_α from Eq. (2.47), finding

$$\frac{K_\alpha^{\mathrm{I}}}{K_\alpha^{\mathrm{II}}} = \prod_{i=0}^{M} P_i^{\sigma_{i\alpha}} \,. \tag{2.58}$$

For equilibrium systems, what determines phase coexistence also determines chemical equilibrium, their equilibria are coupled [113]. A consequence of the coupling between partitioning and chemical equilibrium constants is that the forward and backward rate coefficients $k_\alpha^{\rightleftharpoons}$ in dilute mass action kinetics (Eq. (2.49)) generally break detailed balance of the rates, if it is chosen independently of the partitioning coefficient. The only choice resulting in the system reaching thermodynamic equilibrium is one where $k_\alpha^{\rightharpoonup}/k_\alpha^{\rightharpoondown} = \prod_{i=0}^{M} P_i^{\sigma_{i\alpha}}$.

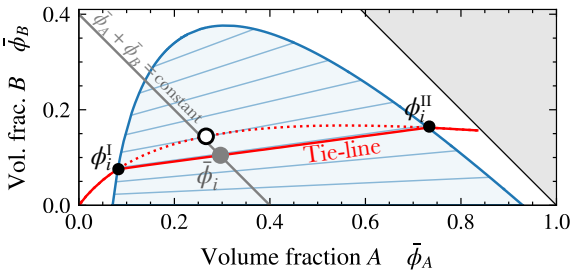


Figure 2.4: **Chemical equilibrium is altered upon phase separation:** For the reaction $A \rightleftharpoons B$, the chemical equilibrium line (red) satisfying $\mu_A = \mu_B$ is altered inside the binodal (thick blue line). When phase separated, chemical equilibrium (solid red line) deviates from chemical equilibrium when mixed (dotted red). The intersection of the chemical equilibrium line with the conserved line (gray) of constant $\bar{\phi}_A + \bar{\phi}_B$ sets the chemical equilibrium, which is different when homogeneous and phase separated. The homogeneous chemical equilibrium (white circle) is away from thermodynamic equilibrium, as diffusive equilibrium is not satisfied. The gray solid dot depicts the thermodynamic equilibrium state.

In figure 2.4, the change in the chemical equilibrium line upon phase separation is depicted. For a phase-separated system, there exists a single tie-line satisfying chemical equilibrium $(\Delta\mu_\alpha) = 0$. Along this tie-line, the system is at thermodynamic equilibrium, as both diffusive and chemical equilibrium are reached, which can be satisfied for all phase volumes $V^{\mathrm{I/II}}$. A mixed system with a composition within the binodal can reach chemical equilibrium, but is not at diffusive equilibrium. As seen in Fig. 2.4, this leads to a chemical equilibrium state different from the one where the system can reach diffusive equilibrium. Overlaying system trajectories with nullclines and binodals/spinodals, as done here, is a powerful way to understand the interplay between reactions, diffusion, and molecular interactions. By contrast, such interplay can often be difficult to infer from the solutions of the continuum equation, Eq. (2.54), alone.

2.3.2 Chemical kinetics at phase coexistence

The reaction-diffusion equation in Eq. (2.54) is highly non-linear with gradients up to fourth order (see Eq. (2.21)), making it challenging to solve analytically and expensive to solve numerically. One approach to reduce the complexity is to consider the interface width, $L_w \propto \sqrt{\kappa\nu/(k_BT)}$, to be much smaller than the system size L. In this limit, regular reaction-diffusion equations of the form $\partial_t\phi_i = D\nabla^2\phi_i + r_i$ can be solved in all phases, with a non-linear boundary condition of local phase equilibrium ($\mu_i^{\mathrm{I}} = \mu_i^{\mathrm{II}}$) for all components. For a thorough discussion, see Ref. [120, 121]. A further simplification can be made when the time scale of diffusion over the system size is much faster than the time scale of chemical reactions.

The dynamic Eq. (2.54) governing phase separation simplifies greatly when the diffusive exchange rate between phases is much faster than the reaction time scale over the system length,

$$L \ll \sqrt{\frac{\Lambda_i}{k_BTk_\alpha}}\,, \tag{2.59}$$

where L is the system size and $\sqrt{\Lambda_i/(k_BTk)}$ is the corresponding reaction-diffusion length scale. In this limit, the two phases (I, II) satisfy the conditions of phase equilibrium at all times (Eq. (2.29)). Furthermore, each phase is homogeneous in composition with different volume fractions in each phase ($\phi_i^{\mathrm{I}}, \phi_i^{\mathrm{II}}$) with corresponding phase volumes V^{I} and $V^{\mathrm{II}} = V - V^{\mathrm{I}}$. This leads to the simplified dynamic equations at phase equilibrium [113]:

$$\frac{\mathrm{d}\phi_i^{\mathrm{I/II}}}{\mathrm{d}t} = r_i^{\mathrm{I/II}} - j_i^{\mathrm{I/II}} - \phi_i^{\mathrm{I/II}}\frac{\dot{V}^{\mathrm{I/II}}}{V^{\mathrm{I/II}}}\,, \tag{2.60a}$$

$$\frac{\dot{V}^{\mathrm{I/II}}}{V^{\mathrm{I/II}}} = \sum_{k=0}^{M}\left(r_k^{\mathrm{I/II}} - j_k^{\mathrm{I/II}}\right)\,, \tag{2.60b}$$

where $i = 1, \ldots, M$, and the diffusive flux $j_i^{\mathrm{I/II}}$ is set to ensure the conditions of phase coexistence. Note that $j_i^{\mathrm{I/II}}$, is different from $\boldsymbol{j}_i$ in Eq. (2.20), as it quantifies the net number of particles of component i which diffuse between the phases per phase volume $V^{\mathrm{I/II}}$. A derivation for the expression of $j_i^{\mathrm{I/II}}$ is given in Appendix B. Phase-averaged quantities, denoted by a bar, for an arbitrary quantity X, become in this limit

$$\bar{X} \equiv \frac{V^{\mathrm{I}}X^{\mathrm{I}} + V^{\mathrm{II}}X^{\mathrm{II}}}{V^{\mathrm{I}} + V^{\mathrm{II}}}\,. \tag{2.61}$$

The kinetic equations in Eqs. (2.60) holds for incompressible systems, where the molecular volumes are constants $\mathrm{d}_t\nu_i = 0$.

A comparison between the evolution of the volume fractions found in the fast diffusion limit and from the continuum equation (Eq. (2.54)) in each phase $\phi_i^{\mathrm{I/II}}$ and the phase-averaged $\bar{\phi}_i$ is depicted in Fig. 2.5 (a) and (b) respectively. In this

example, we consider a simple uni-molecular reaction $A \rightleftharpoons B$. The evolution of the composition in each phase $\phi_i^{\mathrm{I/II}}$ is very close to the binodal for a characteristic system length $L = 0.15\sqrt{\Lambda_i/(k_B T k)}$, and would improve as the system size is reduced. Due to the fast-diffusion limit ignoring the energy cost of the interface, the thermodynamic equilibrium values in each phase deviate slightly from the ones predicted from the full continuum equation.

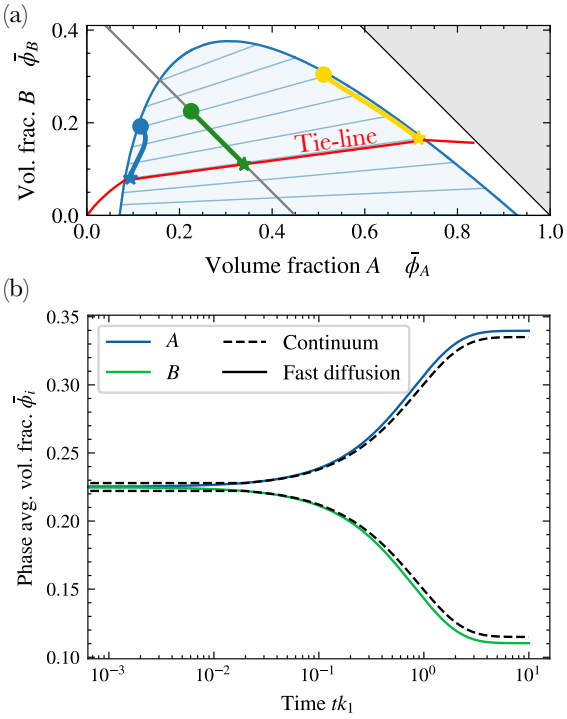


Figure 2.5: **Comparison of fast diffusion with continuum equation:** (a) The evolution of the volume fractions $\phi_i^{\mathrm{I/II}}$ (blue/yellow) when solving the full continuum equations (Eq. (2.54) with a reaction $A \rightleftharpoons B$) for a system size of $L = 0.15\sqrt{Dk}$ is close to the binodal (blue line). (b) For each time point t, the volume fractions are close for both $\bar{\phi}_A$ and $\bar{\phi}_B$, eventually equilibrating with a slight deviation due to the fast-diffusion limit ignoring the surface tension originating from the κ's.

Treating reactions as slow compared to diffusion is a powerful framework. For one, it is an accurate approximation for small system sizes, where reaction rates are commonly on the order of minutes, and diffusion can occur on a time scale of less than seconds [122]. Furthermore, transforming the coupled partial differential equations of Eq. (2.54) to the coupled ordinary differential equations in Eq. (2.60) reduces the numerical costs of solving the equations significantly. With a reduction in the complexity of the equations, we can allow for an increased number of components at non-dilute conditions. One of the fundamental advantages of this approach is that trajectories can be meaningfully and correctly depicted in phase diagrams. Being able to depict the trajectory together with nullclines and binodals/spinodals is a powerful method for understanding the underlying interplay of different processes, such as the effects of interactions between molecules on chemical reactions. The same understanding can be hard to retrieve from analyzing the evolution of volume fractions in space and time, when solving Eq. (2.54).

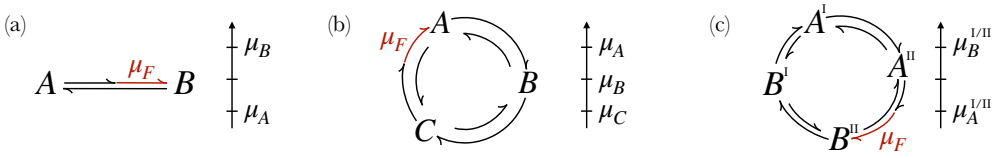


Figure 2.6: **Non-equilibrium steady state reaction flux diagrams:** For three different systems, the steady state reaction fluxes are depicted, with detailed-balance obeying fluxes being displayed by black harpoons, and fueled fluxes by red harpoons. For a unimolecular reaction in a homogeneous system (a), there are no net fluxes in a non-equilibrium steady state. Instead, (b) three components are necessary for a net flux between the states in the steady state. If the system is separated into two phases (I and II), there can be net fluxes for the same reaction network as in (a), as long as the fuel is phase dependent, as shown in (c). Due to the coexisting phases, there can be net fluxes in the non-equilibrium steady state. In (c), some fluxes are reaction fluxes and others are diffusive fluxes.

2.3.3 Non-equilibrium steady states

For a homogeneous system, a unimolecular chemical reaction between A and B, $A \rightleftharpoons B$, will, even in the presence of a fuel, lead to a non-equilibrium steady state with no net fluxes. There is still a measurable entropy production if the individual fluxes from each pathway can be measured, as $\mu_B > \mu_A$ for a fuel aiding the turnover from A to B, as depicted in Fig. 2.6(a). Nevertheless, the net dynamics can be mapped to an equilibrium system, where the chemical potential of the fuel is included in the internal energy of B, that is $\omega_A \to \omega_A + \mu_F$.

To find a non-equilibrium steady state (NESS) in a homogeneous mixture, an additional component in the reaction network is necessary. The minimal example is depicted in Fig. 2.6(b), where the three components A, B, and C are all connected in the reaction network. Here, the fuel aids the turnover from C to A, such that $\mu_A > \mu_B > \mu_C$, allowing a chemical cycle with fluxes in the non-equilibrium steady state.

For a system consisting of multiple phases, the effective number of components increases. Therefore, a uni-molecular reaction, like $A \rightleftharpoons B$, can create net fluxes in the non-equilibrium steady state, if and only if the fuel acts differently in the two phases. In Fig. 2.6(c), such a flux diagram is depicted, where a fuel aids the turnover from A to B in phase II. Non-homogeneous mixtures thus allow for fluxes in the non-equilibrium steady state even for simple reaction networks. The dynamics are therefore not mappable to an equilibrium system, unlike in the equivalent homogeneous system.

For the minimalistic reaction between A and B where a fuel $\Delta\mu^{\mathrm{I/II}}$ acts in both

phases, the reaction rate in each phase follows

$$r^{\mathrm{I/II}} = k^{\mathrm{I/II}} \left(\exp\left\{ \frac{\mu_A^{\mathrm{I/II}} + \Delta\mu^{\mathrm{I/II}}}{k_B T} \right\} - \exp\left\{ \frac{\mu_B^{\mathrm{I/II}}}{k_B T} \right\} \right) . \tag{2.62}$$

Here, $\Delta\mu^{\mathrm{I/II}}$ relates to the chemical potential of the fuel and waste through Eq. (2.53). We will treat the system at diffusive equilibrium, an approximation which holds when diffusion is much faster than reactions, as discussed in the previous section. The nullcline of the reaction, meaning the points in phase space where the net reaction across the system is zero, occurs for two phases if

$$\left(V^{\mathrm{I}} r^{\mathrm{I}} + V^{\mathrm{II}} r^{\mathrm{II}} \right) \big|_{\mathrm{NESS}} = 0 \, . \tag{2.63}$$

For the given reaction (Eq. (2.62)), we find that the volume of the dense phase V^{II} in the non-equilibrium steady state (NESS) obeys

$$V^{\mathrm{II}}_{\mathrm{NESS}} = \frac{k^{\mathrm{I}} \left(1 - \exp\left\{ \frac{\mu_A - \mu_B + \Delta\mu^{\mathrm{I}}}{k_B T} \right\} \right)}{k^{\mathrm{II}} \left(\exp\left\{ \frac{\mu_A - \mu_B + \Delta\mu^{\mathrm{II}}}{k_B T} \right\} - 1 \right) - k^{\mathrm{I}} \left(\exp\left\{ \frac{\mu_A - \mu_B + \Delta\mu^{\mathrm{I}}}{k_B T} \right\} - 1 \right)} \, . \tag{2.64}$$

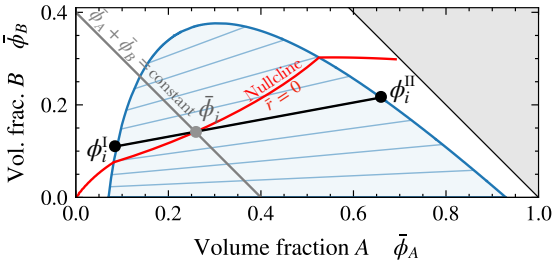


Figure 2.7: **Nullcline for a fueled reaction:** The reaction nullcline for the reaction $A \rightleftharpoons B$ where a phase dependent fuel $\Delta\mu^{\mathrm{I/II}}$ aids the turnover from A to B according to Eq. (2.62) is depicted. The nullcline is set by the points where the system-averaged reaction is zero (Eq. (2.63)). The non-equilibrium steady state is set by the intersection between the nullcline and the constraint line of constant $\bar{\phi}_A + \bar{\phi}_B$.

For a non-fueled chemical reaction, as depicted in Fig. 2.4, the chemical equilibrium is independent of the phase volumes, resulting in a single tie-line satisfying $\mu_A = \mu_B$ being selected at chemical equilibrium. For a system with phase-dependent fuel, there is a single phase volume that results in a steady state for each value of μ_A and μ_B, meaning for each tie-line with physical solutions for the phase volumes. In Fig. 2.7, such a nullcline is depicted. The intersection between the nullcline and the constraint line of constant $\bar{\phi}_A + \bar{\phi}_B$ sets the non-equilibrium steady state. In the steady state, there are non-zero fluxes between the different components in each phase, as shown in Fig. 2.6c. The above analysis holds as long as the system is at diffusive equilibrium, obeying the conditions of phase coexistence.

Chapter 3

Accelerating chemical reactions through external reservoirs

Without enzymatic catalysis, chemical reactions such as oligomerization proceed much more slowly than competing processes like fragmentation. Driving reactions to overcome thermodynamic barriers is therefore one of the central challenges of abiotic chemistry. Recent experiments indicate that wet–dry cycles can substantially alter reaction pathways, potentially accelerating otherwise slow processes [56, 123, 124, 125]. In this chapter, we present a theoretical framework for nondilute chemical reactions under non-equilibrium conditions of evaporation and condensation. The analysis shows that such cycling can change product half-times by more than an order of magnitude, depending on substrate–solvent interactions. Moreover, the cycle frequency strongly influences chemical turnover and can even produce a resonance where turnover is maximal. This behavior allows wet–dry cycles to favor specific reactions, pointing to a potential mechanism for chemical evolution in abiotic systems. This chapter contains results which have been published in Ref. [1], stemming from work done in collaboration with Pranay Jaiswal, and Christoph A. Weber, and in Ref. [6], stemming from work in collaboration with Pranay Jaiswal, Hidde D. Vuijk, and Christoph A. Weber. For reproducibility, figure parameters are listed in Appendix M.1. For reading convenience, we repeat parts of the necessary theory from chapter 2, as we introduce the model.

3.1 Introduction

The presence of catalysts severely alters the speed of chemical reactions by affecting the activation energy, without changing the thermodynamic free energies of the substrate and product state [126]. In contrast to reactants, they are not converted during the turnover of a substrate to a product. The use of catalysts ensures economic

efficiency of most applications in chemical and biomolecular engineering [127, 128]. In living systems, catalytic activity is mediated by enzymes. These enzymes are specialized proteins involved in almost all metabolic processes in the cell [129]. Enzymes do not solely maximize the speed of chemical reactions; they are optimal concerning reaction specificity [130], i.e., they catalyze only specific reactions in a mixture of extremely many reacting components. However, this optimal catalytic property for life had to develop during evolution [131]. In other words, there were likely no enzymes at the molecular origin of life. It remains one of the mysteries of the molecular origin of life how functional products such as self-replicating RNA strands could have formed despite activation barriers significantly larger than k_BT [132].

An alternative perspective on how to speed up chemical reactions is through non-equilibrium conditions [53, 60, 133, 134]. In fact, non-equilibrium applies to many processes in chemical engineering, almost all processes in living cells, and prebiotic soup at early Earth. The reason is that these are all open systems that can exchange matter and entropy with their environment, termed dissipative systems [135]. Continuous dissipation is often achieved by cycles of the system's control parameters, such as temperature or chemical potentials. Recently studied cases are wet-dry cycles [56, 55, 123, 124, 125, 136].

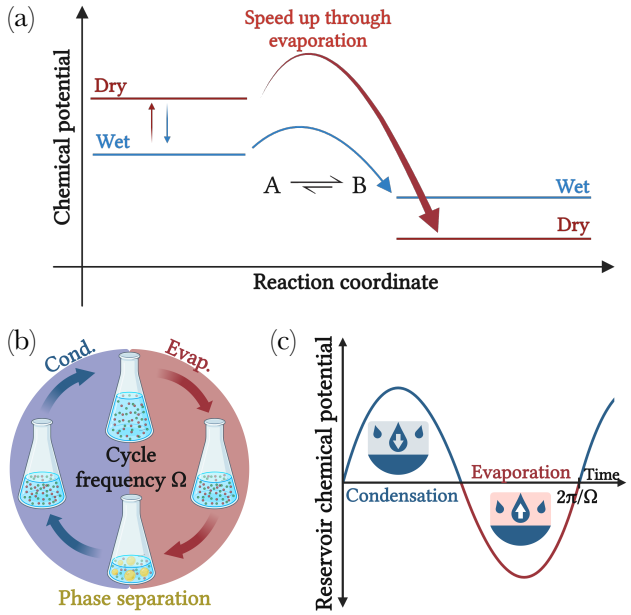


Figure 3.1: **Wet-dry cycles in chemically reacting systems:** (a) Wet-dry cycles can enhance chemical reaction rates by increasing the chemical potential of the substrate μ_A and decreasing the chemical potential μ_B of the product. (b) Wet-dry cycles lead to non-dilute conditions, which affect chemical reactions and may induce liquid-liquid phase separation. (c) The cyclic behavior is included through an oscillating reservoir with periods of evaporation and condensation.

The ability of wet-dry cycles to affect the synthesis of compounds that are also relevant for the molecular origin of life is progressively backed by experiments [137, 138, 123]. Wet-dry cycles have been applied to catalyze the formation of polyacids [139], esters [137, 138], oligopeptides [140], nucleosides [123], polynucleotides [141], phosphorylation [56] and lipid structures [142]. Up to now, the mechanisms by which wet-dry cycles affect chemical processes remain elusive.

Unraveling such mechanisms is complex since reducing the solvent amount through evaporation generally leads to non-dilute conditions. As a result, non-linear thermodynamic contributions to chemical activities emerge. Additionally, non-dilute systems can induce phase transitions, such as liquid-liquid phase separation [87, 80], gel formation [143], and solid precipitation [144, 145]. Recent experiments show that coacervate formation can be triggered by wet-dry cycles [55]. To interpret the existing experimental studies on the effects of wet-dry cycles on the synthesis of chemical building blocks and to decipher the underlying physicochemical mechanisms, a theoretical framework is lacking.

This work proposes a kinetic theory based on thermodynamic principles for a non-dilute chemically reacting mixture coupled to a cyclic wet-dry reservoir that drives the system out of equilibrium; see Fig. 3.1b,c. We show to what extent and under which conditions both evaporation and condensation of solvent can speed up or slow down rates of chemical reactions. In contrast to catalysts that solely affect activation barriers, we find that solvent evaporation and condensation can significantly alter the chemical potentials of substrate and product once the system is driven toward non-dilute conditions, illustrated in Fig. 3.1a. We also study how wet-dry cycles allow for continuous chemical activity by maintaining the system out of equilibrium. For this case, a resonance cycling frequency exists for each evaporation/condensation flux at which the chemical turnover is maximal. This resonance behavior provides a selection pressure for specific chemical reactions, suggesting wet-dry cycles as an abiotic alternative to a catalyst in order to kick-start the molecular evolution of life [146, 147].

3.2 Theory for wet-dry cycles with chemical reactions

We consider a mixture of volume V composed of $M+1$ components, each of particle number N_i ($i = 0, 1, ..., M$) in the Gibbs (T-P-N_i) ensemble with the Gibbs free energy $G(T, P, N_i)$. Evaporation and condensation of molecule i is governed by the difference between the bulk chemical potential $\mu_i = (\partial G/\partial N_i)_{T,p,N_{j\neq i}}$, and the reservoir chemical potential μ_i^{r}. In all our studies, the chemical potential of the reservoir μ_i^{r} is a control parameter that can be constant or vary in time. The evaporation and condensation flux h_i describes the number of molecules passing through the surface area A per unit time:

$$h_i = k_{\mathrm{e},i}\left[\exp\left\{\frac{\mu_i^{\mathrm{r}}}{k_{\mathrm{B}}T}\right\} - \exp\left\{\frac{\mu_i}{k_{\mathrm{B}}T}\right\}\right], \tag{3.1}$$

where $k_{\mathrm{B}}T$ denotes the thermal energy. Here, $k_{\mathrm{e},i}$ is the kinetic evaporation/condensation coefficient for component i. Equilibrium is reached when the chemical potentials between the reservoir and the bulk are equal, $\mu_i = \mu_i^{\mathrm{r}}$, implying no exchange between the bulk and the reservoir, $h_i = 0$. Note that Eq. (3.1) considers a symmetric form, $h_i = h_i^{\mathrm{cond}} - h_i^{\mathrm{evap}}$, of the condensation flux, $h_i^{\mathrm{cond}} = k_{\mathrm{e},i} \exp\{\mu_i^{\mathrm{r}}/k_{\mathrm{B}}T\}$, and evaporation flux, $h_i^{\mathrm{evap}} = k_{\mathrm{e},i} \exp\{\mu_i/k_{\mathrm{B}}T\}$, satisfying detailed balance of the rates, $h_i^{\mathrm{cond}}/h_i^{\mathrm{evap}} = \exp\{(\mu_i^{\mathrm{r}} - \mu_i)/k_{\mathrm{B}}T\}$.

To calculate the chemical potential, we consider the mean field Gibbs free energy G per volume V [148]:

$$\frac{G(T,p,\{\phi_k\})}{V} = k_{\mathrm{B}}T \sum_{i=0}^{M} \left[\frac{\phi_i}{\nu_i} \log \phi_i + \frac{\omega_i \phi_i}{\nu_i k_B T} \right] + \frac{1}{2\nu_0} \sum_{i=0}^{M} \sum_{j=0}^{M} \chi_{ij} \phi_i \phi_j + p \,, \tag{3.2}$$

where χ_{ij} is the interaction parameter characterizing the interaction between component i and j. Moreover, ω_i and ν_i denote the internal energy and molecular volume of component i, respectively. The chemical potential of component i, $\mu_i = \partial G/\partial N_i|_{T,p,N_{j\neq i}}$, thus becomes

$$\mu_i = k_B T \left(\log \phi_i + 1 \right) + \omega_i + \nu_i \left(p - \Lambda \right) + \frac{\nu_i}{\nu_0} \sum_{j=0}^{M} \chi_{ij} \phi_j \,, \tag{3.3}$$

where Λ is given by

$$\Lambda = \frac{1}{2\nu_0} \sum_{i=0}^{M} \sum_{j=0}^{M} \chi_{ij} \phi_i \phi_j + k_B T \sum_{i=0}^{M} \frac{\phi_i}{\nu_i}. \tag{3.4}$$

The volume fraction of a component is defined by the volume it occupies relative to the total system volume, $\phi_i = \nu_i N_i / V$, $i \in \{1, 2, ..., M\}$, where the solvent volume fraction ϕ_0 is set by all other volume fractions

$$\phi_0 = 1 - \sum_{i=1}^{M} \phi_i \,. \tag{3.5}$$

The total volume change due to evaporation and condensation is equal to the volume of molecules removed from or added through the surface area A per unit of time

$$\dot{V} = \sum_{i=0}^{M} \nu_i h_i \,. \tag{3.6}$$

Moreover, the volume change alters the volume fractions of all components, even components where $h_i = 0$, as $\mathrm{d}_t \phi_i = \nu_i h_i / V - \phi_i \dot{V}/V$, and thus affects their chemical

potentials and chemical reactions. A reaction α is written as

$$\sum_{i=0}^{M} C_i \sigma_{i\alpha}^{\rightharpoonup} \rightleftharpoons \sum_{i=0}^{M} C_i \sigma_{i\alpha}^{\leftharpoondown} \,, \tag{3.7}$$

where C_i denoted a chemical component and $\sigma_{i\alpha}^{\rightleftharpoons}$ are the are stoichiometric matrix elements [113]. For the forward and backward chemical reactions rates $r_{\alpha}^{\rightleftharpoons}$ we choose a symmetric form [118]

$$r_{\alpha}^{\rightleftharpoons} = k_{c,\alpha} \exp\left\{ \frac{\sum_{i=0}^{M} \sigma_{i\alpha}^{\rightleftharpoons} \mu_i}{k_B T} \right\}. \tag{3.8}$$

From the chemical potential (3.3), we see that the forward and backward reaction rates are proportional to the product of their volume fractions ϕ_i. This implies that once a component is depleted, reactions slow down and halt upon reaching zero volume fractions. The ratio between the forward and backward reaction rates satisfies detailed balance of the rates [118]:

$$\frac{r_{\alpha}^{\rightharpoonup}}{r_{\alpha}^{\leftharpoondown}} = \exp\left\{ -\frac{\sum_{i=0}^{M} \sigma_{i\alpha} \mu_i}{k_B T} \right\}, \tag{3.9}$$

where we have defined $\sigma_{i\alpha} \equiv \sigma_{i\alpha}^{\leftharpoondown} - \sigma_{i\alpha}^{\rightharpoonup}$. The difference between the forward and backward reaction pathways, for R different reactions, drives the net reaction rate

$$r_i = \sum_{\alpha=1}^{R} \sigma_{i\alpha} \nu_i \left(r_{\alpha}^{\rightharpoonup} - r_{\alpha}^{\leftharpoondown} \right). \tag{3.10}$$

If the interaction strengths cross a certain threshold value, the system will phase separate into two phases (denoted I and II) with different composition $\phi_i^{\mathrm{I/II}}$, but same chemical potential

$$\mu_i^{\mathrm{I}} = \mu_i^{\mathrm{II}}. \tag{3.11}$$

The time evolution of the non-solvent volume fraction of a phase-separated system follows [113, 87, 134, 97]:

$$\partial_t \phi_i^{\mathrm{I/II}} = r_i^{\mathrm{I/II}} - j_i^{\mathrm{I/II}} + \frac{\nu_i h_i^{\mathrm{I/II}}}{V^{\mathrm{I/II}}} - \phi_i^{\mathrm{I/II}} \frac{\dot{V}^{\mathrm{I/II}}}{V^{\mathrm{I/II}}} \,, \tag{3.12}$$

where $j_i^{\mathrm{I/II}}$ gives the diffusive flux between the phases. The volume change of each phase for volume-conserving chemical reactions is given by

$$\frac{\dot{V}^{\mathrm{I/II}}}{V^{\mathrm{I/II}}} = \sum_{i=0}^{M} \left(\frac{\nu_i h_i^{\mathrm{I/II}}}{V^{\mathrm{I/II}}} - j_i^{\mathrm{I/II}} \right). \tag{3.13}$$

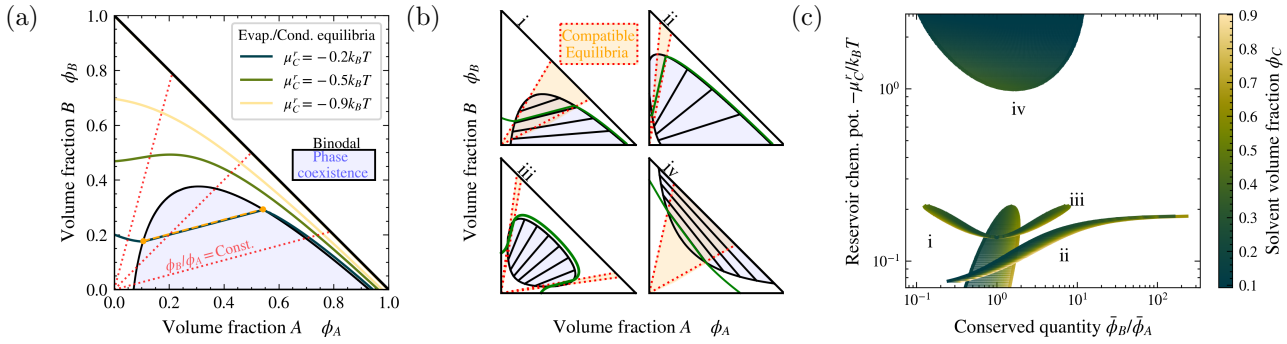


Figure 3.2: **Evaporation/condensation phase diagrams:** (a) Evaporation/condensation equilibrium for different reservoirs is displayed together with the constraint line of constant ϕ_B/ϕ_A. The binodal encloses the blue-shaded area of phase coexistence, where the system phase separates into two phases connected by straight black tie lines. (b) For four sets of interaction strengths, creating four qualitatively different phase diagrams (i-iv), the orange-shaded region displays the initial states resulting in a phase-separated equilibrium. (c) The constraint of constant ϕ_B/ϕ_A at phase-separated equilibria is qualitatively different for the different types of diagrams when varying the reservoir through μ_C^r.

To determine the diffusive flux j_i, we use the $(M+1)$ constraints from phase equilibria (Eq. (3.11)). See chapter 2.3.2 for further discussion and appendix B for a derivation. For a homogeneous system, the governing equations are retrieved by setting $j_i^{\mathrm{I/II}} = 0$, and dropping the phase specifying superscript.

For a phase-separated system, chemical reactions occur simultaneously in both phases, while the fluxes related to evaporation/condensation are considered to occur at a surface area enclosing the dilute phase. As the two phases are at phase equilibrium, the choice of which phase is in contact with the reservoir becomes irrelevant. Furthermore, we will assume that the reaction rate coefficient does not depend on the phase $k_{c,\alpha}^{\mathrm{I}} = k_{c,\alpha}^{\mathrm{II}}$. We consider that evaporation/condensation of macromolecules is negligible compared to the solvent [149], such that $k_{e,i} = k_e\delta_{iC}$. When only the solvent is evaporating/condensing, and no chemical reactions occur, the ratio of each non-solvent volume fraction, ϕ_i/ϕ_j, is conserved. To reduce dimensionality and the number of free parameters, we will consider a single unimolecular reaction $A \rightleftharpoons B$, such that $\sigma_{i\alpha} = \delta_{iA} - \delta_{iB}$, and $k_{c,\alpha} = k_c$. The kinetic algorithm to evolve the system according to Eq. (3.12) builds on a method derived in Ref. [113]; see chapter 2.3.2 for how the theory accounts for the phase equilibrium constraint, and for the properties a system should have for phase equilibrium to hold well at each time during the kinetics.

3.3 Thermodynamics of evaporation/condensation

To understand chemical kinetics in systems that can phase separate and undergo wet-dry cycles, it is instrumental first to understand equilibrium states in the absence of chemical reactions. Equilibrium states composed of coexisting phases form through evaporation/condensation if the equilibria are compatible. Compatible equilibria refer to the case where the phase equilibrium condition (Eq. (3.11)) can be concomitantly satisfied together with one or more other equilibrium conditions [113]. In the absence of chemical processes, this other condition corresponds to evaporation-condensation equilibrium. Graphically, such compatible equilibria correspond to the intersections in Fig. 3.2a between the binodal and the nullcline that is set by the reservoir chemical potential μ_C^{r} (orange dashed line). During evaporation and condensation, ϕ_B/ϕ_A is conserved. Thus, the domain of compatible equilibria is cone-shaped and is spanned by two distinct conservation lines of constant ϕ_B/ϕ_A (orange domains in Fig. 3.2b). Any initial state will move during its evaporation/condensation kinetics and settle in a compatible equilibrium with coexisting phases. The number and the size of domains with compatible equilibria depend on the interactions among the molecules. The area of compatible domains increases when A and B interact more differently with the solvent (case i in contrast to ii and iii in Fig. 3.2b,c), or when A and B phase separate independently of the solvent (case iv in Fig. 3.2b,c).

The equilibrium states are homogeneous when evaporation-condensation and phase equilibrium are incompatible. Such incompatible states graphically correspond to the complement of the orange shaded domains in Fig. 3.2b, i.e., the values of ϕ_B/ϕ_A that lie outside of the nullcline's intersection with the binodal. Incompatible equilibria are favored when the A and B interact similarly with the solvent, and when neither A nor B phase separates with the solvent, as discussed for compatible equilibria. All equilibria are incompatible for some values of the reservoir chemical potential (green and yellow lines in Fig. 3.2a).

The phase behavior of the evaporating/condensing mixtures (i-iv) without chemical reactions can be summarized by the thermodynamic phase diagram spanned by the reservoir chemical potential μ_C^{r} and the conserved variable ϕ_B/ϕ_A. The different interactions (i-iv) lead to very different shapes where two phases can coexist with very different values of the conserved variable ϕ_B/ϕ_A, as shown in Fig. 3.2c. For case iv, where the coexisting phases have a similar solvent volume fraction, the non-solvent components phase-separate independently of the solvent. As a result, each phase has a large difference in ϕ_B/ϕ_A, creating a significant phase-separated domain.

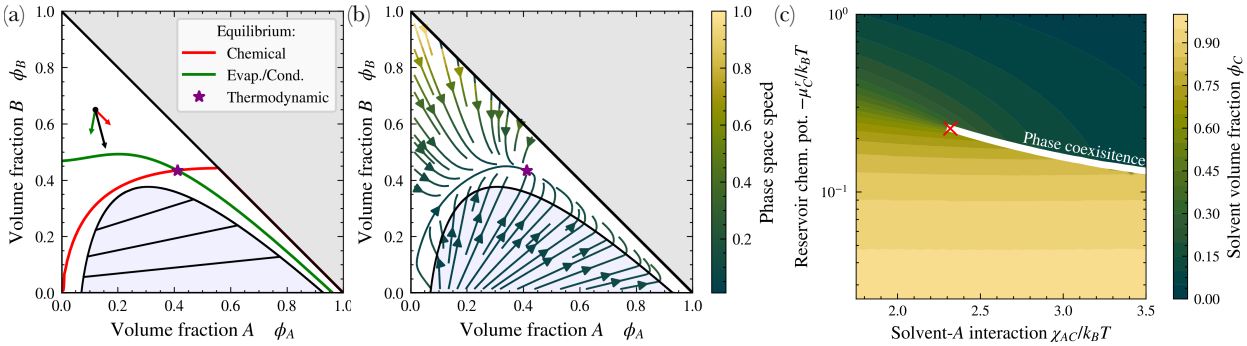


Figure 3.3: **Evaporation/condensation and chemically reacting phase diagrams:** (a) The intersection of the equilibrium lines of evaporation/condensation and the chemical reaction gives the thermodynamic equilibrium point that all initial states converge towards. An example of the change in volume fraction is displayed as a black arrow, which is the sum of a chemical (red) and evaporation/condensation component (green). (b) The set of all such points produces a flow field in the phase diagram, all converging at the thermodynamic equilibrium point, displayed for $k_e/k_c = 4$. (c) Phase-separated thermodynamic equilibrium states are only achievable for a single reservoir value for each interaction strength. The phase coexistence line separates solvent-poor and solvent-rich states until the critical point (red cross).

3.4 Evaporation/condensation of chemically reacting mixtures

In the presence of chemical reactions, ϕ_B/ϕ_A is not conserved, and compatible equilibria are no longer achievable. Thermodynamic equilibrium, therefore, always corresponds to a homogeneous state where the conditions of chemical equilibrium ($\mu_A = \mu_B$) and evaporation-condensation equilibrium ($\mu_C = \mu_C^r$) are fulfilled; see the intersection between the red and green lines in Fig. 3.3a. Any initial state (ϕ_A, ϕ_B) will evolve following a flow field in the phase diagram toward the fixed point of thermodynamic equilibrium; see Fig. 3.3b. Each point in this flow diagram has two independent directions that characterize the rates of change in the average volume fractions, which correspond to the chemical turnover flux of constant $\phi_A + \phi_B$ and evaporation-condensation flux of constant ϕ_B/ϕ_A; see vectors shown in Fig. 3.3a. Such fluxes of strength visualized by the length of each vector are set by k_e and k_c. Gibbs' phase rule (discussed in chapter 2.1.5) makes the domain of coexisting states inaccessible at thermodynamic equilibrium. This results in small changes in reservoir chemical potential μ_C^r close to the binodal line, which yields significant changes in the equilibrium compositions. This behavior is shown in Fig. 3.3c, where solvent-rich and solvent-poor equilibrium states are separated by the phase coexistence line ending up at the critical point (red cross) in Fig. 3.3c.

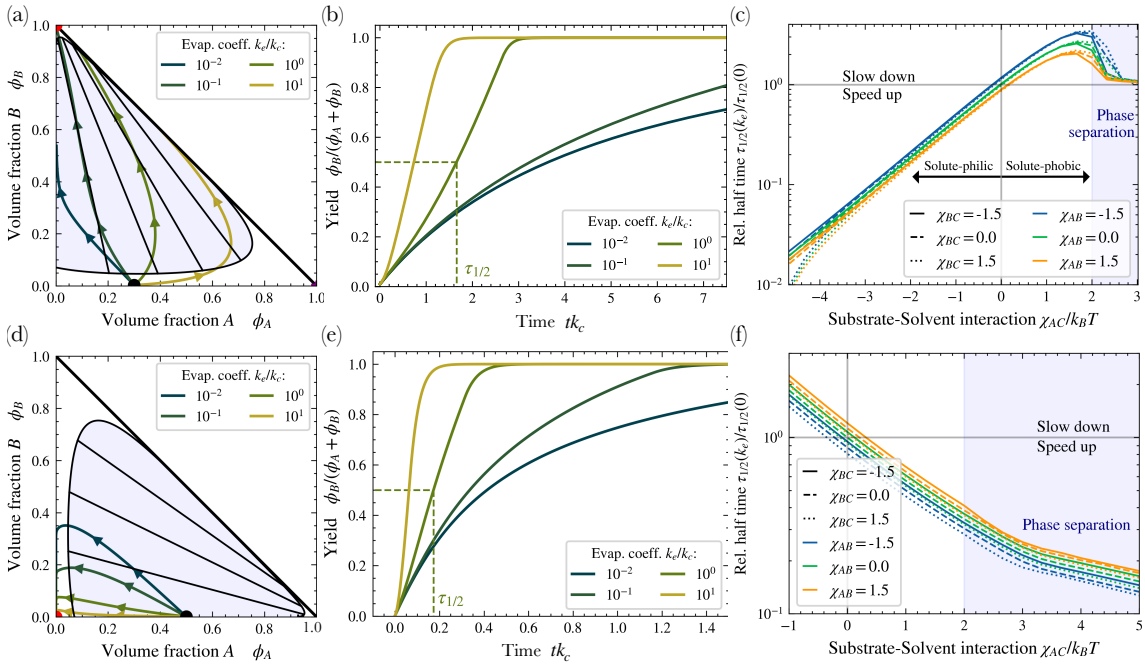


Figure 3.4: **Chemical speed-up through evaporation/condensation:** The pathway through the phase diagram during evaporation (a-c) and condensation (d-f) towards the same equilibrium (red dot) for different k_e/k_c is shown in (a) and (d), respectively. The yield of the product B, $\phi_B/(\phi_A+\phi_B)$, is displayed as a function of time in (b) and (e), showing a significant speedup through the different pathways. How this speedup depends on the interaction strengths for $k_e/k_c = 10$ is displayed in (c) and (f). The top row (a-c) shows the results for evaporation, and the bottom row for condensation (d-f).

3.5 Accelerating chemical reactions through evaporation and condensation

Evaporation and condensation can speed up chemical reactions by driving the system along different phase diagram trajectories. For unidirectional chemical reactions, where $\mu^0_{\text{product}} \ll \mu^0_{\text{substrate}}$, as defined in Eq. (2.16), the trajectories are characterized by solvent-poor or solvent-rich conditions during the relaxation to thermodynamic equilibrium (Fig. 3.4a,d). These different solvent conditions have a dramatic impact on the time ($\tau_{1/2}$) it takes to turn over 50% of A to B, seen as the half-time of the yield $\phi_B/(\phi_A+\phi_B)$ in Fig. 3.4b,e. For the depicted molecular interactions, i.e., A favors the solvent while B does not, a quick removal of the solvent (large k_e/k_c) causes an increase in the chemical potential of A. This increase in the substrate's chemical potential, visualized in Fig. 3.1a, speeds up the turnover from A to B.

The threshold between slowdown and speedup of chemical reactions corresponds to no interactions between the components ($\chi_{ij} = 0$), as seen in Fig. 3.4c and 3.4f. Thus, solely up-concentrating the components does not affect the speedup, which is purely a consequence of non-dilute conditions and the interactions among reacting components.

The role of interactions on speedup can be understood through the chemical potentials. For unidirectional chemical reactions, where $\omega_{\text{product}} \ll \omega_{\text{substrate}}$, the direction from substrate to product dominates, as the backward pathway relative to the forward pathway is exponentially damped by the factor $(\mu^0_{\text{product}} - \mu^0_{\text{substrate}})/k_{\text{B}}T$. The chemical reaction in equation (3.10) can, under these conditions, be approximately written as

$$r_{A\rightharpoonup B} - r_{A\leftharpoondown B} \approx \tilde{k}_c \phi_A \exp\left\{\frac{\gamma_A}{k_{\text{B}}T}\right\}, \tag{3.14}$$

when A is acting as the substrate and B as the product. We have defined an effective reaction rate coefficient $\tilde{k_c} \equiv k_c \exp\{\mu^0_A/k_{\text{B}}T\}$, and γ_A is the activity coefficient, defined as

$$\gamma_A = \frac{\nu_A}{\nu_C}\phi_B\chi_{AB} + \frac{\nu_A}{\nu_C}\phi_C\chi_{AC} - \nu_A\Lambda\,. \tag{3.15}$$

To speed up the chemical reaction, this quantity should be maximized. During evaporation, we reduce ϕ_C, increasing ϕ_A and ϕ_B, such that a negative value of χ_{AC} increases the substrate's chemical potential. While for condensation, ϕ_C is increased at the expense of ϕ_A and ϕ_B, such that the chemical potential is increased for positive values of χ_{AC}. Though the interaction between χ_{AB} might appear as important as χ_{AC}, the volume fraction of the product will generally vary less during evaporation/condensation. However, it will become important towards the latter part of the turnover process. Importantly, Λ, defined in Eq. (3.4), also depends on the volume fractions and will vary during evaporation/condensation. The changes in Λ will contribute less to the overall speedup as its interaction terms contribute only as second order in the ϕ's.

The sign of χ determines whether evaporation or condensation aids the chemical turnover, while the magnitude determines the speedup. In other words, reactions of solvophilic substrates are sped up through evaporation, and solvophobic substrates under condensation. Though other interactions (χ_{AB}, χ_{BS}) also affect the speedup, they are less relevant than the substrate-solvent interaction (χ_{AS}), as seen in figures 3.4c and 3.4f. All statements about interaction strengths are qualitatively reversed once we switch from evaporation (Figs. 3.4a,b,c) to condensation (Figs. 3.4d,e,f).

3.6 Chemical reactions subject to wet-dry cycles

Now we discuss the effects on chemical reactions when the system is subject to cycles of the reservoir chemical potential

$$\mu^r_C(t) = \langle \mu^r_C \rangle + \mu^{\text{amp}}_{C,r} \sin(\Omega t)\,. \tag{3.16}$$

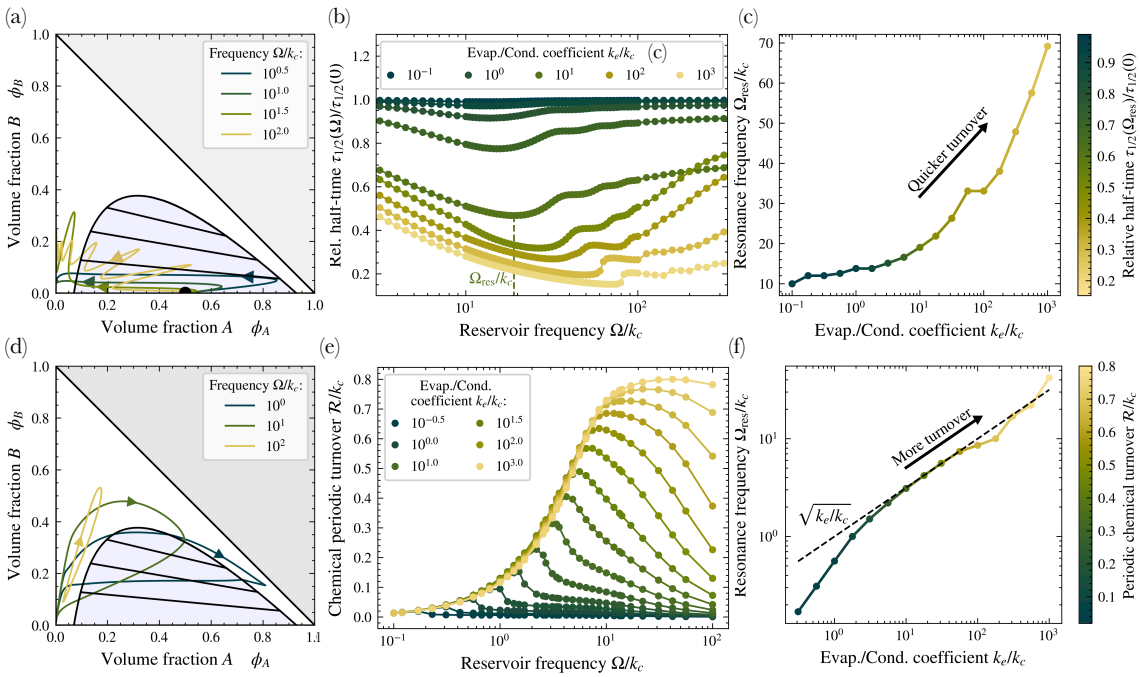


Figure 3.5: **Resonance behavior in chemical reactions with an oscillating reservoir:** The top row (a-c) shows the conversion from A to B on the path towards an orbit, while the bottom row (d-f) shows equivalent quantities during the orbits. The change in the phase-diagram pathway with cycling frequency Ω is displayed in (a) and (d) for $k_e/k_c = 10^{1.5}$ and $10^{2.75}$, respectively. Both cases exhibit a resonance behavior of the chemical turnover, quantified by the conversion half-time (b,c) and periodic chemical turnover (Eq. (3.17)) (e,f). Both values increase with k_e/k_c at the resonance frequency.

Here, Ω denotes the cycling frequency, $\langle \mu_C^r \rangle$ the average reservoir value and $\mu_{C,r}^{\text{amp}}$ the oscillation amplitude. We will consider two cases of cycles: (i) the chemical path toward the orbit, and (ii) the long-time periodic behavior of the orbit. For case (i), we are interested in the chemical turnover for a unidirectional chemical reaction, as before, and how it is affected by cycles. After the initial phase, the trajectory becomes periodic. In case (ii), we are interested in the chemical turnover rate during such an orbit.

(i) We find that wet-dry cycles can speed up chemical turnover relative to no cycling, with a resonance peak at a specific cycling frequency. Different cycling frequencies Ω result in the different pathways displayed in Fig. 3.5a, with different conversion half-times. Oscillating reservoirs decrease the conversion half-time relative to no oscillations, $\tau_{1/2}(\Omega)/\tau_{1/2}(0)$, by up to a factor of 5 for the parameters used in Fig. 3.5b. Weaker local maxima are observed for the harmonics of the resonance frequency, i.e. $\Omega = n\Omega_{\text{res}}$ for $n = 2, 3, \ldots$. The harmonics emerge due to the system being in the dry/wet-state an additional time for each higher harmonic. The resonance cycling frequency Ω_{res} is stable, changing less than a decade when varying k_e/k_c over four decades, displayed in Fig. 3.5c.

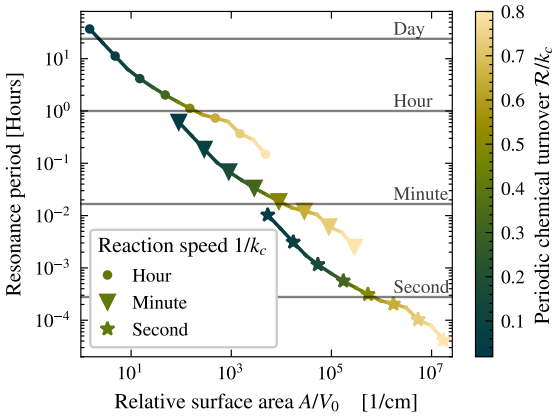


Figure 3.6: **Resonance behavior determined by system's geometry:** The resonance frequency of the periodic chemical turnover in Fig. 3.5f can be reached by varying the geometry of the mixture through its surface-to-volume ratio. Different reaction timescales lead to different resonance frequencies, making the resonance behavior act as a selection mechanism. For this figure, the evaporation speed is 5 cm/day.

(ii) With repeated cycling, the system will eventually enter an orbit, where the path in the phase diagram is a closed loop (Fig. 3.5d). To quantify the chemical activity per period, we define the chemical turnover per cycle,

$$\mathcal{R} = \frac{\Omega}{2\pi} \int_0^{2\pi/\Omega} \mathrm{d}t |\vec{r} - \overleftarrow{r}| \,. \tag{3.17}$$

This quantity is at its minimal $\mathcal{R} = 0$ if no chemical reactions take place during the cycle, and maximal $\mathcal{R} = \Omega/\pi$ if all molecules undergo both the forward and backward reaction per cycle.

Different cycling frequencies Ω/k_c produce different orbits in the phase diagram, as shown in Fig. 3.5d. Small Ω/k_c allows the system to remain close to chemical equilibrium, while large values do not allow the system to respond to the change in

the reservoir. In between, there exists a frequency where the system has time to react to the reservoir and be driven away from the chemical equilibrium line, maximizing $\mathcal{R}$. For larger values of k_e/k_c, the ability of the system to respond to the reservoir increases, making $\mathcal{R}$ larger as well. With increasing k_e/k_c, the system is able to follow the reservoir equilibrium, saturating $\mathcal{R}$. Without interactions ($\chi_{ij} = 0$), the chemical equilibrium ($\mu_A = \mu_B$) is set by $\phi_B/\phi_A =$ constant (see Eq. (2.47)). As the equilibrium condition is identical to the evaporation/condensation constraint, reservoir oscillations cannot drive the system away from chemical equilibrium. Thus, again, changes in the solvent composition can only affect chemical reactions when non-dilute interactions among reacting components are considered.

The resonance behavior provides a selection mechanism for specific chemical reactions. This selection is achieved by the speedup of specific chemical reactions, while other reactions do not benefit from the wet-dry cycles. The selection conditions depend on the value of the evaporation flux coefficient k_e, which, however, does not change much for water for typical temperature and pressure oscillation on Earth [150, 151, 152]. A specific resonance frequency, and thus the selection of specific chemical reactions, can, however, be realized by varying the system's geometry, i.e., surface-to-volume ratio A/V_0 of the system (Fig. 3.6). Increasing the surface area to volume ratio from a puddle to a thin film increases the resonance frequency.

3.7 Comparison to reactant cycles

Until now, we have considered the case of wet-dry cycles, where only the solvent can be exchanged with an external reservoir. The investigation can be generalized to environments where the reactants can be exchanged with the environment instead. For the ternary case studied here, this can be done by allowing the A-component to exchange with the reservoir $k_{A,e} \neq 0$ and $k_{C,e} = 0$, as defined in Eq. (3.1). In Ref. [6], we investigated this case and compared it to the behavior of wet-dry cycles. In this section, we give a short summary of the main results presented in the reference.

In addition to generalizing to other types of reservoirs, we can refine the notion of chemical activity per period $\mathcal{R}$, as defined in Eq. (3.17). Instead of quantifying solely the number of reaction events per volume, we can quantify the dissipated energy over a period. Such a definition quantifies the amount of energy harvested from the reservoir, which is given by

$$\mathcal{W} = -\lim_{t_0\to\infty} \frac{1}{\nu} \int_{t_0}^{t_0+2\pi/\Omega} \int_{V(t)} \mathrm{d}^3\boldsymbol{x}\, \mathrm{d}t\, r\, \Delta\mu\,, \tag{3.18}$$

where $\Delta\mu$ is the thermodynamic force of the chemical reaction $\Delta\mu = \mu_A - \mu_B$, ν is the molecular volume $\nu = \nu_A = \nu_B$, and r is the corresponding reaction flux defined

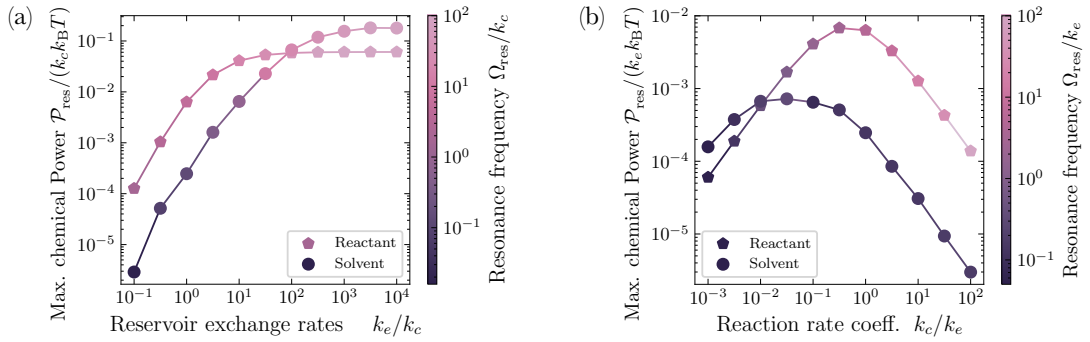


Figure 3.7: **Comparing solvent and reactant cycles at resonance:** (a) The maximal chemical power increases and saturates for both types of reservoir cycles as a function of the exchange rate with the reservoir. Interestingly, there is a crossover below which reactant cycles lead to more chemical power, while above, wet-dry cycles dominate. (b) As a function of the reaction rate coefficient, the maximum chemical power no longer saturates. Instead, there exists an optimal reaction rate that maximizes the power for both wet-dry and reactant cycles.

in Eq. (3.8). To find the dissipation per unit time, we define the chemical power of the cycle as

$$\mathcal{P} = \frac{\Omega}{2\pi}\mathcal{W}\,. \tag{3.19}$$

Interestingly, the resonance behavior of $\mathcal{R}$ for wet-dry cycles further holds for the chemical power $\mathcal{P}$ for both wet-dry and reactant cycles. In Fig. 3.7, the chemical power at resonance is compared between the two reservoirs. In (a), we see that for slow exchange with the reservoir, reactant cycles yield a larger power than wet-dry (solvent) cycles, while wet-dry cycles are more efficient for larger exchange rates. For wet-dry cycles, the resonance frequency is slightly lower than for reactant cycles. The same behavior is seen in Fig. 3.7(b), where wet-dry cycles yield a larger power at resonance for slow chemical reaction rates. Increasing the exchange rate with the reservoir always increases the chemical power (Fig. 3.7(a)), which is not the case as the reaction rate is increased. Instead, there exists a specific reaction rate coefficient where the chemical power at resonance is maxi-

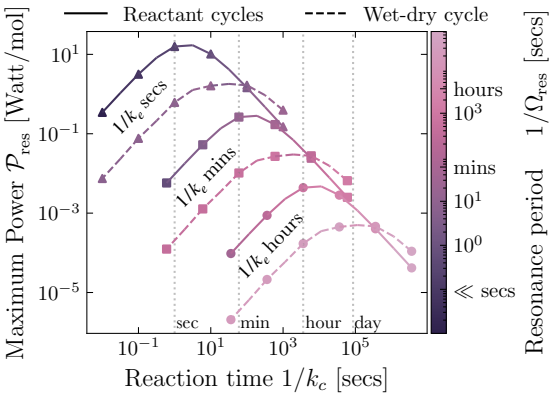


Figure 3.8: **Comparison between cycles with units:** There exists an optimal speed of chemical reaction that enhances the maximum chemical power $\mathcal{P}_{\rm res}$. Different reservoir exchange rates k_e produce different power outputs. For each k_e, there exists an optimal reaction time k_c that maximizes the chemical power. The vertical lines indicate reaction time scales of seconds, minutes, hours, and days.

mal. The optimal reaction rate coefficient for wet-dry cycles is around an order of magnitude lower than for reactant cycles and occurs at a lower resonance frequency. The discrepancy can be understood by noting that wet-dry cycles only implicitly affect the chemistry through interactions between components, and therefore need a slower cycling frequency to affect the turnover. Whereas reactant cycles directly affect the concentration of components undergoing chemical reactions.

In Fig. 3.8, wet-dry and reactant cycles are compared for different reaction-time scales. Despite reactant cycles yielding a larger power for most reaction time scales, wet-dry cycles become more efficient for the fastest reaction time scales. Intuitively, adding and removing components that directly undergo chemical reactions should be the most efficient, as it is a direct way of driving the chemical reactions. Nevertheless, interactions between components can, in some parameter regimes, make wet-dry cycles more suitable for harvesting energy from the reservoir to drive chemical reactions.

3.8 Discussion and conclusion

In this work, we develop a thermodynamic description for evaporation and condensation of non-dilute reacting mixtures and study their effect on chemical reactions. We find that the kinetics of evaporation and condensation can significantly speed up chemical reactions. This speedup stems from evaporation and condensation driving the system away from equilibrium. For wet-dry cycles, the turnover from substrate to product can be sped up even further. We found that the cycling frequency strongly affects the rate of chemical reactions per cycle. A key finding is the existence of a resonance cycling frequency where turnover is maximal. This resonance behavior can act as a selection mechanism to mainly speed up specific reactions.

The mechanism for how evaporation and condensation for constant and cycling reservoirs affect chemical reactions relies on effectively lowering the free energy barrier for a reaction to occur (Fig. 3.1a). By reducing the amount of solvent through evaporation, a solvophilic component will interact less with the solvent and more with the other components, making it less energetically favorable to be solvophilic. Likewise, solvophobic substrates interacting more with the solvent make it less energetically favorable to be solvophobic. In both cases, the substrate chemical potential μ_A increases, thereby reducing the free energy barrier ΔE. For realistic thermodynamic and kinetic parameters of chemical compounds in aqueous solutions [150, 152, 153, 154, 155], speedups of chemical reactions by more than a factor of 10 are possible through wet-dry cycles. Importantly, if assuming dilute chemistry ($\chi_{ij} = 0$), no speedup or resonance would appear; the effect of wet-dry cycles results from dense chemistry interactions. This pronounced effect on chemical reactions

suggests a strong contender as a naturally occurring process analogous to enzymes. However, while enzymes speed up reactions by lowering the kinetic contributions of the energetic barrier between the substrate and the product, evaporation/condensation can increase the chemical potential of the substrate μ_A and lower the chemical potential of the product. As a result, the speed of reactions can be enhanced (Fig. 3.1a).

Most chemical reactions in biology or chemical engineering involve enzymes or catalysts to enhance the speed of reactions [129, 156]. Subjecting such catalyzed reactions to wet-dry cycles could speed up the reactions even further. In particular, active sites of enzymes are typically solvent-depleted [157, 158, 159, 160] implying that enzymatic activity is enhanced in the dry state of the cycle. Pronounced enhancement is expected when changing the time-dependent reservoir chemical potential from symmetric wet-dry cycles (Eq. (3.16)) to strongly asymmetric cycles with long dry and short wet periods.

In this chapter, the change in chemical reactions stems from the non-dilute conditions and the interactions among reacting components. Interestingly, phase separation did not qualitatively alter the speedup or resonance phenomena observed. In Ref. [6], we show that phase separation can qualitatively alter the resonance phenomena, but only for wet-dry cycles. In fact, for certain interactions, phase separation results in two resonance frequencies. Phase separation would become more important for the behavior if reaction rate coefficients were different between the two phases. By including phase-dependent reaction coefficients or explicitly accounting for reactions on the interface between the two phases, we expect that phase separation can alter the resonance behavior.

In this chapter, we have focused on how chemical reactions are affected by wet-dry cycles in ternary mixtures. However, the theoretical framework provided can be applied to an arbitrary number of components with higher-order chemical reactions or reaction networks, as outlined in the theory section. For such complex reaction schemes, we expect that the speed up of chemical reactions gets even more pronounced as the reaction rates become proportional to a product of all the substrate volume fractions (Eq. (3.8)). This allows wet-dry cycles to affect chemical reaction rates even when assuming dilute chemistry [136]. Due to volume oscillations affecting all components equally, it is not able to bias the reaction towards specific reaction pathways. To select a specific reaction pathway, non-dilute reactions are necessary. Moreover, the resonance behavior should persist as it arises from a generic coupling between reservoir driving and reaction rates. Exploiting this resonance behavior in reacting mixtures with many components poses challenges at the interface between theory and experiments, which include the proper characterization of interaction parameters and kinetic rates.

Our findings suggest that wet-dry cycles could have played a vital role in speeding up prebiotic chemistry on the early Earth, where biological enzymes were absent. During early Earth settings, systems were likely exposed to repeated wet-dry cycles. The reservoir oscillations can originate from weather changes such as temperature, humidity, or pressure [161, 162], or other phenomena such as salt deliquescence [163], or freeze-thaw cycles [164, 165, 57]. Alternatively, oscillations in the solvent can be found in foams [124], or porous rock containing trapped gas-bubbles [166, 167].

An important implication of our findings is that the resonance behavior in the frequency of wet-dry cycles provides a selection mechanism for chemical reactions. Since the system's geometry determines the resonance frequency, different chemical reactions are selected in a puddle versus a mesoscopic aqueous droplet. This suggests that pores of different sizes subject to wet-dry cycles can provide a setting for chemical selection and evolution at the molecular origin of life.

Chapter 4

Excitability and self sustaining oscillations of active droplets

In the previous chapter, we studied how external oscillations can affect chemistry to select and speed up specific reaction pathways. In this chapter, we investigate the opposite limit: how can chemical reactions in the bulk create self-sustained internally driven oscillations? To arrive at an answer to this question, we first show how fueled chemical reactions in phase-separated systems create bistability. We show how this bistability can occur in minimalistic ternary systems and in polymer systems. By including an additional reaction, we show how the system becomes excitable. The excitable droplet behaves as an excitable medium, allowing for self-sustained oscillations. We show and discuss the properties of a single droplet and of an emulsion. This chapter contains results which have been published in Ref. [5], stemming from work done in collaboration with Hidde D Vuijk, Job Boekhoven, and Christoph A. Weber. For reproducibility, figure parameters are listed in Appendix M.2. For reading convenience, we repeat parts of the necessary theory from chapter 2, as we introduce the model.

4.1 Introduction

Oscillations of concentrations are relevant to numerous biological processes, including the circadian rhythm [168, 169] and the cell cycle [170, 171]. Other famous examples are synthetic life-inspired transcriptional oscillators [172, 173]. In chemistry, extensive experimental and theoretical work has revealed the underlying mechanism of oscillating chemical reactions, such as the Briggs–Rauscher reaction [174, 175], the Belousov–Zhabotinsky reaction [176, 177], and the Brusselator [178, 179] as paradigm examples. The oscillations emerge as a consequence of non-linearities in the chemical kinetic equations, giving rise to an unstable region in phase space

through a Hopf-bifurcation [180, 181]. Such non-linearities can further lead to excitable concentration fronts and pulses that propagate, as well as stationary Turing patterns [182, 183, 184].

An alternative mechanism for spatial patterns in chemically active systems is through phase separation of liquid condensates [80, 81]. The patterns in phase-separated systems do not require non-linearities in the chemical rates. Instead, attractive interactions between the components give rise to a phase boundary mediating non-linear transport between phases [87, 185]. When the chemical reactions break detailed balance of the rates, the system is maintained away from equilibrium, making the droplets active [85, 86, 87]. Active droplets can behave fundamentally different from their equilibrium counterpart, exhibiting properties such as droplet division [92], accelerated ripening [89], size control [91, 92, 93], shell-formation [96], and various other spatial morphologies and patterns [97, 98].

Oscillating chemical reactions rely on the precise spatial and temporal regulation of chemical processes, a property that can be provided by biomolecular condensates [91, 80, 81]. A key question is whether liquid condensates could not only regulate the chemical processes to enhance the stability of oscillating chemical reactions [186], but create the oscillations themselves. Cyclic reaction networks coupled to the liquid condensate could lead to oscillations in concentrations and even of the droplet itself. Such a mechanism is of biological relevance, as cycles of formation and dissolution of active biomolecular condensates have been shown to facilitate DNA repair [187, 188, 189].

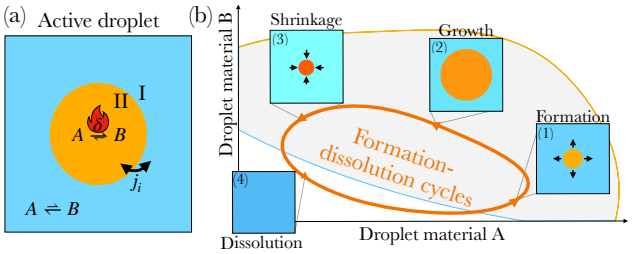


Figure 4.1: **Formation and dissolution cycles of active droplets:** (a) Chemically active droplets break detailed balance of the rates by the turnover of a fuel δ (Eq. (4.2)). This fuel drives the turnover of component A to B in the dense phase (II), while in the dilute phase (I), there is no fuel, and the pathway from B to A dominates. We find that for a large enough fuel strength δ, a saddle-node bifurcation occurs, making the system bistable. (b) Upon adding an additional chemical reaction (Eq. (4.28)), the active droplet becomes excitable and capable of self-sustained cycles, where they form (1) and grow (2), before shrinking (3) and dissolving (4), starting the cycle again.

In this chapter, we have presented a minimal theoretical model to achieve self-sustained oscillation of droplets; see Fig. 4.1. The oscillations do not emerge due to a complex non-linear reaction network but through the combination of minimalistic chemistry away from equilibrium and the thermodynamics of phase separation, leading to an active droplet. We show that a saddle-node bifurcation in the droplet volume emerges as active droplets are driven far away from equilibrium, making

them bistable. We demonstrate this bistability in a minimalistic ternary system and in a system of many interacting polymers. With the addition of a second reaction that breaks the conservation law of the first chemical reaction, the droplets behave as an excitable medium, exhibiting both wave propagation with a refractory period and emergent self-sustained oscillations.

4.2 Theory of active droplets

The evolution of volume fractions $\phi_i(\boldsymbol{x}, t)$ in space $\boldsymbol{x}$ and time t are governed by diffusive fluxes $\boldsymbol{j}_i$ and reaction rates r_i of component i with the corresponding continuum equations [88, 119, 87]:

$$\partial_t \phi_i(\boldsymbol{x}, t) = -\nabla \cdot \boldsymbol{j}_i + r_i \,, \qquad i = A, B. \tag{4.1}$$

We consider a ternary mixture (A, B, C) with $\phi_C = 1 - \phi_A - \phi_B$ that is initially limited to a reversible unimolecular chemical reaction between A and B,

$$r_1 : \quad \mathrm{A} \underset{k_1}{\overset{k_1 \exp\left\{\tilde{\delta}/(k_B T)\right\}}{\rightleftharpoons}} \mathrm{B} \,, \tag{4.2}$$

where the net reaction rates are $r_1 = r_B = -r_A$, $r_C = 0$, and k_1 is the reaction rate coefficient. The reaction rates are governed by the chemical potentials μ_i [148, 113],

$$r_1 = k_1 \left[\exp\left\{ \frac{\mu_A + \tilde{\delta}(\phi_A, \phi_B)}{k_B T} \right\} - \exp\left\{ \frac{\mu_B}{k_B T} \right\} \right] , \tag{4.3}$$

where $k_B T$ is the thermal energy. Moreover, $\tilde{\delta}(\phi_A, \phi_B)$ captures the effects of a composition-dependent fuel [190, 191] that enhances the forward reaction pathway, where $\exp\{\tilde{\delta}/(k_B T)\}$ can be understood as a chemostatted fuel activity equal to $\Delta\mu_{\alpha,\beta}$ in Eq. (2.53). A non-vanishing $\tilde{\delta}(\phi_A, \phi_B)$ breaks detailed balance of the rates [118, 87], $r_1^{\rightharpoonup}/r_1^{\leftharpoondown} = \exp\{(\Delta\mu + \tilde{\delta})/(k_B T)\}$, where $\Delta\mu = \mu_A - \mu_B$ is the chemical potential difference between the two states which only vanishes at chemical equilibrium ($\Delta\mu = 0$).

According to Eq. (4.1), the volume fraction fields ϕ_i further evolve due to chemical potential gradients in space, leading to diffusive fluxes. In linear response, diffusive exchange fluxes read $\boldsymbol{j}_i = -\Lambda_i \nabla \bar{\mu}_i$, where $\Lambda_i = \Lambda_{i,0} \phi_i \phi_C$ are the mobility coefficients [111] and $\bar{\mu}_i$ are the exchange chemical potentials $\bar{\mu}_i = \mu_i - \nu_i \mu_C / \nu_C$, and $\Lambda_{i,0}$ relates to the diffusion coefficient as $\Lambda_{i,0} = D_i k_B T$. For sufficiently strong interactions between the components, the system can phase-separate, creating stable inhomogeneous volume-fraction profiles.

The dynamic Eqs. (4.1) governing phase separation greatly simplify [113] when the diffusive exchange rate between phases is much faster than the reaction rate coefficient, $k_1 \ll D_i/L^2$, where L is the system size and $\sqrt{D_i/k_1}$ is the reaction-diffusion length scale. In this limit, each phase is homogeneous in composition with different volume fractions in each phase $(\phi_i^{\mathrm{I}}, \phi_i^{\mathrm{II}})$ with corresponding phase volumes V^{I} and $V^{\mathrm{II}} = V - V^{\mathrm{I}}$, where V is the system volume. Importantly, the chemical kinetics can be represented as a trajectory of average volume fractions in the thermodynamic phase diagram; see e.g. Fig. 4.1 for a cycling trajectory (orbit) in a phase diagram where A and B can phase separate from the solvent C. The phase compositions (I, II) are constrained to the binodal line and satisfy phase equilibrium at all times: $\mu_i^{\mathrm{I}} = \mu_i^{\mathrm{II}}$. These simplified dynamic equations in each phase read [113]:

$$\frac{\mathrm{d}\phi_i^{\mathrm{I/II}}}{\mathrm{d}t} = r_i^{\mathrm{I/II}} - j_i^{\mathrm{I/II}} - \phi_i^{\mathrm{I/II}} \frac{\dot{V}^{\mathrm{I/II}}}{V^{\mathrm{I/II}}}\,, \tag{4.4a}$$

$$\frac{\dot{V}^{\mathrm{I/II}}}{V^{\mathrm{I/II}}} = \sum_{k=A,B,C} \left(r_k^{\mathrm{I/II}} - j_k^{\mathrm{I/II}} \right)\,, \tag{4.4b}$$

where $i = A, B$, and the diffusive flux $j_i^{\mathrm{I/II}}$ ensures the conditions of phase coexistence, see section 2.3.2 for a derivation. Phase-averaged quantities are denoted by a bar, which, for an arbitrary quantity X at phase equilibrium, is

$$\bar{X} \equiv \frac{V^{\mathrm{I}} X^{\mathrm{I}} + V^{\mathrm{II}} X^{\mathrm{II}}}{V^{\mathrm{I}} + V^{\mathrm{II}}}\,. \tag{4.5}$$

For a ternary system with chemical reaction 1 (Eq. (4.2)), a natural choice of quantities is the phase-averaged total volume fraction $\bar{\psi}$, which is conserved during reaction 1, and the phase-averaged reaction extent $\bar{\xi}$,

$$\bar{\psi} \equiv \frac{\bar{\phi}_A + \bar{\phi}_B}{2}\,, \qquad\qquad \bar{\xi} \equiv \frac{\bar{\phi}_B - \bar{\phi}_A}{2}\,. \tag{4.6}$$

From Eqs. (4.4), we see that the evolution of these phase-averaged quantities obey

$$\mathrm{d}_t \bar{\psi} = 0\,, \qquad\qquad \mathrm{d}_t \bar{\xi} = \bar{r}_1\,. \tag{4.7}$$

Steady states, which corresponds to $\bar{r}_1 = 0$, defines the reaction nullcline of r_1:

$$V^{\mathrm{I}}\, r_1^{\mathrm{I}} + V^{\mathrm{II}}\, r_1^{\mathrm{II}} = 0\,. \tag{4.8}$$

For reactions obeying detailed balance of the rates ($\tilde{\delta} = 0$), the nullcline is the entire tie-line satisfying $\mu_A = \mu_B$, making it independent of the phase volumes (V^{I} and V^{II}). For chemical reactions that break the detailed balance of the rates, the

nullcline for phase-separated systems depends on the phase volumes. The system is kept away from equilibrium by the composition-dependent fueling

$$\tilde{\delta}(\psi) = \frac{\delta}{2}\left[1 + \tanh\left(\frac{\psi - \psi_\delta}{\kappa}\right)\right], \tag{4.9}$$

where $\delta \geq 0$ is the fuel strength that enhances turnover from A to B. For $\delta = 0$, the system can settle to thermodynamic equilibrium. The parameter $\kappa \ll 1$ sets the transition range of ψ values around ψ_δ, where the system transits from the non-fueled to the fueled case. For $\psi > \psi_\delta$, the fuel is activated. We choose $\psi_\delta = 1/4$ such that the dilute phase is approximately unfueled, while the dense phase is fueled ($\tilde{\delta}^{\mathrm{I}} \simeq 0$, $\tilde{\delta}^{\mathrm{II}} \simeq \delta$), resulting in a volume-dependent reaction nullcline when phase-separated. The phase-dependent fueling gives rise to non-equilibrium steady states (NESS) with continuous chemical turnover in both phases, balanced by diffusive fluxes across the phase boundary, see Fig. 4.1(a). For the fueling described by Eq. (4.9), component B is produced in the dense phase (large ψ) and diffuses to the dilute phase (small ψ), reacting to A before diffusing back to the droplet. Eq. (4.8) dictates a corresponding steady state droplet volume, finding

$$\frac{V^{\mathrm{II}}}{V} = \frac{k_1^{\mathrm{I}}\left[e^{\frac{\Delta\mu+\delta^{\mathrm{I}}}{k_B T}} - 1\right]}{k_1^{\mathrm{II}}\left[e^{\frac{\Delta\mu+\delta^{\mathrm{II}}}{k_B T}} - 1\right] - k_1^{\mathrm{I}}\left[e^{\frac{\Delta\mu+\delta^{\mathrm{I}}}{k_B T}} - 1\right]}, \tag{4.10}$$

where $\delta^{\mathrm{I/II}} = \tilde{\delta}(\psi^{\mathrm{I/II}})$ and $k_1^{\mathrm{I/II}}$ denote the phase-dependent reaction rate coefficients of reaction 1 (Eq. (4.2)). Where the unfueled case $\delta^{\mathrm{I}} = \delta^{\mathrm{II}} = 0$ gives a solution for all V^{II} with $\Delta\mu = 0$. Unless stated otherwise, we solve Eqs. (4.4) numerically.

4.3 Saddle-node bifurcation to a bistable droplet

4.3.1 Bistability of a ternary system

Without fuel ($\delta = 0$), the steady state of the system is homogeneous, corresponding to thermodynamic equilibrium. Upon increasing the fuel strength δ, active droplets emerge as non-equilibrium steady states that are locally stable. Their emergence requires a large enough fueling, bending the reaction nullcline strongly enough (Fig. 4.3(a,b)) and thereby creating multiple intersections with the same conserved line of constant total volume fraction $\bar{\psi}$. We note that the homogeneous steady state remains stable as δ is increased and a locally stable active droplet is created. Thus, the system becomes bistable with two locally stable steady states.

Bistability between an active droplet state and a homogeneous state emerges through a saddle-node bifurcation at which the steady droplet volume V^{II} jumps

from zero to a finite value; see Fig. 4.3(c). The bifurcation occurs at a critical fuel strength δ^*. An approximate expression for the critical fuel strength is given as (see Appendix F for the derivation)

$$\frac{\delta^*}{k_B T} \simeq \log\left(1 + \frac{k_1^{\mathrm{I}}}{k_1^{\mathrm{II}}}\Theta + \frac{\bar{\psi} - \psi^{\mathrm{I}}}{\psi^{\mathrm{II}} - \psi^{\mathrm{I}}}\,\Theta\left(1 - \frac{k_1^{\mathrm{I}}}{k_1^{\mathrm{II}}}\left(1 + \frac{\Theta}{2}\right)\right)\right). \tag{4.11}$$

where Θ depends solely on thermodynamic parameters (see Eq. (F.5)). Importantly, the minimal fuel strength δ^*_{min} for bistability to occur simplifies to

$$\delta^*_{\mathrm{min}} \simeq k_B T \log\left(1 + \frac{k_1^{\mathrm{I}}}{k_1^{\mathrm{II}}}\Theta\right), \tag{4.12}$$

We see that the bifurcation is favored when the reactions are faster in the dense phase compared to the dilute phase ($k_1^{\mathrm{II}} > k_1^{\mathrm{I}}$). Moreover, the prerequisite for the bifurcation ($\Theta > 0$) sets a restriction to certain phase diagram types. For example, the bifurcation can only occur for systems where the product B of the fueled reaction is less soluble than its precursor A, for which $\Theta > 0$. Furthermore, Θ never approaches zero, implying that a fuel is always necessary for the emergence of bistability.

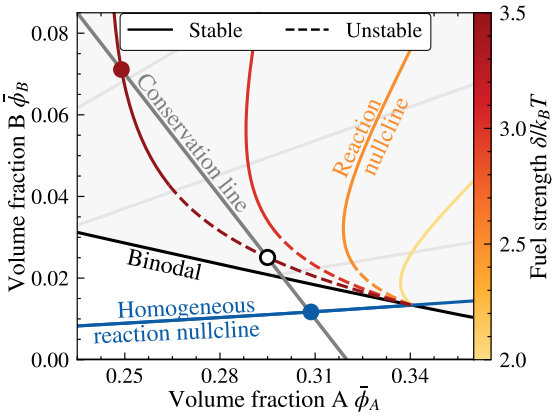


Figure 4.2: **Bistability of an active droplet:** Upon increasing the fuel strength δ, the reaction nullcline (Eq. (4.10)) for reaction r_1 (Eq. (4.2)) in the phase-separated domain (gray shaded) bends, leading to three non-equilibrium steady states (NESS) for a fixed conservation line $\bar{\psi} = (\bar{\phi}_A + \bar{\phi}_B)/2$. Closed symbols are stable and open symbols are unstable fixed points. The homogeneous reaction nullcline remains unchanged by changes in the fuel strength δ. The full phase diagram is displayed in Fig. G.1 in Appendix G.

We further find that bistability is only possible in a range of total volume fractions $\bar{\psi}$-values and this range increases with fuel strength δ (orange domain in Fig. 4.3(e)). The analytic expression for the critical fuel strength δ^* given in Eq. (4.12) agrees well with the complete numerical solution close to the active droplet domain in Fig. 4.3(d). As the total volume fraction $\bar{\psi}$ enters the binodal for a fixed fuel strength $\delta > \delta^*$, a jump occurs in both phase volume and average composition (Fig. 4.3(c)), unlike the case for $\delta < \delta^*$.

The onset of bistability can be understood through a normal form obtained from Taylor expanding the dynamic Eq. (4.7) for the reaction extent $\bar{\xi}$. By expanding the phase-separated and homogeneous case separately, at the points $(\bar{\xi}_0, \bar{\psi}_0)$ and $(\bar{\xi}_2, \bar{\psi}_0)$

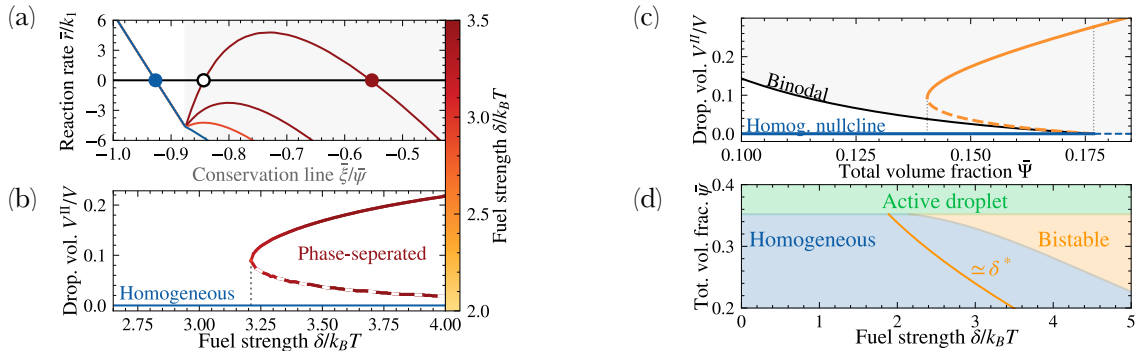


Figure 4.3: **Saddle-node bifurcation to an active droplet:** (a) The net reaction rate along the conservation line only has multiple steady states for sufficiently large fuel strength $\delta/(k_BT)$, where the center point is unstable. (b) The multiple steady points arise from a saddle-node bifurcation in droplet volume upon a critical fuel strength δ^* (vertical dotted line). Above this threshold value δ^*, there are three NESS' where two are locally stable (bistability). (c) At fixed fuel strength, The occurrence of the bifurcation depends on the value of the quantity $\bar{\psi}$ conserved by reaction (4.2). There is a range of $\bar{\psi}$-values (vertical dotted lines), where the system is bistable. (d) Bistability is only possible within a range of $\bar{\psi}$ values that grow with fuel strength δ. The analytical approximation for the critical fuel strength δ^* from Eq. (4.11) agrees well for fuel strength δ close to the onset of bistability.

respectively, we find (for derivation see Appendix D),

$$\mathrm{d}_t\bar{\xi} \simeq \begin{cases} -\left(\bar{\xi}-\bar{\xi}_0\right) h_1, & \bar{\xi} \le \xi^{\mathrm{I}}, \\ \bar{r}_1(\bar{\xi}_2,\bar{\psi}_0) - \left(\bar{\xi}-\bar{\xi}_2\right)^2 h_2 & \bar{\xi} > \xi^{\mathrm{I}}, \end{cases} \tag{4.13}$$

with the coefficients h_1 and h_2 given in Eq. (D.3) and (D.9) respectively, $\bar{r}_1$ is the phase-averaged reaction rate evaluated at the phase-separated expansion point, and ξ^{I} is the reaction extent at the dilute branch of the binodal for the given $\bar{\psi}_0$. As seen in Fig. 4.3(b), for homogeneous values of the reaction extent ($\bar{\xi} \le \xi^{\mathrm{I}}$), the reaction rate is linear and stable ($h_1 > 0$) around its steady point $\bar{\xi}_0$. For phase-separated values of the reaction extent ($\bar{\xi} > \xi^{\mathrm{I}}$), a second-order expansion is necessary to capture the behavior for strong fueling δ. Since the normal form in the phase-separated domain corresponds to a concave parabola ($h_2 > 0$), the requirement for bistability is a positive offset value at the expansion point:

$$r_1^{\rightharpoonup}\left(\overline{\exp\left\{\frac{\delta}{k_BT}\right\} - \exp\left\{-\frac{\Delta\mu}{k_BT}\right\}}\right)\Bigg|_{\bar{\xi}_2,\,\bar{\psi}_0} > 0\,. \tag{4.14}$$

This equation shows that the saddle-node bifurcation occurs when the thermodynamic force ($\Delta\mu$) is dominated by the non-equilibrium driving (δ). The bar denotes a phase average (Eq. (4.5)). We note that for $\bar{\psi}_0$ values with a homogeneous steady point, the chemical potential difference is negative when phase separated ($\Delta\mu < 0$).

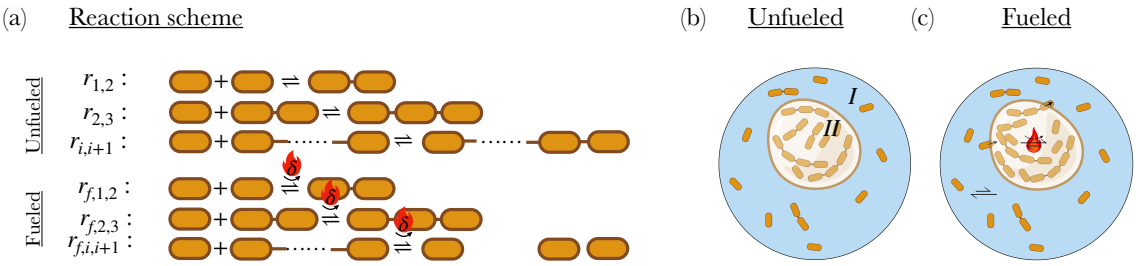


Figure 4.4: **System illustrations:** Polymers aggregate through reversible monomer pickup and release, a subset of the reaction network is depicted in (a) for both unfueled reaction (top, Eq. (4.17)), which reach thermodynamic equilibrium, and the fueled case (bottom, Eq. (4.20)). (b) As the polymers grow, they enhance their propensity to phase separate, which could lead to the coexistence between a dense phase which is oligomer rich (II) and a dilute phase which is oligomer poor (I). (c) For the case where the monomer-pickup is aided by a fuel δ which partitions to the dense phase ($\delta^{\mathrm{I}} \simeq 0,\ \delta^{\mathrm{II}} \simeq \delta$), the system reaches a non-equilibrium steady state with continuous fluxes of aggregation and exchange between the phases. This figure panel is made with biorender [192].

4.3.2 Bistability of active phase separating polymers

Bistability of droplets is a phenomenon not limited to simple ternary mixtures. It can also occur in condensed phases formed out of polymers. To model polymerization, we will follow the description that we developed in collaboration with Giacomo Bartolucci, Thomas C. T. Michael, and Christoph A. Weber in Ref. [2]. This framework describes the interactions and assembly properties of aggregating polymers. The polymers assemble through monomer pickup and release, and are described by the number of constituting units i. A polymer of length i, denoted P_i, can increasing its size by monomer pickup, as

$$P_i + P_1 \rightleftharpoons P_{i+1}\,, \tag{4.15}$$

this process is assumed to be reversible. An illustration of the reaction scheme is given in Fig. 4.4a. The net reaction of a sequence i is the difference between the reaction rate of production and loss of the polymer, such that

$$r_i = i\,(r_{i-1,i} - r_{i,i+1}) \quad i > 1, \qquad \text{and} \quad r_1 = -r_{1,2} - \sum_{i=1}^{M} r_{i,i+1}\,, \tag{4.16}$$

where M is the max length, we will investigate the case of $M \to \infty$. All volume fractions ϕ_i^α evolve in time as in the previous section, that is Eqs. (4.4), where ϕ_i^α is the volume fraction of polymer P_i in phase α. The chemical potentials govern the rate of polymerization according to

$$r_{i,i+1} = k\left[\exp\left\{\frac{i\bar{\mu}_i + \bar{\mu}_1}{k_B T}\right\} - \exp\left\{\frac{(i+1)\bar{\mu}_{i+1}}{k_B T}\right\}\right]\,, \tag{4.17}$$

where $\bar{\mu}_i$ is the exchange chemical potential per length, defined as $\bar{\mu}_i = \nu_1 \left(\partial f/\partial \phi_i\right)_{\phi_{j\neq i}}$, where ν_1 is the monomer molecular volume, and f is the Flory-Huggins free energy density

$$f = \frac{k_B T}{\nu_0}\left[\sum_{i=1}^{M}\left(\frac{\phi_i}{\rho_i}\log\left(\frac{\phi_i}{\rho_i}\right) + \frac{\omega_i \phi_i}{k_B T} + \chi_{i0}\phi_i\phi_0\right) + \phi_0 \log\phi_0 + \frac{\phi_0\omega_0}{k_B T}\right]. \tag{4.18}$$

Here, the 0 subscript denotes the solvent, and index i denotes a polymer of length i, with molecular volume relative to the solvent of $\rho_i = \nu_i/\nu_0$. We assume that the molecular volume scales linearly with length, such that $\nu_i = i\nu_0$. As shown in Ref. [2], the polymer-polymer interactions for some classes of systems are zero, such that $\chi_{ij} = 0$ for $i, j > 0$. Furthermore, the internal energies are defined to be per polymer length i, such that $\omega_i = \omega_\infty - \Delta\omega/i$, where ω_∞ is the internal energy per length of an infinitely long polymer, and $\Delta\omega$ is the bond energy. The solvent-interaction per length of polymer i will be chosen independent of length, such that $\chi_{i0} = \chi$, where χ is the solvent-monomer interaction strength. Due to the solvent-polymer interaction, the system can phase separate forming two phases I and II obeying the conditions of phase coexistence $\Pi^{\mathrm{I}} = \Pi^{\mathrm{II}}$ and $\bar{\mu}_i^{\mathrm{I}} = \bar{\mu}_i^{\mathrm{II}}$ for all polymer lengths i. Phase I is polymer-poor, and phase II is polymer-rich. Throughout the system, the total volume of polymers is conserved, i.e.

$$\bar{\phi}_{\mathrm{tot}} = \sum_{i=1}^{M} \bar{\phi}_i \tag{4.19}$$

is a constant. An illustration of the phase-separated system at thermodynamic equilibrium is given in Fig. 4.4b.

In our work in Ref. [2], the thermodynamic equilibrium state was found for reversible polymerization. Here, we extend this work by including a phase-dependent fuel with strength $\delta^{\mathrm{I/II}}$ that aids monomer pickup, such that the fueled reaction rate $r_{f,i,i+i}^{\mathrm{I/II}}$ follows

$$r_{f,i,i+1}^{\mathrm{I/II}} = k^{\mathrm{I/II}}\left[\exp\left\{\frac{i\bar{\mu}_i + \bar{\mu}_1 + \delta^{\mathrm{I/II}}}{k_B T}\right\} - \exp\left\{\frac{(i+1)\bar{\mu}_{i+1}}{k_B T}\right\}\right]. \tag{4.20}$$

We will investigate the case, as in the previous section, where the fuel partitions into the dense phase, such that $\delta^{\mathrm{I}} \simeq 0$ and $\delta^{\mathrm{II}} \simeq \delta$, as depicted in Fig. 4.4c.

The non-equilibrium steady state (NESS) is reached when the system average of all reaction events $\bar{r}_{f,i}$ is zero. In the limit of $M \to \infty$, this allows us to derive an expression for the NESS distribution of polymers

$$\phi_i^{\mathrm{NESS}} = i\left(\frac{\phi_1}{\phi^*}\right)^i \phi^*, \tag{4.21}$$

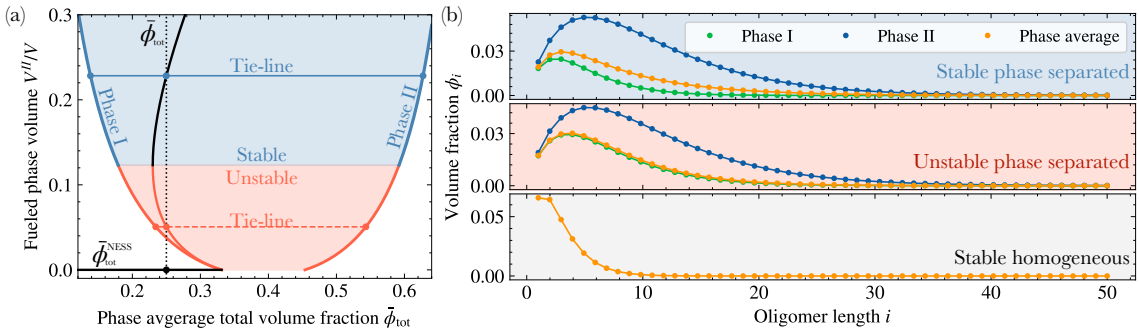


Figure 4.5: **Bistable phase diagram and length distributions:** (a) The phase diagram at the non-equilibrium steady state (NESS) is unstable for small phase volumes. For the given $\bar{\phi}_{\rm tot}$ (black dotted line), there are three intersections with the tie-line selecting curve (black solid line when stable and red solid line when unstable), each intersection is a steady state. Since two of the states are stable, the system is bistable. The corresponding length distributions are depicted in (b), where the blue shaded background is the stable tie-line in (a), while the red is the unstable homogeneous state, and the gray corresponds to the stable homogeneous state.

where ϕ^* is characteristic volume fraction

$$\phi^* = \frac{V \exp\left\{\frac{\Delta\omega}{k_B T} - 1\right\}}{V^{\rm I} \exp\left\{\frac{\delta^{\rm I}}{k_B T}\right\} + V^{\rm II} \exp\left\{\frac{\delta^{\rm II}}{k_B T}\right\}} \,. \tag{4.22}$$

Note that due to the conditions of equal chemical potentials between the phases, the effect of the fuel in either phase affects the NESS distribution in both phases equally. The monomer volume fraction ϕ_1 can be expressed in terms of $\phi_{\rm tot}$ in Eq. (4.19) by calculating the sum, finding

$$\frac{\phi_1^{\rm NESS}}{\phi^*} = \frac{1 + 2\frac{\phi_{\rm tot}}{\phi^*} - \sqrt{1 + 4\frac{\phi_{\rm tot}}{\phi^*}}}{\frac{2\phi_{\rm tot}}{\phi^*}} \,. \tag{4.23}$$

With the expression for $\phi_1^{\rm NESS}$, the entire polymer distribution $\phi_i^{\rm NESS}$ is known, allowing us to recast the free energy density in Eq. (4.18) in terms of the total volume fraction $\phi_{\rm tot}$ and the characteristic volume fraction ϕ_* at the non-equilibrium steady state, finding

$$f_{\rm NESS} = \frac{k_B T}{\nu_0} \left\{ \frac{\phi_* \alpha}{\rho_1(\alpha - 1)} \left[\frac{\log(\alpha)}{\alpha - 1} - \log\left(\frac{\phi^*}{\rho_1}\right) + \Delta\omega\rho_1 \right] + \frac{\phi_{\rm tot}\omega_\infty}{k_B T} + \chi\phi_{\rm tot}(1 - \phi_{\rm tot}) \right\} , \tag{4.24}$$

where $\alpha(\phi_{\rm tot}, \phi^*)$ is defined as $\alpha = \phi_1^{\rm NESS}/\phi^*$ from Eq. (4.23). Note that the free energy density $f_{\rm NESS}$ is not just a function of thermodynamic parameters, as the arguments of $f_{\rm NESS}(\phi_{\rm tot}, V^{\rm I}/V, \delta^{\rm I}, \delta^{\rm II})$ are not solely thermodynamic quantities. Despite the $M + 1$ components, the system can be described in one dimension by

utilizing the constraints of the non-equilibrium steady state ($\bar{r}_{f,i} = 0$).

For the case of an unfueled system ($\delta^{\mathrm{I}} = \delta^{\mathrm{II}} = 0$), the free energy density is similar to a binary system, and the Maxwell construction can be used to find the composition of two coexisting phases with total volume fraction $\phi_{\mathrm{tot}}^{\mathrm{I/II}}$, as we have done in Ref. [2]. Once the $\phi_{\mathrm{tot}}^{\mathrm{I/II}}$ is known, the entire polymer distribution can be determined from Eq. (4.21) and Eq. (4.23).

For the case of a fueled system, where either fuel is non-zero ($\delta^{\mathrm{I}} \neq 0$ and/or $\delta^{\mathrm{II}} \neq 0$) and importantly phase dependent ($\delta^{\mathrm{I}} \neq \delta^{\mathrm{II}}$), the phase volumes now enter the total free energy density in Eq. (4.24). One must therefore use a Maxwell construction to find the total volume fraction of the coexisting phases $\phi_{\mathrm{tot}}^{\mathrm{I/II,NESS}}$ for each phase volume, and self-consistently find the phase-averaged total volume fraction from

$$\bar{\phi}_{\mathrm{tot}} = \frac{V^{\mathrm{I}}\phi_{\mathrm{tot}}^{\mathrm{I}} + V^{\mathrm{II}}\phi_{\mathrm{tot}}^{\mathrm{II}}}{V^{\mathrm{I}} + V^{\mathrm{II}}}\,. \tag{4.25}$$

Following this procedure, the phase diagram in Fig. 4.5 is found. Unlike for an equilibrium phase diagram, both the phase volume and total composition must be specified. For most values of the phase-averaged total volume fraction $\bar{\phi}_{\mathrm{tot}}$, there exists a single phase volume V^{II} at the non-equilibrium steady state. But for a specific interval of $\bar{\phi}_{\mathrm{tot}}$, two different phase volumes are solutions to the non-equilibrium steady state condition of $\bar{r}_{f,i} = 0$, where the lower volume is unstable. By unstable, we mean perturbations away from the NESS results in a trajectory towards either of the stable NESS of the same $\bar{\phi}_{\mathrm{tot}}$, as in Fig. 4.3(a).

The bistability occurs because the fuel increases the propensity to phase separate. As the fuel aids monomer-pickup, the average polymer length increases with the phase-volume according to [2]

$$\langle i \rangle^{\mathrm{I/II}} \simeq \sqrt{\frac{\phi_{\mathrm{tot}}^{\mathrm{I/II}}\left(V^{\mathrm{I}}\exp\left\{\frac{\delta^{\mathrm{I}}}{k_BT}\right\} + V^{\mathrm{II}}\exp\left\{\frac{\delta^{\mathrm{II}}}{k_BT}\right\}\right)}{V\exp\left\{\frac{\Delta\omega}{k_BT} - 1\right\}}}\,. \tag{4.26}$$

With increasing length, the importance of mixing entropy decreases (see Eq. (4.18)), which enhances the propensity to phase separate. Thus, as the volume of the fueled phase increases, the propensity to phase separate increases, making the dense phase denser and the dilute phase more dilute. This results in the behavior seen in Fig. 4.5.

A very similar system was recently theoretically studied by Chen et al. [193]. In their investigation, they found bistability when increasing and decreasing the fuel strength, which aids monomer pickup. They exploited this bistability to allow for a history dependence, where the non-equilibrium steady state depends on whether the system was initialized as homogeneous or phase separated. In their work, the fuel strength is chosen to be proportional to ϕ_{tot}, but the mechanism of bistability

is likely the same.

Bistability can be achieved both in the minimalistic case of a ternary system and for the high-dimensional case of infinitely many polymers. For both cases, bistability occurs because the phase-dependent fuel enhances the propensity to phase separate. In the ternary case, the product of the fueled reaction repels the solvent more than the substrate ($\chi_{BC} > \chi_{AC}$), while for the polymers, the increase in length from fueling enhances the propensity to phase separate.

In the next section, we will see how an additional reaction will make the bistable system excitable. To understand excitability, we will return to the minimalistic ternary system. Excitability is achievable for the polymer system, but will not be discussed here.

4.4 An excitable active droplet

Excitable systems can lead to wave propagation upon a perturbation, followed by a refractory period in which the system is no longer excitable [182, 183, 184, 194]. Here, we show that the bistable system discussed in the previous paragraph becomes excitable upon adding a second reaction r_2 that allows the otherwise conserved quantity $\bar{\psi}$ to change in time,

$$r_2: \quad \mathrm{A} + \mathrm{B} \underset{}{\overset{k_2}{\rightleftharpoons}} 2\mathrm{C}\,, \tag{4.27}$$

where the reaction obeys detailed balance of the rates

$$r_2 = k_2 \left(\exp\left\{ \frac{2\mu_C}{k_B T} \right\} - \exp\left\{ \frac{\mu_A + \mu_B}{k_B T} \right\} \right) . \tag{4.28}$$

Thus, the net reaction rates r_i in Eq. (4.1) and Eqs. (4.4) is $r_A = r_2 - r_1$ and $r_B = r_2 + r_1$. The dynamics of the average reaction extent and the total volume fraction (Eq. (4.6)) become

$$\mathrm{d}_t \bar{\xi} = \bar{r}_1\,, \qquad \mathrm{d}_t \bar{\psi} = \bar{r}_2\,. \tag{4.29}$$

For a bistable droplet ($\delta > \delta^*$, Eq. (4.12)), we find that a compositional perturbation which destabilizes the droplet, resulting in an excited trajectory in the phase diagram (Fig. 4.6a). Such a perturbation leads to the dissolution of the droplet. Once dissolved, the reactions drive the average composition into the binodal, nucleating an active droplet, and returning the system to its initial non-equilibrium steady state. Along its excited trajectory, the system exhibits a refractory period when it is no longer excitable. The resulting evolution of the droplet volume in time is reminiscent of a wave pulse. Similar to the excitable droplet displayed in Fig. 4.6a,

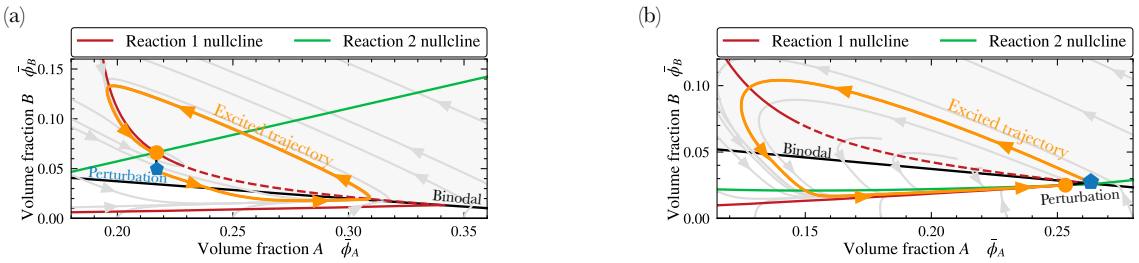


Figure 4.6: **Excitable of and to an active droplet:** When considering two chemical reactions, reaction r_1 (Eq. (4.2)) and reaction r_2 (Eq. (4.28)), the active droplet system becomes excitable. A compositional perturbation (blue pentagon) can excite the system. As a result, the system follows an extended trajectory (orange) with changes in average composition far beyond the initial perturbation. During this excited trajectory, the system has a refractory period when it is no longer excitable. It relaxes back to the NESS, where the reaction nullclines for both reactions intersect (green and red lines). The active droplet system can be considered as a bistable excitable medium. In (a), the perturbation drives the phase-separated NESS outside the binodal, while in (b) it drives a homogeneous NESS into the binodal. Both cases cause an excited trajectory, and the dilute branch of the binodal acts as an excitation threshold for both.

a homogeneous system with the two reactions r_1 and r_2 can become excitable upon a droplet-forming perturbation. To achieve this, the intersection of the two nullclines must occur outside the binodal. Thus, a droplet forming perturbation into the binodal results in an excited trajectory much longer than the perturbation, as seen in Fig. 4.6b. The dilute branch of the binodal corresponds to an excitation threshold. During the excited trajectory, the droplet volume and composition change until it dissolves and returns to its homogeneous steady state. The system is not excitable along the trajectory, thus exhibiting a refractory period. The full phase space for Fig. 4.6 is displayed in Fig. G.2 in Appendix G.

4.5 An oscillating active droplet

When the excitable fixed point becomes unstable, the excitable system undergoes cycles of formation and dissolution of the active droplet. Graphically speaking, destabilization corresponds to both nullclines of r_1 and r_2 intersecting inside the phase-separated domain of the phase diagram. As a result, a tiny droplet grows because the phase-dependent fueling favors the turnover from A to B, increasing its volume (Fig. 4.7(a)). These compositional changes cause the second reaction rate $\bar{r}_2$ to flip sign, thereby decreasing the total volume fraction $\bar{\psi}$. This reduction in $\bar{\psi}$ lowers the droplet volume until crossing the unstable branch of the saddle-node bifurcation (Fig. 4.3(c,d), dashed lines). Consequently, the droplet dissolves, and the effect of fueling approximately disappears. The lack of fuel reverses the direction of the net reaction rate $\bar{r}_1$, now converting B back to A, leading to a flipped sign

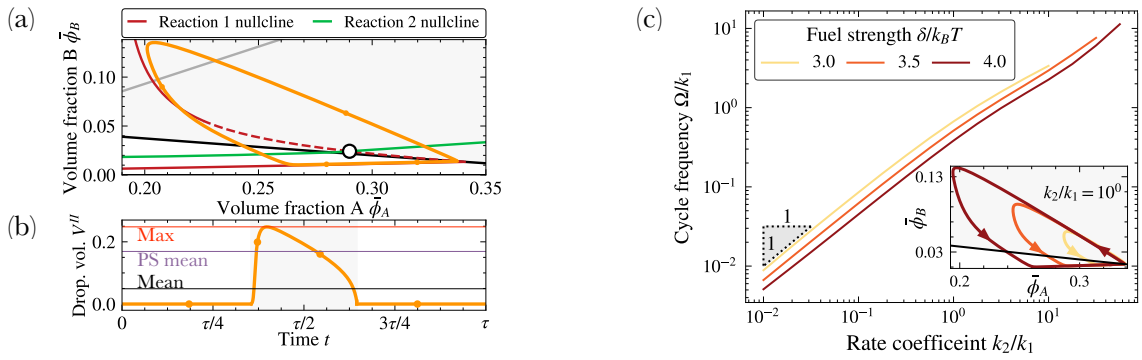


Figure 4.7: **Self-sustained cycles of droplet formation and dissolution:** (a) For cases where the intersection between the two reaction nullclines (green and red lines) occurs for unstable droplet volumes (white dot), it cannot concomitantly satisfy its two steady-state conditions. The result is self-sustained oscillations of an active droplet cycling between formation and dissolution. (b) The droplet continuously forms and dissolves in time. When nucleated at the same position, droplet formation and dissolution correspond to a standing wave. (c) The cycle frequency Ω is weakly affected by the fuel strength δ. However, the cyclic trajectories in phase space (inset) are significantly bigger, indicating that the compositional changes are much more pronounced. The cycle frequency follows a power-law with the reaction rate coefficient k_2, $\Omega \propto k_2$, due to the fuel making r_1 relax much faster than r_2 upon formation and dissolution.

of $\bar{r}_2$. A droplet nucleates, reaching its initial state. The cycle of formation and dissolution of the active droplet continues indefinitely.

The transitions between an excitable system (Fig. 4.6) and a system with oscillations of an active droplet (Fig. 4.7(a)) occur through a Hopf bifurcation [182, 180, 181, 194]. At the Hopf bifurcation, the eigenvalues of the linearization around the fixed point become imaginary. The transition to oscillations can, for example, be triggered by changing the reference chemical potential of component C, as defined in Eq. (2.16), or by extending the unstable domain of r_1 by increasing the fuel strength δ.

The Hopf bifurcation can be understood through the system's normal form (derived in Appendix E). It is obtained by performing a Taylor expansion on both sides of the binodal (see Appendix E for the derivation), yielding

$$\mathrm{d}_t\bar{\psi} \simeq -\left(\bar{\xi}-\bar{\xi}_1\right) f_\xi - \left(\bar{\psi}-\bar{\psi}_1\right) f_\psi \,, \tag{4.30a}$$

$$\mathrm{d}_t\bar{\xi} \simeq \begin{cases} -\left(\bar{\psi}-\bar{\psi}_0\right) h_\psi - \left(\bar{\xi}-\bar{\xi}_0\right) h_\xi \,, & \bar{\xi} \le \xi^{\mathrm{I}}, \\ \bar{r}_1 - \left(\bar{\psi}-\bar{\psi}_1\right) \tilde{h}_\psi - h_{\xi\xi}\left(\bar{\xi}-\bar{\xi}_2\right)^2 \,, & \bar{\xi} > \xi^{\mathrm{I}}. \end{cases} \tag{4.30b}$$

The key difference to the normal form with a single reaction r_1 (Eq. (4.13)) is that $\bar{\psi}$ is no longer a conserved quantity and changes during the kinetics. The consequences are coupling terms between the two dynamic quantities $\bar{\psi}(t)$ and $\bar{\xi}(t)$, characterized by the coefficients f_ξ, h_ψ and $\tilde{h}_\psi$ (definitions see Eqs. (E.3), (E.5), and (E.6)).

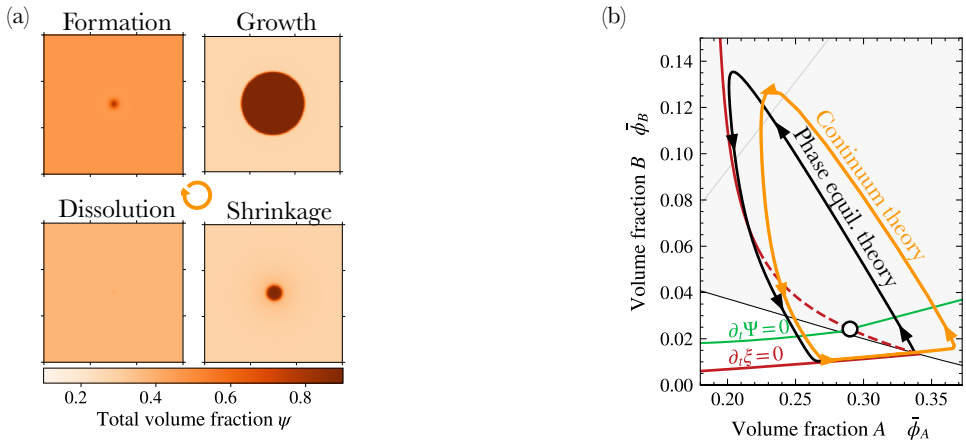


Figure 4.8: **Spatial profile of an oscillating droplet:** We confirm active droplets' formation and dissolution cycles by solving the continuum equation for the volume fraction fields (Eq. (4.1)). We find a good agreement of the phase space trajectory of the spatially averaged volume fractions for phase equilibrium theory (Eq. (4.4)), where the main discrepancy originates from the droplet's nucleation time. A movie of the spatial profile together with its trajectory in phase space can be found in Ref. [196].

We find that $f_\xi > 0$, $h_\psi < 0$, and $\tilde{h}_\psi > 0$, leads to negative feedback between $\bar{\psi}(t)$ and $\bar{\xi}(t)$ (see discussion in Appendix E). The signs of f_ξ and h_ψ are set by thermodynamics, while the difference in sign between $\tilde{h}_\psi$ and h_ψ is ensured from the fuel. Negative feedback loops, together with the instability at the intersection of the two nullclines, represent a well-established fundamental ingredient for the emergence of oscillations in non-linear systems [182, 195]. For cases where the two quantities can relax to a stable fixed point, i.e., $\bar{\psi}_0 = \bar{\psi}_1$ and $\bar{\xi}_0 = \bar{\xi}_1$ (homogeneous), or $\bar{\xi}_2 = \bar{\xi}_1$ (phase-separated), the system does not oscillate, however can remain excitable.

The oscillation frequency is dominantly determined by the reaction rate coefficient k_2 of the second reaction (4.28). This property is apparent in the scaling of the cycle frequency $\Omega \propto k_2$, as shown in Fig. 4.7(b). The dominant role of k_2 can be explained as follows: The first reaction (4.2) is fast since phase separation switches on the fueled forward pathway converting A to B, a prerequisite for having an active droplet. To complete the cycle, the slower unfueled reaction 2 with $k_2 \ll k_1 \exp\{\delta/(k_B T)\}$ becomes rate limiting, thereby determining the cycle frequency Ω. This separation of time-scales is consistent with the fuel strength δ having only a little effect on the cycle frequency Ω (different colors in Fig. 4.7(b)). Notably, the compositional changes along a cycle are more pronounced for larger fuel strength δ (Fig. 4.7(b), inset).

Consistently, we find formation-dissolution cycles when numerically solving the continuous dynamic equations for the volume fraction fields $\phi_i(\boldsymbol{x}, t)$ $(i = A, B)$ (Eq. (4.1)). Fig. 4.8(a) depicts representative snapshots of the volume fraction fields indicating the formation, growth, shrinkage, and dissolution of an active droplet

over the period $(2\pi/\Omega)$. Fig. 4.8(b) shows the corresponding phase space trajectory, which agrees well with the fast-diffusion limit. The self-sustaining oscillations are, therefore, well captured by the dynamic equations derived in the fast-diffusion limit which paved the way to determine the normal forms of the saddle-node bifurcation (Eq. (4.13)) of an active droplet and the Hopf bifurcation (Eq. (4.30)) to cycles of formation and dissolution of an active droplet.

In order to nucleate a single droplet, and not an emulsion, we include a nucleation potential $U(\boldsymbol{x})$ in our free energy G

$$G = \int_V \mathrm{d}^3\boldsymbol{x} \left[g(\phi_i(\boldsymbol{x})) + \frac{\kappa_A}{2} \left(\nabla\phi_A\right)^2 + \frac{\kappa_B}{2} \left(\nabla\phi_B\right)^2 + U(\boldsymbol{x})(\phi_A + \phi_B) \right] . \tag{4.31}$$

Here, g is the Gibbs-free energy density, i.e., the right-hand side of Eq. (3.2), κ_i is the free energy cost associated with gradients of component i, and U is the external potential affecting A and B. The chemical potential of component i for this free energy is calculated from the functional derivative of G with respect to volume fraction,

$$\tilde{\mu}_i = \nu_i \frac{\delta G}{\delta \phi_i} . \tag{4.32}$$

The inclusion of the external potential in the free energy yields an additional term in the chemical potential with respect to Eq. (2.16),

$$\tilde{\mu}_i = \mu_i + U(\boldsymbol{x}) \left[1 - \delta_{i,C}\right] . \tag{4.33}$$

The free energy cost of gradients through κ_i is related to surface tension, which leads to a droplet size-dependent Laplace pressure, thereby increasing the respective volume fraction in both domains. Since the binodal displayed in Fig. 4.8(b) is calculated without such surface tension effects, it does not exactly correspond to the binodal for the continuum equation. Still, it serves as a good approximation when the droplet's radius is much larger than the capillary length [87].

So far, we have considered the case of a system where both A and B independently phase-separate from the solvent. The same reaction network can produce fundamentally different oscillations by changing the interactions between the components. For example, it could make large droplet volumes unstable, though this is similar to the system we studied here.

An interesting case is a system where the reaction nullcline becomes unstable in the center of the binodal instead of by either the dilute or dense branch, as depicted in Fig. 4.9. Such a system does not result in formation and dissolution cycles but droplet volume oscillations. The oscillations emerge from an interval of droplet compositions being unstable. As the droplet grows and changes composition,

it becomes unstable, leading to a fast change in composition and a decrease in droplet volume. This fast composition change crosses the nullcline of r_2, leading to a turnover from C to A and B. Again, this turnover results in an unstable droplet composition where diffusive and reactive fluxes cannot balance, causing a rapid increase in droplet volume. The oscillations continue indefinitely.

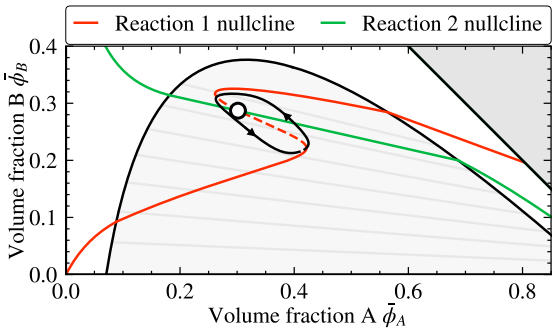


Figure 4.9: **Oscillations in droplet volumes:** The same chemical reaction network can produce oscillations in the droplet volume. These oscillations are achieved by changing the interaction strengths between the components, altering the binodal and, therefore, the curvature of the reaction nullcline.

The discussed oscillation type originates from phase diagrams where the dilute phase can become dense enough for the fuel to partition into both phases. This reduces the effect of the fuel, as it affects reactions similarly in both phases. In the limiting case of the fuel affecting the reaction equally in both phases, the system is mappable to an equilibrium system with a change in internal energy $\tilde{\omega}_A = \omega_A + \delta$. For the case in Fig. 4.9, the fueling threshold is $\psi_\delta = 0.225$, with a fuel transition width $\kappa = 0.04$.

4.6 Entropy production of an excitable droplet

The detailed-balance breaking reaction produces entropy. The entropy production originates from the fact that the fuel is chemostatted, implying an entropy-producing boundary flux to maintain the chemostatted fuel. The entropy production rate reads [197, 119]

$$-T\partial_t s = \partial_t g + j_F \,. \tag{4.34}$$

The flux of fuel into the system j_F balances the loss of fuel by the fueled reaction, $j_F = -\mu_F r_F/\nu_F$, where we now explicitly keep track of the chemostatted fuel with a constant chemical potential μ_F and molecular volume ν_F. In a non-equilibrium steady state $\partial_t g = 0$, and additionally integrates to zero over a closed cycle, such that all entropy production originates from the dissipative fueled reaction r_F. From the fueled reaction rate (derived in appendix C), we find that the boundary flux of the fuel yields an entropy density production

$$\frac{\dot{s}}{k_B} = \frac{\mu_F k_\mathrm{f}}{k_B T \nu_F} \exp\left\{\frac{\mu_A + \mu_F}{k_B T}\right\} . \tag{4.35}$$

Performing a volume integral to find the entropy production rate $\dot{S} = \int \dot{s}\, \mathrm{d}^3x$ yields

$$\frac{\dot{S}}{k_B} = \frac{\mu_F}{k_B T \nu_F} \exp\left\{\frac{\mu_A + \mu_F}{k_B T}\right\} \left(V^{\mathrm{I}} k_{\mathrm{f}}^{\mathrm{I}} + V^{\mathrm{II}} k_{\mathrm{f}}^{\mathrm{II}}\right). \tag{4.36}$$

At a non-equilibrium steady state, we can use the expression of the droplet volume V^{II} (Eq. (4.10)) and the relation between fuel strength $\tilde{\delta}$ and k_{f} in Eq. (C.5), where $\tilde{\delta}^{\mathrm{I}} = 0$ and $\tilde{\delta}^{\mathrm{II}} = \delta$. The entropy production at the non-equilibrium steady state simplifies to

$$\frac{\dot{S}_{\mathrm{NESS}}}{k_B} = \frac{\mu_F k_1 V}{k_B T \nu_F} \left(\exp\left\{\frac{\mu_B}{k_B T}\right\} - \exp\left\{\frac{\mu_A}{k_B T}\right\}\right). \tag{4.37}$$

Since the fuel is aiding the transition from A to B ($\mu_F > 0$), we have that $\mu_A < \mu_B$, ensuring that the entropy production is positive. For the case where the fuel strength goes to zero, the tie-line corresponding to chemical equilibrium, i.e., $\mu_A = \mu_B$, will be selected, making the entropy production rate zero. Interestingly, the entropy production is independent of the fuel strength and the phase volumes.

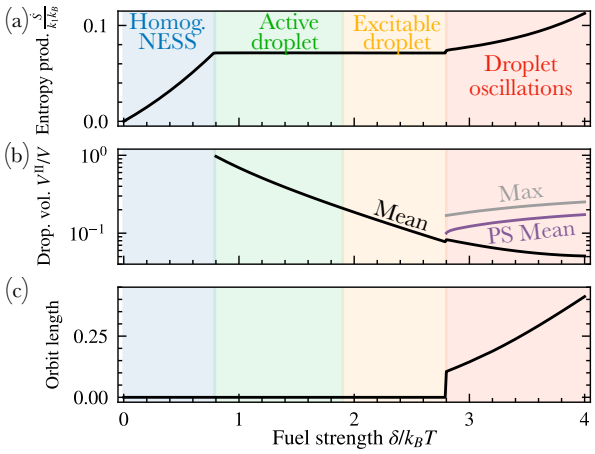


Figure 4.10: **Discontinuous entropy production at the onset of oscillations.** (a) Upon increasing the fuel strength δ, homogeneous NESS transits to the active droplet state with a stationary droplet volume with a kink in the entropy production $\dot{S}$. Increasing δ further decreases the droplet volume (b) but keeps the entropy production (a) constant. At the onset of formation-dissolution cycles, the entropy production $\dot{S}$ jumps (transition between the green and yellow domains in (a)). The volume averaged over a period jumps slightly, while the peak of the wave (Max) and its mean volume when phase separated (PS Mean) (see Fig. 4.7(b)) show a significant jump, as does the Orbit length shown in (c).

As the chemical reaction r_2 obeys detailed balance of the rates (Eq. (4.27)), the chemical potentials μ_A and μ_B are fixed by the tie-line selected by the nullcline of r_2 (green line in Fig. 4.6 and Fig. 4.7(a)). Thus, the entropy production for active droplets is independent of the fuel strength δ, and is instead a constant set by the properties of the non-dissipative chemical reaction 2. For a fixed $\Delta\mu = \mu_A - \mu_B$, we see from Eq. (4.10) that an increase in fuel strength δ decreases the droplet volume. This reduced volume exactly balances the increased dissipation from a stronger fuel, keeping the entropy production constant, as seen in Fig. 4.10 and Eq. (4.37). We note again that without a second reaction, which obeys detailed balance, the entropy productions will increase with the fuel strength δ [119].

Increasing the fuel strength further, the droplet transitions from an active droplet to an excitable droplet, where the steady state is stable. Upon further increasing the fuel strength δ, cycles of formation and dissolution emerge. The transition from stationary, excitable droplets to oscillations creates a discontinuous jump in the entropy production rate $\dot{S}$. Over a cycle, the average entropy production rate is calculated by integrating the entropy production (Eq. (4.36)) over a period,

$$\langle \dot{S} \rangle = \frac{\Omega}{2\pi} \int_0^{2\pi/\Omega} \dot{S} \, \mathrm{d}t \,, \tag{4.38}$$

which continues to grow with the fuel strength δ. As the entropy production rate jumps, so does the average droplet volume; see Fig. 4.10(b). An even larger jump occurs when comparing the average droplet volume during phase separation (PS mean) and the maximum droplet volume over a cycle (Max). The path of the non-equilibrium steady state is trivially zero, but discontinuously jumps upon the onset of cycles (Fig. 4.10(c)). The large jump upon oscillations is necessary, as the trajectory must cross the dilute branch of the binodal twice. Formation-dissolution cycles thus permit a discontinuous increase in entropy production and larger droplet volumes as the system is driven away from equilibrium.

4.7 Emulsions of excitable and oscillating droplets

So far, we have only investigated the excitability and oscillations of a single droplet. We will now study the same system by solving the continuum Eq. (4.1), which allows the system to enter the spinodal and create an emulsion of droplets. The results presented in this section are preliminary and subject to further analysis.

Droplets in an emulsion behave differently from a single droplet. In an emulsion, larger droplets absorb material from smaller droplets, a process known as ripening [112]. Furthermore, droplets can meet and coalesce into a single larger droplet. Additionally, the droplets can interact with one another through the shared dilute phase between them. The behavior of a single droplet away from equilibrium might be fundamentally different from its behavior in an emulsion.

In Fig. 4.11, we vary the reaction rate coefficient of the unfueld reaction (k_2 in Eq. (4.28)) and the fuel strength δ (Eq. (4.9)). For the lowest fuel values, the system is monostable and cannot undergo oscillations. Increasing the fuel strength beyond the critical fuel strength δ^*_{min} does not result in oscillations. Instead, the non-equilibrium steady state is an emulsion of droplets with similar radii. For an emulsion without a fueled reaction, the droplets would ripen and coarsen until a single droplet remains. Here, the ripening is arrested by counteracting non-equilibrium fluxes, which stabilize the droplets at a finite mean size and prevent complete coalescence.

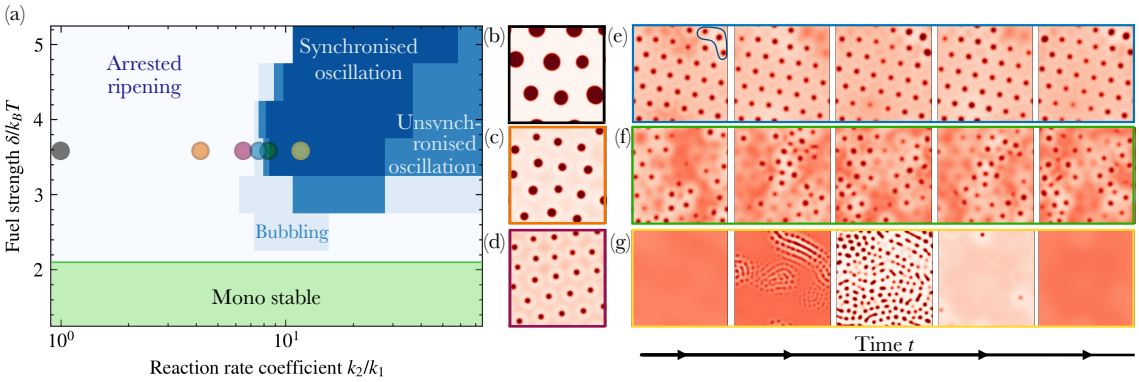


Figure 4.11: **Phase diagram for an emulsion of excitable droplets:** At low fuel strengths, the system is monostable and unable to oscillate. Above the critical fuel δ^*_{min}, the system reaches a non-equilibrium steady state of droplets with finite radii, where arrested ripening arises from counteracting reaction fluxes. For fixed δ, increasing k_2 reduces the mean droplet size (panels b–d). At intermediate k_2, a subset of droplets undergo persistent formation–dissolution cycles ("bubbling," panel e). Further increase of k_2 drives a transition from asynchronous oscillations (panel f) to system-wide synchronized formation dissolution cycles (panel g), before synchronization is lost again at higher k_2. The length of the box is two reaction-diffusion length scales $L = 2\sqrt{D/k_1}$. The corresponding movies can be found in Ref. [196].

For a fixed fuel strength δ, the average size of the droplets decreases as the reaction rate coefficient k_2 is increased (Fig. 4.11(b, c, d)). Kinetic rates determine the droplet properties at the steady state, rather than thermodynamic parameters.

Increasing the rate coefficient slightly further, a small subset of the arrested droplets start formation-dissolution cycles (Fig. 4.11e). This bubbling persists indefinitely in our simulations. Increasing the rate coefficients further, the arrested ripening halts and the formation dissolution cycles emerge, as seen in Fig. 4.11f. The oscillations are not synchronized. Some regions are in a phase above the nucleation threshold, while other regions are below. At a given time, some regions have droplets while others do not. Over a period, there is no point where the entire system is homogeneous. Enhancing the rate coefficient further creates a state where the full system undergoes synchronous cycles of formation and dissolution, as seen in Fig. 4.11g. The whole system nucleates droplets approximately simultaneously, where the droplets grow and eventually dissolve in phase, making the system homogeneous again. For even faster rate coefficients k_2, the oscillations are no longer synchronized, as displayed in Fig. 4.11f.

The arrested ripening regime in Fig. 4.11 must occur due to interactions between the droplets through the shared dilute phase between them. To understand this arresting of Oswald ripening, we depict the system-averaged volume fraction in the phase diagram in Fig. 4.12(a). Due to non-negligible interface gradients, the heavily coarse-grained phase-averaged composition only grants us a partial understanding of the system behavior.

For low k_2, the system averaged volume fractions $\bar{\phi}_i$ rest close to the nullcline of r_1 for their respective fuel strength δ. The non-equilibrium steady state is far away from the nullcline of the second reaction r_2, and this compositional distance increases with the fuel strength. Due to the fuel, the system has steady gradients of chemical potential in space, particularly across the droplet interfaces. These inhomogeneous chemical potentials result in a state where r_2 has different signs in the dense and dilute phases. Droplet material is produced in the dilute phase and degraded in the dense phase. The droplets, therefore, compete for the droplet material produced in the dilute phase, arresting their growth. As the reaction rate coefficient k_2 increases, the difference in production between the phases increases, and the phase-averaged volume fractions move towards the nullcline of r_2. For sufficiently large k_2, the system-averaged volume fractions reach the unstable branch of the nullcline of r_1. For sufficiently large k_2, there is not sufficient droplet material being produced, making some droplets dissolve and form again, as seen in Fig 4.11. For even larger k_2, all droplets oscillate.

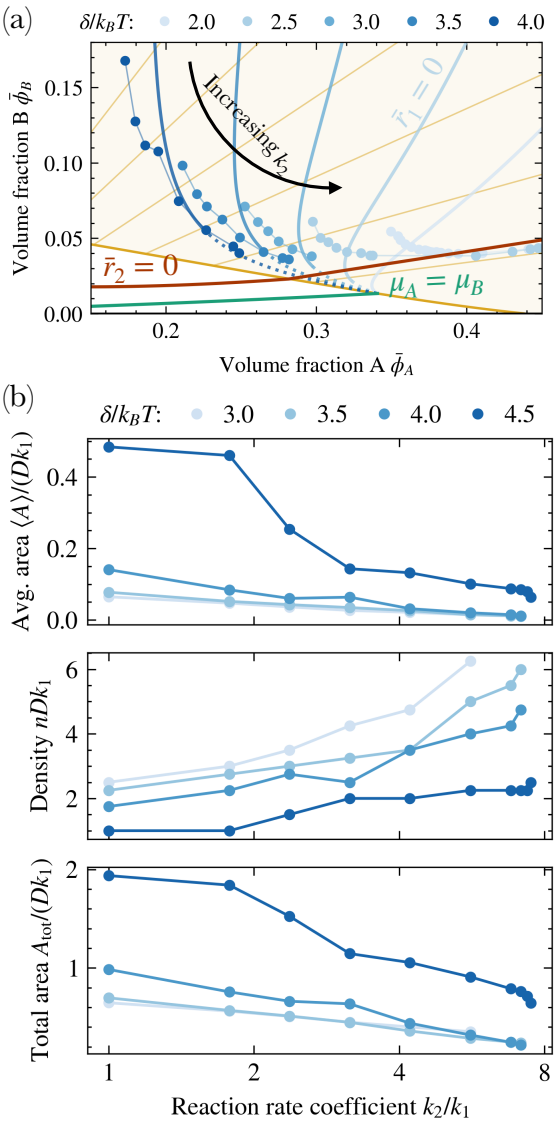


Figure 4.12: **Droplet behavior in regime of arrested ripening:** (a) System-averaged volume fractions $\bar{\phi}_i$ for varying reaction rate coefficients k_2 and fuel strengths δ, shown relative to the nullclines of the two reactions r_1 and r_2. Increasing k_2 shifts the averages toward the r_2 nullcline, eventually reaching the unstable branch of r_1's nullcline, where they start oscillating. (b) The average droplet area and total droplet area increase with k_2, while the number density decreases.

With the increase in k_2, both the average droplet area and total droplet area decrease, while the number density of droplets increases, see Fig. 4.12b. With increasing k_2, the competition for droplet material in the dilute phase increases, resulting in a higher density and smaller droplets in the steady state. The total condensed area decreases as the average area of the droplets decreases, despite the

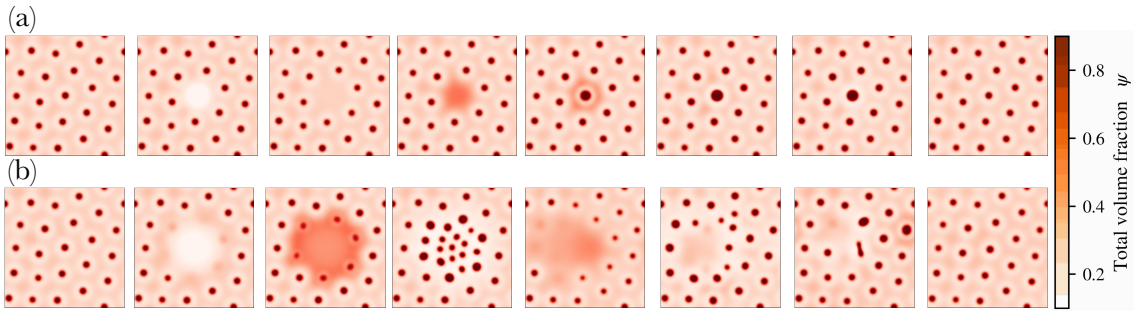


Figure 4.13: **Local perturbations in the arrested ripening regime:** A time-dependent flux (Eq. (4.39)) to remove A and B locally is added to the reaction rate in Eq. (4.1), which peaks at $tk_1 = 1/2$ before decaying. For both small ($\lambda = 0.15\sqrt{D/k_1}$, (a)) and large ($\lambda = 0.35\sqrt{D/k_1}$, (b)) perturbations, the system returns to the same arrested state. While the small perturbation requires only one nucleation event, the larger perturbation drives repeated nucleation–dissolution cycles before relaxation. The system size is two reaction diffusion length scales $L = \sqrt{D/k_1}$, and time evolves in the right direction for (a) and (b). The corresponding movies can be found in Ref. [196].

number density increasing. The increase in number density forces the droplets further together, resulting in a stronger competition for the droplet material produced in the surrounding dilute phase.

The stability of the regime of arrested ripening is tested by a local perturbation. The local perturbation converts A and B to C by a time-dependent externally controlled flux on the form

$$p_i = k_p \, \phi_i \, \exp\{\boldsymbol{x}^2/\lambda^2\} \, t \, \exp\{-2tk_1\} \, , \tag{4.39}$$

which enters (4.1) as $r_i \to r_i - p_i$. Here, k_p sets the rate, while λ determines the length scale of the perturbation. The time-scale is fixed, reaching a maximal rate at $tk_1 = 1/2$, before decaying to zero. In Fig. 4.13, the arrested ripening regime returns to its initial value for both a small ($\lambda = 0.15\sqrt{D/k_1}$ in (a)) and a large ($\lambda = 0.35\sqrt{D/k_1}$ in (b)) perturbation sizes. The final state for both runs has the same number density and droplet size as the initial condition. The two runs differ only by their trajectory back to their arrested state. In (b), the perturbed regions oscillate, nucleating and dissolving the droplets three times before reaching the non-equilibrium steady state. In (a), only one nucleation event is necessary.

In Fig. 4.14, the oscillating regime of Fig. 4.11 is analyzed. The frequency of oscillations increases approximately proportionally to the rate coefficient k_2. Different oscillation types, as depicted in Fig. 4.11a, are denoted by different markers. From bubbling (triangles), the frequency increases upon the onset of unsynchronized oscillations (crosses), and continues to grow as k_2 increases further, leading to synchronized oscillations (circles). The oscillation frequency is compared to the frequency of a single oscillating droplet, as studied in section 4.5, with the same fuel

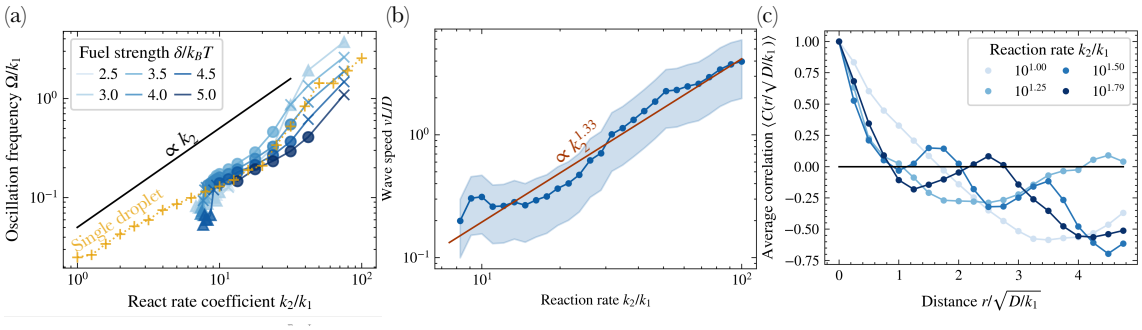


Figure 4.14: **Oscillation frequency, wave speed, and spatial auto-correlations:** (a) The oscillation frequency grows roughly proportionally with k_2, with different oscillation types indicated by markers (triangles: bubbling, crosses: unsynchronized, circles: synchronized). Frequencies match those of a single droplet at $\delta = 4k_BT$, but in emulsions a large k_2 is required, showing the role of interactions between the droplets. (b) The wave speed, defined as the time for nucleation across the system length L, scales slightly faster than k_2 and can exceed the diffusive speeds, enabling rapid signal transfer. (c) The spatial autocorrelation of $\psi = \phi_A + \phi_B$, defined in (4.40), shows oscillatory structure with decreasing wavelength as k_2 increases. The corresponding movies can be found in Ref. [196].

strength $\delta = 4k_BT$. Strikingly, the frequency almost exactly matches the oscillation frequency of many droplets for the same fuel strength. The main difference between the two is that for a single droplet, there is no critical reaction rate coefficient k_2 necessary for oscillations. This behavior further indicates how the arrested ripening regime is an effect of the interactions between droplets in the emulsion, which hinders oscillations.

In Fig. 4.14b, we have defined the wave speed as the average time it takes for droplet nucleation to occur over the system length L. Interestingly, the speed scales slightly faster than the reaction rate coefficient k_2. At a specific reaction rate coefficient k_2, the wave speed becomes faster than the diffusion speed over the same distance D/L. Exceeding diffusive speeds provides a mechanism for fast signal transfer.

In Fig. 4.14c, the time-averaged auto-correlation of ψ,

$$C(r) = \frac{\int_{T_0}^{T} \mathrm{d}t \int_0^L \mathrm{d}x \int_0^H \mathrm{d}y \, (\psi(x,y,t) - \langle\psi\rangle)\,(\psi(x+r,y,t) - \langle\psi\rangle)}{\int_{T_0}^{T} \mathrm{d}t \int_0^L \mathrm{d}x \int_0^H \mathrm{d}y \, (\psi(x,y,t) - \langle\psi\rangle)^2} \,. \tag{4.40}$$

is displayed as a function of horizontal distance r. The time average is performed over the time interval $[T_0, T]$, which is much larger than most of the reactive time scales of the system T, T_0, $T - T_0 \gg k_1^{-1}$, k_2^{-1}, and larger than the diffusive timescale of the system L^2/D. The time-dependent system average total volume fraction $\langle\psi\rangle(t)$ is calculated from

$$\langle\psi\rangle = \frac{1}{HL} \int_0^L \mathrm{d}x \int_0^H \mathrm{d}y \, \psi(x,y,t) \,, \tag{4.41}$$

where $\psi = \phi_A + \phi_B$. Note that the system size $(L \times H)$ is $(10 \times 0.5)\sqrt{D/k_1}$, where the short horizontal distance H is averaged over, such that only correlations along the long axis of the system are displayed. For different reaction rate coefficients k_2, different local maxima and minima appear as a function of the distance r. For $k_2/k_1 = 10$, a single local minimum exists. As k_2 increases, a local maximum appears as well. The wavelength of the spatial oscillations of the correlation function decreases as the rate coefficient increases, except for the largest rate coefficients.

We continue by initializing the high-aspect ratio system of size $(L \times H)$, with a sinusoidal profile for $\psi = \phi_A + \phi_B$ along the long axis. With an order parameter of synchronization

$$O = \left| \frac{1}{N} \sum_{k=0}^{N-1} e^{i\varphi_k} \right|, \qquad (4.42)$$

the initial profile is asynchronous $(O \simeq 0)$. In the definition of the order parameter, φ_k is the local phase calculated from the average total volume fraction $\psi = \phi_A + \phi_B$ over an interval along the long-axis

$$\Psi_k(t) = \frac{1}{H\Delta} \int_{k\Delta}^{(k+1)\Delta} \mathrm{d}x \int_0^H \mathrm{d}y\, \psi\,, \quad (4.43)$$

where $\Delta = L/N$ and $k = 0, \ldots, N-1$. From the $\Psi_k(t)$, we calculate the instantaneous phase φ_k from a Hilbert-transform [198].

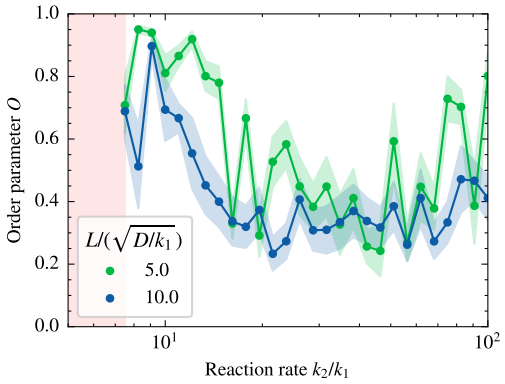


Figure 4.15: **Synchronization of oscillations:** The synchronization order parameter O, defined in (4.42), evolves from 0 to a finite value over time for two different system sizes L. For slow rate coefficients k_2, both systems are closed to completely synchronization $(O \approx 1)$, after which O decreases with increasing k_2 and saturates around 0.3. For the smaller system size, the synchronisation varies more, and exhibits large variations for small changes in the rate coefficient, though the general trend is a decrease with k_2. Even at large k_2, the system remains partially synchronized, with the reduction in O originating from the limited diffusive coupling across the long axis compared to the oscillation period. The regime without oscillations is colored red. Representative movies can be found in Ref. [196].

With an asynchronous initial profile, the long-time average order parameter is displayed in Fig. 4.15 as the rate coefficient k_2 is varied. The synchronization is displayed for two different system lengths $L = 10\ell$ and $L = 5\ell$, both with $H = 0.5\ell$, where ℓ is the reaction-diffusion length scale $\ell = \sqrt{D/k_1}$. For small rate coefficients, the order parameter is maximal. With increasing rate coefficient, the order parameter decreases until stabilizing at a value around 0.3. The smaller system size has a larger value of the order parameter than the longer system, but also exhibits strong variations for small changes in the rate coefficient k_2. Even for large k_2, resulting in high

frequency oscillations, the system is never completely asynchronous. The reduction in the order parameter with k_2 likely originates from the diffusive distance over a period being much smaller than the system size. Note that the initial profile was asynchronous, meaning the system can self-synchronize through diffusion, where a longer oscillation period aids in synchronization.

4.8 Discussion and conclusion

Cyclic reaction networks in homogeneous systems, such as the Brusselator [178, 179], rely on a complex non-linear reaction network of many components with a large variety of time scales [199]. The oscillations studied here occur with only two independent components. Our phase-separated system can instead use the phase boundary to create a simple cyclic network by effectively increasing the number of degrees of freedom through the phase-dependent concentrations. A cyclic reaction network is possible with a single detailed-balance-breaking reaction, as $A^{\mathrm{I}} \to A^{\mathrm{II}} \to B^{\mathrm{II}} \to B^{\mathrm{I}} \to A^{\mathrm{I}}$, where diffusive exchange between the phases I and II acts as effective reaction steps.

Bistability in homogeneous chemical systems relies on higher-order reaction networks to achieve a cubic normal form [184]. Our phase-separated system gives rise to the effective properties of a cubic form through two phases with a piece-wise combination of a linear and quadratic function (Eq. (4.13)). In other words, phase separation enables three stable points despite only two independent chemical states. The bistability found in Fig. 4.3 is, therefore, not a consequence of the complexity of the reaction network as such, it is a property of active droplets. Specifically, bistability arises from the coupling between phase separation and detailed-balance-breaking chemical reactions far from equilibrium, which we have shown can give rise to excitability and self-sustaining chemical oscillations.

The relevance of thermodynamic interactions among components for the bistability enables the same chemical reaction network to give fundamentally different oscillations. Changing thermodynamic parameters leads to different binodals, strongly altering the composition changes during a cycle. We find, for example, self-sustained oscillations that form and dissolve the droplet repeatedly, and cases where only the active droplet's volume oscillates without ever dissolving.

In this chapter, we have not investigated the role of noise. For the bistable system in section 4.3, sufficiently strong noise could allow the system to transition between the two stable steady states. For this case, one steady state would be homogeneous, while the other is phase separated. Alternatively, for the case in Fig. 4.9, noise could allow transitions between droplets of different volumes. Furthermore, the inclusion of noise in our equations would allow nucleation of droplets outside of the spinodal, resulting in a different trajectory and oscillation frequency. The role of stochasticity

will be addressed in future work.

In our work, we present a minimal theoretical model to achieve bistability, excitability, and oscillations in phase-separated systems with chemical reaction maintained away from equilibrium. We show that this non-equilibrium driving enables a saddle-node bifurcation in droplet volumes, resulting in two locally stable chemical states (bistability). By including an additional reaction, the system becomes excitable. For cases where the non-equilibrium steady state is on either of the stable branches, perturbations trigger an excited trajectory in composition, during which the system exhibits a refractory period. Upon enhancing the non-equilibrium driving, a Hopf bifurcation occurs to a self-sustained oscillating droplet undergoing indefinite formation and dissolution cycles. The behavior of an emulsion of oscillating droplets is qualitatively different from the single droplet case. Most importantly, the emulsion only oscillates for a sufficiently large rate coefficient k_2, below which the system reaches a stable non-equilibrium steady state of droplets with similar radii. In the oscillating regime, the system synchronizes itself over time, with a local maximum in the order parameter as a function of the reaction rate k_2.

The results presented for the emulsion of droplets are preliminary and represent ongoing work that remains subject to further analysis. In particular, the role of system size in the synchronization behavior observed in Fig. 4.15 should be examined more carefully. The relationship between the order parameter and the rate coefficient k_2 is non-monotonic and varies strongly even for small changes in k_2, especially for shorter systems. Whether synchronization persists for larger system sizes remains to be determined.

Excitable and oscillating active droplets can be realized experimentally in multiple ways. The general requirement is a phase-separating system with a phase-dependent fueled reaction pathway and a second chemical reaction that alters the quantity conserved by the fueled reaction. Moreover, the product of the fueled pathway should phase-separate more strongly than the precursor reactant. A potential experimental system could rely on transient post-translational modifications of proteins, where phase separation is activated through arginine methylation or phosphorylation [200]. Alternatively, various synthetic systems were recently proposed that use chemical fuels to transiently activate thermodynamically unstable products for active droplet formation [201], either through phosphorylation of peptides [202] or fuel-driven condensing reactions [96, 203].

The emergence of oscillations through non-equilibrium phase separation makes them an ideal functional module for *de novo* life systems [3]. Robust self-sustained chemical oscillations are important for *de novo* life to facilitate clock-like functions similar to extant life, such as the circadian rhythm and cell cycle [168, 169, 170], and could thereby provide other biological functions such as DNA repair [187, 188, 189].

Chapter 5

Selecting sequence patterns through phase separation

In the two previous chapters, we have seen how interactions between components can lead to speed-up of a selected reaction pathway (chapter 3) and create self-sustaining chemical oscillations (chapter 4) when maintained away from equilibrium. To understand the mechanisms, these functions were studied for the minimalistic system of three components. In order to describe systems of information-storing oligomers, we need to increase the complexity in terms of the number of components and the reaction network. In this chapter, we develop a thermodynamic theory for the evolution of sequences in coexisting phases. We discuss the enrichment of sequences through phase separation towards and at thermodynamic equilibrium, and how this enrichment can be modified when the system is maintained away from equilibrium through fragmentation. This chapter contains results that have been published in Ref. [4], stemming from work done in collaboration with Giacomo Bartolucci, Dieter Braun, and Christoph A Weber. For reproducibility, figure parameters are listed in Appendix M.3. For reading convenience, we repeat parts of the necessary theory from chapter 2, as we introduce the model.

5.1 Introduction

The ability to store genetic information and catalyze chemical reactions makes oligonucleotides, such as DNA and RNA, key building blocks for the molecular origin of life [28, 29, 32, 31]. Without any specialized biological molecules, such as protein-based enzymes on early Earth, the mechanisms that can give rise to the emergence of long oligomers with prebiotic functions remain elusive [47, 48, 49, 50, 51]. A particularly aspired function is folding into a secondary configuration capable of replicating unfolded oligonucleotide sequences. However, a fundamental problem is

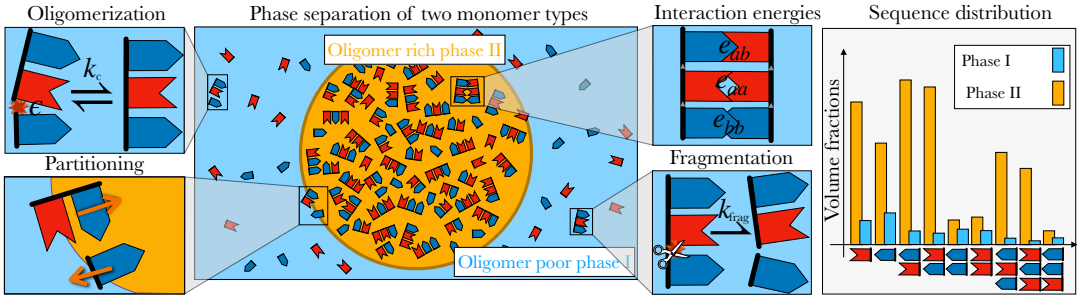


Figure 5.1: **Different monomers oligomerize to sequences that can fragment and interact, leading to oligomer-rich and oligomer-poor phases:** Two distinct monomeric units a and b reversibly oligomerize through monomer pickup and release with a rate coefficient k_c and a free energy ϵ associated with each bond along the oligomer backbone. Sequence-dependent interactions between oligomers can lead to phase separation, resulting in the coexistence of an oligomer-rich and an oligomer-poor phase, mediating the partitioning of monomers and sequences. Sequences interact through nucleotide base-specific interactions, with interaction parameters e_{ab}, e_{aa} and e_{bb}. We also discuss the case when the system is maintained away from equilibrium by a fragmentation pathway occurring with a rate coefficient k_{frag}. Due to non-dilute conditions, phase separation and reversible oligomerization are mutually coupled. This coupling gives rise to complex sequence distributions that differ between the two coexisting phases, I and II.

the exponentially large number of different sequences, particularly for the minimal sequence length required for complex functionalities such as replication [45]. This problem renders the need for selection mechanisms at the molecular origin of life in the early stages, when sequences grow in length and start exploring their exponential sequence space. Various selection mechanisms were proposed based on the non-equilibrium conditions presumably acting on early Earth [204], such as salt and temperature gradients [53, 54], wet-dry [55, 56] or freeze-thaw cycles [57], shear-driven replication [58, 59], accumulation at liquid-vapor interfaces [60] or mineral surfaces [61], template aided-ligation [205, 206, 207], and finally via phase separated condensates [134].

The significance of phase-separated condensates in providing a selection mechanism at the molecular origin of life was suggested roughly a century ago [100]. These ideas have been recently revisited due to advances in sequencing technologies capable of resolving large sequence distributions [208], complemented by observations of phase separation of oligonucleotide sequences under prebiotic conditions. In particular, it was shown that the attractive interactions among different oligomeric sequences can result in the formation of additional phases such as hydrogels [209, 210], liquid crystalline phases [211, 212], and coacervates [213, 134]. Interestingly, the saturation concentrations above which oligonucleotides such as single-stranded DNA and RNA phase separate can be rather low, in the order of μM, due to the strong interaction strength ($\sim 5k_BT$) of complementary base pairs [213, 214]. Such low saturation concentrations make oligonucleotide phase separation a likely and robust

scenario, even under varying physicochemical conditions at early Earth. Furthermore, non-equilibrium thermodynamics implies that the chemical processes such as polymerization, ligation, and fragmentation of oligonucleotides are generically coupled with phase separation of oligonucleotides [113, 215]. This coupling gives rise to feedback between the evolving sequence distribution and its propensity to form condensed phases. In other words, when considering the condensed phases as phenotypes and their local sequence distributions as genotypes, thermodynamics provides a coupling between phenotype and genotype – a prerequisite for Darwinian evolution.

The coupling between chemical processes and phase separation is expected to mediate a selection mechanism that depends on the inter-sequence interactions. Indeed, phase separation was shown to enrich gradually specific sequences from a pool composed of a few designed sequences [134]. However, it remains an open question whether the interaction among sequences, and particularly phase separation, can direct the oligomerization of sequences. If phase separation can efficiently direct oligomerization, it could represent a physicochemical mechanism that reduces the occupations of the exponentially growing sequence space. The coupling between sequence distributions and phase separation becomes especially interesting in the presence of competing effects to aggregation, namely hydrolysis [216, 207]. Phase separation creates a phase that is both oligomer-rich and solvent-poor, possibly serving as a protective environment against hydrolysis, steering the fragmentation of specific sequences.

In this work, we developed a theoretical framework using non-equilibrium thermodynamics to describe the oligomerization of sequences at non-dilute conditions. Such conditions require accounting for the interactions among all sequences, including the possibility of phase coexistence, as illustrated in Fig. 5.1. Using our theoretical framework, we find that the interactions among sequences direct oligomerization toward sequences that phase separate from the solvent and form a condensed phase. We show that phase separation favors forming more cooperatively interacting sequences over specialized sequences that interact only weakly with most other sequences. To unravel the role of phase separation in sequence selection, we maintain the system out of equilibrium by fragmentation inspired by the hydrolysis of oligonucleotides. We find that slow fragmentation rates select for sequences with high information content, while low-information sequences with highly correlated sequence motifs emerge at sufficiently high fragmentation rates. Thus, phase separation can mediate a selection pressure for sequences through selective oligomerization.

5.2 Theory for oligomerization at phase equilibrium

To describe the dynamics of sequences in phase-separated systems, we use the volume fraction, ϕ_i^α, where α denotes the phase. Here, the component label i includes monomers and oligomers summarised in the set σ. For two monomeric building blocks, denoted a and b, the union of the set of monomers $\sigma_m = \{a, b\}$ and sequences $\sigma_o = \{aa, ab, ba, bb, \ldots\}$ gives all non-solvent components $\sigma = \sigma_m \cup \sigma_o$. The volume fraction of a component, $\phi_i^\alpha = \nu_i N_i^\alpha / V^\alpha$, quantifies the volume occupied by N_i^α molecules of type i with molecular volume ν_i, relative to the phase volume V^α. For simplicity, we consider an incompressible system where molecular volumes ν_i of each component are constants.

In this work, we build on a theoretical framework developed for non-dilute chemical kinetics where, if phase separation occurs, the different phases α are at phase equilibrium [113]. In this framework, the time evolution of the components volume fractions $\phi_i^\alpha(t)$ in each phase α is governed by oligomerization rates r_i^α, the diffusive exchange fluxes between phases j_i^α, and changes in the phase volumes V^α (see Fig. 5.1 for an illustration of the different physical processes):

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi_i^\alpha = r_i^\alpha - j_i^\alpha - \frac{\phi_i^\alpha}{V^\alpha}\frac{\mathrm{d}}{\mathrm{d}t}V^\alpha\,, \tag{5.1a}$$

with the component label $i \in \sigma$ denoting monomers and oligomer sequences. In an incompressible system, oligomerization conserves volume ($\sum_{i\in\sigma} r_i^\alpha = 0$), such that the phase volume $V^\alpha(t)$ changes in time only due to partitioning fluxes of solvent j_s^α, and partitioning fluxes of monomers and sequences j_i^α ($i \in \sigma$) through the phase boundary:

$$\frac{1}{V^\alpha}\frac{\mathrm{d}}{\mathrm{d}t}V^\alpha = -j_s^\alpha - \sum_{i\in\sigma} j_i^\alpha\,. \tag{5.1b}$$

The oligomerization rate r_i^α of component i results from the gain $r_{j+m\rightleftharpoons i}^\alpha$ and loss $r_{i+m\rightleftharpoons j}^\alpha$ contributions related to the possible oligomerization pathways. For simplicity, we discuss monomer pickup and release (Fig. 5.2), leading to

$$r_i^\alpha = \sum_{j\in\sigma}\sum_{m\in\sigma_m}\left(r_{j+m\rightleftharpoons i}^\alpha - r_{i+m\rightleftharpoons j}^\alpha\right)\,. \tag{5.1c}$$

This oligomerization rate conserves the total volume (and mass) of monomers and monomers incorporated into sequences. In a non-dilute system, the gain and loss contributions can be written as follows [117]:

$$r_{i+m\rightleftharpoons j}^\alpha = k_{imj}^\alpha\left[\exp\left\{\frac{\bar{\mu}_i^\alpha + \bar{\mu}_m^\alpha}{k_B T}\right\} - \exp\left\{\frac{\bar{\mu}_j^\alpha}{k_B T}\right\}\right]\,, \tag{5.1d}$$

where the terms in the square bracket denote the forward $r_{i+m\rightharpoonup j}$ and backward $r_{i+m\rightharpoonup j}$ reaction rates respectively, and $\bar{\mu}_j^\alpha$ denote the exchange chemical potential of component j in phase α. In Eq. (5.1d), forward and backward rates obey detailed balance of the rates [118]:

$$\frac{r^\alpha_{i+m\rightharpoonup j}}{r^\alpha_{i+m\rightharpoonup j}} = \exp\left\{\frac{\bar{\mu}_i^\alpha + \bar{\mu}_m^\alpha - \bar{\mu}_j^\alpha}{k_B T}\right\}. \tag{5.1e}$$

There is a single kinetic rate coefficient k^α_{imj} for each chemical reaction $i+m \rightleftharpoons j$ governing the relaxation toward chemical equilibrium, $\bar{\mu}_i^\alpha + \bar{\mu}_m^\alpha = \bar{\mu}_j^\alpha$. For simplicity, we consider the kinetic rate coefficients constant, phase independent, and agnostic to the reactants, such that $k^\alpha_{imj} = k_c$.

The non-dilute conditions require accounting for interactions among all components, the solvent, monomers, and oligomers. On a coarse-grained level, these interactions can be described by contributions to the free energy density $f(\{\phi_i\})$, which makes the chemical potential that characterizes the free energy cost to add component i to a mixture, $\mu_i(\{\phi_j\}_{j\in\sigma}) = \nu_i \partial f/\partial\phi_i|_{\phi_{j\neq i}}$, dependent on all components $j \in \sigma$. The free energy density f contains mixing entropy contributions and enthalpic terms up to second order in each sequence volume fractions ϕ_i; see section 2.1. First-order enthalpic terms in f correspond to internal free energies that govern oligomerization equilibria in the dilute limit. In our model, the internal free energies of sequences depend on sequence length, and internal free energy scales are set by backbone bond energy ϵ (Fig. 5.1). We consider ϵ as sequence-independent and study values in the order of a few k_BT, consistent with biofilaments such as DNA and RNA [217]. We note that for peptides, nearest neighbor interactions along the sequences are crucial, leading to a sequence-dependent internal free energy [218, 219]. The interactions between sequences i and j are captured by second-order terms in the free energy and are of the form $e_{ij}\phi_i\phi_j$, where e_{ij} is the average interaction parameter. The quantity e_{ij} is calculated by performing a Boltzmann average over all possible combinations of nucleotide interactions (corresponding to base-pairs for DNA and RNA) between sequence i and j, as illustrated in Fig. 5.2(b). Inspired by RNA and DNA, we investigate cases where nucleotides interact favourably with their opposite, i.e., $e_{aa} = e_{bb} = 0$ and $e_{ab} < 0$ (attractive due to Watson-Crick base pairing). For more information about the interactions, see Appendix H.

Interactions between sequences can lead to phase separation once enough sequences have formed via oligomerization. If diffusion rates of sequences between phases are fast compared to oligomerization rates, phases are at phase equilibrium at each time of the oligomerization kinetics (see section. 2.3.2 for a detailed discussion and derivation). For two coexisting phases $\alpha \in \{\mathrm{I}, \mathrm{II}\}$, chemical potentials $\bar{\mu}_i^\alpha$ and osmotic pressures $\Pi^\alpha = f^\alpha - \sum_{i\in\sigma}\phi_i^\alpha\bar{\mu}_i/\nu_i$ are balanced: $\bar{\mu}_i^{\mathrm{I}} = \bar{\mu}_i^{\mathrm{II}}$ and $\Pi^{\mathrm{I}} = \Pi^{\mathrm{II}}$.

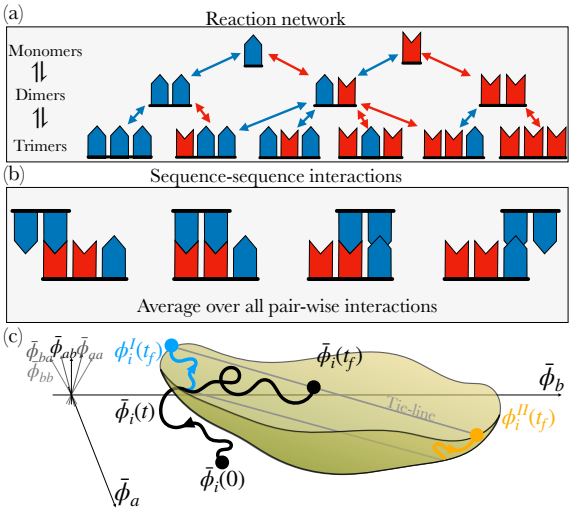


Figure 5.2: **Reaction network, sequence-sequence interactions, and oligomerization kinetics in the high-dimensional sequence space:** (a) The reaction network for monomer pickup and release is illustrated for a subset of the sequence space (monomers, dimers, and trimers). (b) Interactions between sequences are determined by sliding over all possible pair-wise sequence configurations (Eq. (H.2)). We average over such configurations to calculate an interaction parameter for each sequence pair (Eq. (H.3)). (c) Oligomerization creates a trajectory in the high-dimensional sequence space. Initializing the system in a homogeneous state with an average volume fraction $\bar{\phi}_i(t=0)$, the oligomerization trajectory $\bar{\phi}_i(t)$ (black) may intersect with the binodal manifold (green surface), leading to two coexisting phases. The phases I and II exhibit different compositions of oligomers (oligomer-rich and oligomer-poor) that are connected by a tie-line. Phase compositions $\phi_i^{\mathrm{I/II}}(t)$ move along the binodal manifold (blue and orange) and relax to thermodynamic equilibrium or a non-equilibrium steady state (e.g., with non-equilibrium fragmentation) at time t_f.

In this work, we focus on the case of two coexisting phases and compare the results to the case where phase separation is suppressed (mixed case). To support this focus, we looked for multi-phase equilibria for the parameters investigated in this work, finding exclusively two-phase coexistence. During the oligomerization kinetics in the presence of two phases at phase equilibrium, the oligomer-rich (II) and oligomer-poor (I) phases are constrained to the surface of a high-dimensional binodal manifold, where tie-lines connect the two coexisting phases; see Fig. 5.2c. The constraint of phase equilibrium during oligomerization is satisfied by solving $\partial_t \bar{\mu}_i^{\mathrm{I}} = \partial_t \bar{\mu}_i^{\mathrm{II}}$ and $\partial_t \Pi^{\mathrm{I}} = \partial_t \Pi^{\mathrm{II}}$, and self-consistently calculating the diffusive fluxes between the phases, j_i^{α}; for more details see section 2.3.2. Note that the condition of phase coexistence implies the α-superscript of the chemical potentials in Eqs. (5.1c), (5.1d),(5.1e), can be omitted. Furthermore, as the reaction rate coefficient k_c is phase-independent, it implies an equal oligomerization rate r_i in the oligomer-poor and -rich phase throughout the chemical kinetics.

Since chemical potentials govern both the oligomerization rates (Eq. (5.1d)) as well as the diffusive fluxes between the coexisting phases, there is a mutual coupling between oligomerization and phase separation. This coupling is a fundamental thermodynamic property of chemically reacting, non-dilute mixtures (also away from equilibrium); see section 2.3.

To understand the role of phase separation in the oligomerization kinetics, we consider a mixed system as a reference. The mixed system is homogeneous, and the

reaction kinetics follow

$$\mathrm{d}\phi_i/\mathrm{d}t = r_i, \tag{5.2}$$

with the oligomerization rate given in Eqs. (5.1c) and (5.1d). In simple terms, it is the same system as the phase-separated system, i.e., with the same physical parameters, but where phase coexistence is suppressed. We note that the mixed state has a higher free energy, requiring non-equilibrium mixing to maintain the system away from phase equilibrium.

In the following, we numerically integrate Eqs. (5.1) to evolve the volume fractions of monomers and sequences, $\phi_i^\alpha(t)$ with $i \in \sigma$ and $\alpha \in \{\mathrm{I}, \mathrm{II}\}$. The average volume fractions are found from a weighted sum over both phases,

$$V\bar{\phi}_i = V^{\mathrm{I}}\phi_i^{\mathrm{I}} + V^{\mathrm{II}}\phi_i^{\mathrm{II}}, \tag{5.3}$$

where V is the total system volume $V = V^{\mathrm{I}} + V^{\mathrm{II}}$. For all studies in this work, we initialize the system at $t = 0$ with an equal amount of a and b monomers and no other non-solvent components. Specifically, $\bar{\phi}_a(t = 0) = \bar{\phi}_b(t = 0)$, and $\bar{\phi}_i(0) = 0$ $\forall i \in \sigma_o$, which further sets the total sequences volume fraction $\bar{\phi}_{\mathrm{tot}} = \sum_{i\in\sigma} \bar{\phi}_i$ that is conserved during the oligomerization kinetics.

5.3 Sequence evolution toward thermodynamic equilibrium

Dimers, trimers, and longer oligomers emerge from a pool of monomers through oligomerization. As more and longer oligomers form, the decrease in mixing entropy and the stronger attractive interactions enhance the propensity to phase separate. When phase separation occurs at $t = t^*$ (Fig. 5.3a), an oligomer-rich phase (phase II) forms that coexists with an oligomer-poor phase (phase I). The two phases can differ by several orders of magnitude in their total oligomer volume fraction. The corresponding phase average typically exceeds the one of the mixed system governed by Eq. (5.2), where phase separation is suppressed.

The interactions among sequences in the presence of phase separation bias the oligomerization kinetics, leading to the enrichment of specific sequences. As examples, we show the evolution of two different 5-mers, an alternating sequence *ababa* (Fig. 5.3b) and a homo-5-mer *aaaaa* (Fig. 5.3c). Upon phase separation, the volume fractions in the two phases differ significantly between the two sequences. However, this difference is much more pronounced for the alternating sequence than for the homopolymeric sequence. This is because phase separation arises from attractive interactions between sequences. Alternating sequences tend to interact favorably with

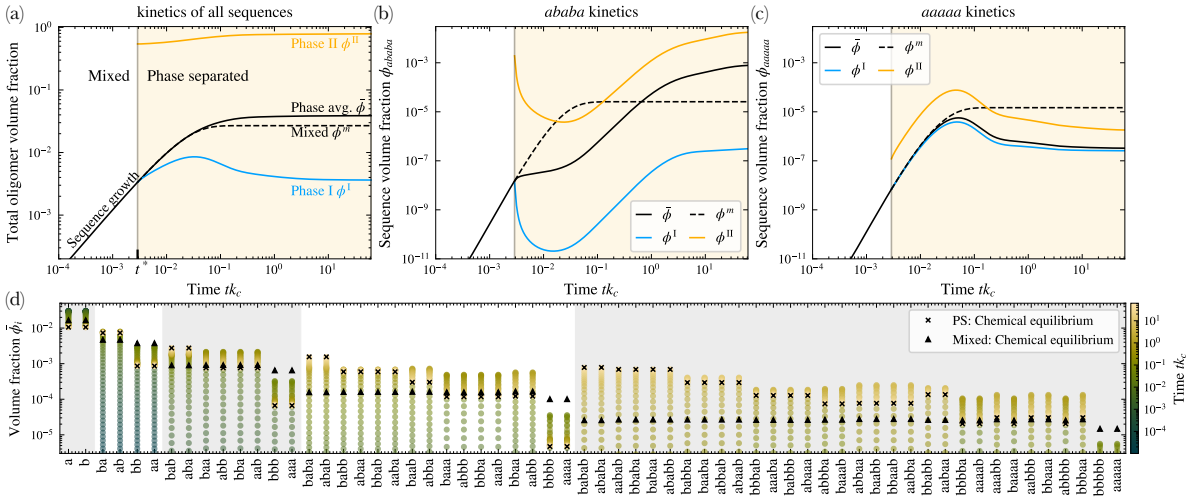


Figure 5.3: **Oligomerization kinetics triggers phase separation leading to the enrichment of specific sequences:** After initializing the system with the two monomers a and b, oligomer sequences form, increasing their volume fractions $\phi_i(t)$, causing the total oligomer volume fraction $\sum_{n=2} \bar{\Phi}(n)$ (Eq. (5.4)) to increase (a). Once sufficient oligomers have formed, the interactions among sequences trigger the formation of two coexisting, phase I (blue) and II (orange), with different volume fractions in each phase. After phase separation, different sequences undergo specific kinetics, leading to different phase compositions at thermodynamic equilibrium. Here, we display the dynamics of the specific oligomers $i = ababa$ (b) and $i = aaaaa$ (c). In the phase separation domain, we also show the average volume fraction (black, Eq. (5.3)). As a control, we compare it to the mixed system where phase separation is suppressed (dashed). (d) The complex time evolution of the sequence distribution ranging from monomers to five-mers is generally non-monotonic, where time is indicated by the color bar. While the mixed reference case (triangle) shows an almost flat distribution for each sequence length at chemical equilibrium, there are significant variations in the average volume fraction $\bar{\phi}_i(t)$ when the system can phase-separate (cross). Note that the sequences are ordered by their Hamming distance to the alternating sequence of equal length, and sub-ordered by the number of a-b-neighbors.

many others in the system (we later refer to these oligomers as cooperative), whereas homosequences interact strongly only with their homopolymeric complements. Most importantly, the sequence interactions also affect the oligomerization, biasing the formation toward such strongly interacting, phase-separation-prone sequences. Conceptually, the mutual feedback between oligomerization and phase equilibrium arises from both processes being governed by the free energy (see Sect. 5.5 for the discussion when the system is actively maintained away from equilibrium). This bias is reflected by the larger values of the average volume fractions with $\bar{\phi}_{ababa}(t) \gg \bar{\phi}_{aaaa}(t)$ for late times t. It also explains why many more similarly alternating sequences form in the entire system at thermodynamic equilibrium (black crosses in Fig. 5.3d).

To unravel the role of the mutual feedback between phase separation and oligomerization leading to the strong enrichment of specific sequences, we considered a mixed reference system (see dashed curves in Fig. 5.3a-c). Physically, a mixed system is

homogeneous, and phase separation is suppressed. Mathematically, this case corresponds to only considering the oligomerization kinetics in the non-dilute mixture by solving Eq. (5.2). In Fig. 5.3a, we see that the mixed system leads to slightly less total oligomer volume fraction $\sum_{n=2} \bar{\Phi}(n)$ than the phase-separated system, where $\bar{\Phi}(n)$ is the phase-averaged total volume fraction of oligomers of length n, defined as

$$\Phi^{\alpha}(n) = \sum_{i \in \sigma} \phi_i^{\alpha} \delta_{n,n_i} \,. \tag{5.4}$$

In other words, phase separation only slightly enhances the net trend of monomers forming oligomers. For specific sequences, however, the time evolution towards chemical equilibrium is completely altered by phase separation. This effect can be seen by comparing the dashed (mixed) with blue (phase I) and orange (phase II) lines in Fig. 5.3b-c for the two 5-mers, and when considering the complete sequences distribution shown in Fig. 5.3d. The two sequences depicted in Fig. 5.3b and 5.3c differ by less than a factor of two when mixed, while for the phase average, the same ratio is more than 10^3. Phase separation enhances *ababa* multiple orders of magnitude and suppresses *aaaaa* to the extent that it is more abundant for the mixed system. In Fig. 5.3d, the final chemical equilibrium volume fractions for each sequence are indicated by crosses, while triangles display the stationary values for the mixed state. We observe that the mixed case shows little diversity between sequences of the same length, resulting in an almost flat sequence distribution. Note that the mixed system behaves completely differently from the phase-separated case, which gives rise to specific sequences being strongly enriched while others are strongly depleted.

A striking property of our theoretical framework is that it allows us to follow the kinetics of the full sequence distribution in time (Fig. 5.3d). Longer oligomers form at the expense of depleting the amount of initialized monomers. At early times, the kinetics is dominated by the entropic gain to distribute the monomer mass to larger oligomers. This fast entropy-driven growth at intermediate times causes the oligomer volume fractions to exceed their equilibrium value before relaxing to it due to sequence-sequence interactions (see e.g., Fig. 3c). At later times, the enthalpic gains through interactions balance with the entropic costs of unequally distributed mass between sequences of the same length. In the mixed case, the system is not dense enough for the interactions to alter the chemical kinetics significantly, resulting in a simple monotonic increase in the volume fractions (dashed lines in Fig. 5.3b and c). Eventually, thermodynamic equilibrium is reached where phase coexistence and chemical equilibrium are concomitantly satisfied, and the sequence distribution in each phase becomes stationary (black crosses in Fig. 5.3d). At thermodynamic equilibrium, the most and least abundant sequences among 5-mers differ in vol-

ume fraction by more than three orders of magnitude, as shown in Fig. 5.3b and c. Phase separation thus strongly alters kinetics and the sequence distribution at thermodynamic equilibrium.

5.4 Sequence distribution at thermodynamic equilibrium

At chemical equilibrium, the chemical potentials of all sequences i are given by the chemical potential of the monomers

$$\bar{\mu}_i^\alpha = a_i \bar{\mu}_a^\alpha + b_i \bar{\mu}_b^\alpha \,, \tag{5.5}$$

where a_i is the number of monomers of type a in oligomer i, and b_i correspondingly for type b, such that the sequence length n_i is $n_i = a_i + b_i$. This condition represents that at chemical equilibrium, the chemical potential of oligomer i must be equal to the chemical potential of its constituting monomers. When this condition is fulfilled, all oligomerization fluxes vanish ($r_i = 0$). If this condition is concomitantly satisfied with the conditions of phase coexistence, the system is at thermodynamic equilibrium. Mixed systems with a composition inside the binodal can satisfy chemical equilibrium, but will not be at thermodynamic equilibrium.

The condition of chemical equilibrium in Eq. (5.5) states that the chemical potentials of all oligomers i can be expressed as a weighted average of the monomer chemical potentials. Thus, it introduces $\mathcal{S} = \mathcal{C} - 2$ additional constraints to Gibbs' phase rule (Eq. (2.31)), yielding

$$\mathcal{P} \leq 3. \tag{5.6}$$

Chemical equilibrium thus heavily constrains the number of possible phases at thermodynamic equilibrium. With the given interaction parameters ($\chi_{AB} < 0$), there is no possibility of three-phase coexistence. The maximum number of phases at chemical equilibrium is therefore two.

From the chemical equilibrium condition, we see that it is the chemical potential of the monomeric units that determines the chemical potential of all other sequences through Eq. (5.5), and hence their volume fraction. This leaves two independent degrees of freedom, namely the total volume fraction for each nucleotide basis,

$$\phi_A^\alpha = \sum_{i \in \sigma} \frac{a_i \phi_i^\alpha}{n_i}, \qquad \phi_B^\alpha = \sum_{i \in \sigma} \frac{b_i \phi_i^\alpha}{n_i}. \tag{5.7}$$

These two quantities uniquely determine the chemical equilibrium sequence distribution in a given phase. As oligomerization conserves nucleotides, the total oligomer

volume fraction of each nucleotide type, namely $\bar{\phi}_A$ and $\bar{\phi}_B$, are conserved during the kinetics and must be specified when initializing the system. To find the length distribution, the total volume fraction can be decomposed into a total oligomer volume fraction for each length defined in Eq. (5.4). To determine the sequence distribution, we must calculate the chemical potentials. From the free energy in Eq. (2.12), the chemical potential of an oligomer (Eq. (5.5)) reads

$$\frac{\bar{\mu}_i}{\rho_i\, k_B T} = \frac{1}{\rho_i}\log\left(\frac{\phi_i}{\rho_i}\right) - \log\left(\phi_s\right) + \frac{1}{\rho_i} - 1 + \frac{\omega_i - \omega_s}{k_B T} + \sum_{j\in\sigma}\phi_j\frac{\chi_{ij}-\chi_{js}}{k_B T} + \frac{\phi_s\chi_{is}}{k_B T}\,. \tag{5.8}$$

Using the chemical equilibrium condition from Eq. (5.5), the sequence distribution is found to be

$$\begin{aligned}\phi_i^\alpha = \rho_i\left(\frac{\phi_a^\alpha}{\rho_1}\right)^{a_i}\left(\frac{\phi_b^\alpha}{\rho_1}\right)^{b_i}\exp\{\rho_i-\rho_1\}\times\exp\left\{-\frac{\rho_i}{k_B T}\left(\omega_i - \frac{a_i\omega_a + b_i\omega_b}{a_i+b_i}\right)\right\} \\ \times\exp\left\{-\frac{\rho_i\phi_s^\alpha}{k_B T}\left(\chi_{is} - \frac{a_i\chi_{as}+b_i\chi_{bs}}{a_i+b_i}\right)\right\}\times\exp\left\{-\frac{\rho_i}{k_B T}\sum_{j\in\sigma}\phi_j^\alpha\left[\chi_{ij} - \frac{a_i\chi_{aj}+b_i\chi_{bj}}{a_i+b_i}\right]\right\}.\end{aligned} \tag{5.9}$$

The first factor before the $\times$ represents the mixing entropy, resulting in an exponential decay with length. The remaining terms characterize the energy change from distributing a_i and b_i monomers into an oligomer i, changing the internal energy ω_i, and interaction energies with the solvent χ_{is}, and among sequences χ_{ij}. The more negative the energy change, the larger ϕ_i^α will be.

With our choice of interactions and internal energies, one can write the equilibrium distribution as

$$\begin{aligned}\phi_i^\alpha = \rho_i\left(\frac{\phi_a^\alpha}{\rho_1}\right)^{a_i}\left(\frac{\phi_b^\alpha}{\rho_1}\right)^{b_i}\exp\left\{(\rho_i-\rho_1)\left(1-\frac{\epsilon}{k_B T}\right)\right\} \\ \times\exp\left\{-\frac{\rho_i}{k_B T}\sum_{j\in\sigma}\phi_j^\alpha\left[e_{ij} - \frac{a_i e_{aj}+b_i e_{bj}}{a_i+b_i}\right]\right\}.\end{aligned} \tag{5.10}$$

Notably, the solvent interactions e_{is} do not affect the chemical equilibrium sequence distribution, though they affect the chemical reaction kinetics and phase compositions ϕ_A^α, ϕ_B^α. For two oligomers of equal length in a phase where $\phi_A^\alpha = \phi_B^\alpha$, or alternatively $a_i = a_j$ and $b_i = b_j$, we find that the non-solvent interaction terms

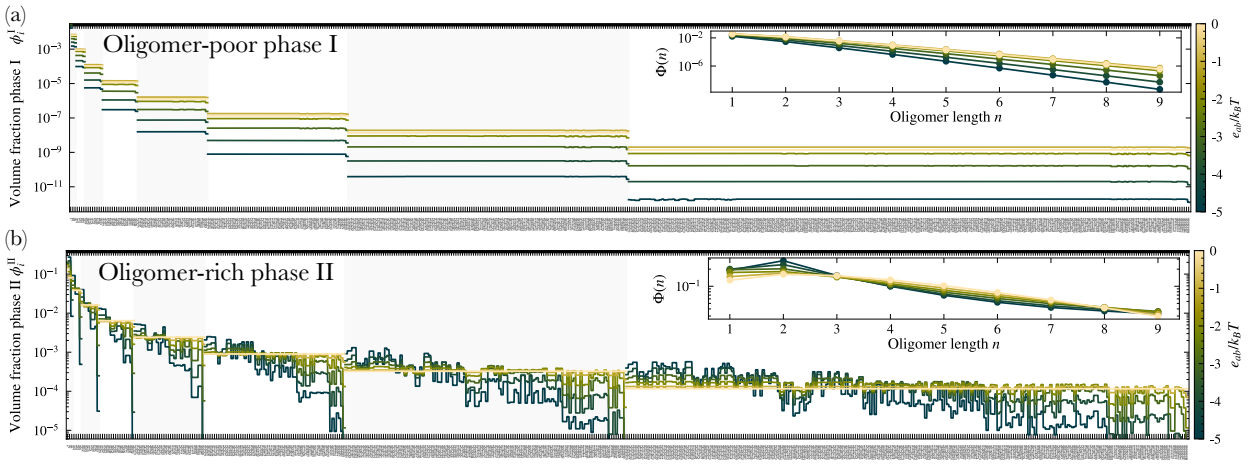


Figure 5.4: **Sequence interactions yield complicated sequence distributions in the dense phase:** Sequence distributions in the oligomer-poor phase I (a) and the oligomer-rich phase II (b). While the sequences are approximately equally distributed in the oligomer-poor phase I for each oligomer length l, there are strong variations in the oligomer-rich phase II. The prerequisite for the enrichment or depletion of specific sequences via phase separation is a strong enough interaction strength e_{ab} (values in k_BT indicated by color code). Sequences are ordered by their Hamming distance to the alternating sequence of equal length, and sub-ordered by the number of a-b-neighbors. Inset: The length distribution of oligomers is weakly affected by the interaction strength.

solely set the relative volume fraction

$$\frac{\phi_i^\alpha}{\phi_j^\alpha} = \exp\left\{-\frac{\rho_i}{k_BT}\left(\sum_{k\in\sigma}\phi_k^\alpha\left[e_{ik}-e_{jk}\right]\right)\right\} \times \exp\left\{-\frac{\rho_1}{k_BT}\sum_{k\in\sigma}\phi_k^\alpha\left(e_{ak}\left(a_i-a_j\right)+e_{bk}\left(b_i-b_j\right)\right)\right\}. \tag{5.11}$$

We see that in a dilute system where $\bar{\phi}_{\text{tot}} \to 0$, all sequences of each length will be equally abundant. As the total sequence volume fraction $\bar{\phi}_{\text{tot}}$ increases, the oligomer-oligomer interactions can co-enrich sequences that adapt to the sequence patterns in the respective phase, leading to more cooperative sequences. Each sequence is affected by interactions with all other sequences and, in turn, affects them as well.

The full sequence distribution in the oligomer-poor and oligomer-rich phases at thermodynamic equilibrium is depicted in Fig. 5.4. As Eq. (5.11) predicts, the oligomer-poor phase (5.4a) does not contain sufficient oligomer material for enrichment or depletion of specific sequences. Thus, entropy dominates, leading to sequences of the same length being equally distributed, independent of the interaction strength e_{ab}. This is not the case in the oligomer-rich phase (Fig. 5.4b), where a strong enrichment and depletion is observed. For each length, the x-axis is ordered after the number of a-b-neighbours. The sequence distribution shows the general trend of an enrichment of sequence with many a-b-neighbours, and depletion

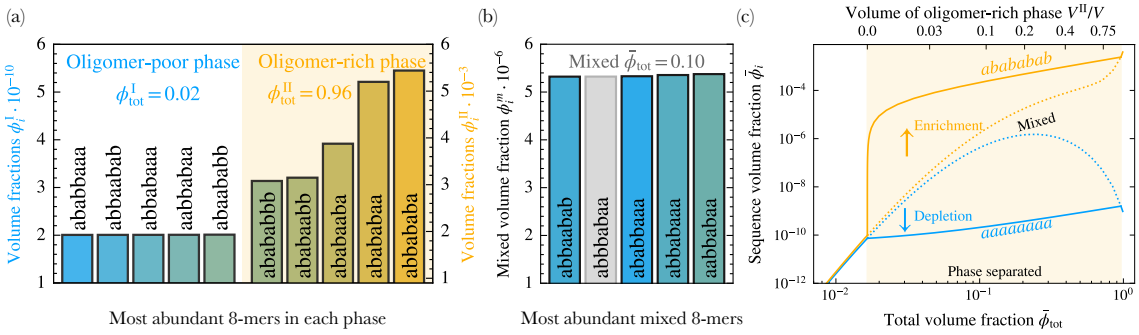


Figure 5.5: **Phase separation strongly enriches and depletes specific sequences:** (a) The volume fraction of the five most abundant 8-mer sequences indicates that the oligomer-rich phase II biases oligomerization toward alternating sequences (within approx. one Hamming distance) by a compositional enhancement of around seven orders of magnitude relative to the oligomer-poor phase I. Moreover, in contrast to the oligomer-rich phase II, phase I also shows no significant compositional variations among sequences. (b) In the mixed system ($\bar{\phi}_{\mathrm{tot}} = 0.10$), the five most abundant sequences are similar to the oligomer-poor phase I and show only little compositional variations. Increasing $\bar{\phi}_{\mathrm{tot}}$ toward $\phi^{\mathrm{II}}_{\mathrm{tot}}$, the mixed phases consistently approach the composition of the oligomer-rich phase II (see Fig. 5.6). (c) In the phase-separated regime ($\phi^{\mathrm{I}}_{\mathrm{tot}} < \bar{\phi}_{\mathrm{tot}} < \phi^{\mathrm{II}}_{\mathrm{tot}}$), the relative enrichment and depletion is the largest for small volumes of the oligomer-rich phase II, V^{II} (see upper plot axis). While homopolymers (e.g., *aaaaaaaa*) get depleted at thermodynamic equilibrium, alternating sequences (e.g., *abababab*) become enriched by many orders of magnitude compared to the respective mixed cases (dotted).

of those with few.

The key property of phase separation is that the coexisting phases act as local hubs with very different total oligomer mass and, thereby, very different inter-sequence interactions (Fig. 5.3a). These interactions affect oligomerization locally, enriching or depleting specific sequences. This hub-like effect is evident by comparing the five most abundant sequences in each phase (Fig. 5.5a). While the oligomer-poor (I) phase has no clear preference for any sequence (see also Fig. 5.4a), the oligomer-rich (II) phase shows a clear preference towards alternating sequences, each of varying abundance. For the corresponding mixed system with a total volume fraction $\bar{\phi}_{\mathrm{tot}} = 0.10$ (Fig. 5.5b), the distribution is almost flat, similar to the oligomer-poor phase. Most importantly, phase separation has a dual effect: It changes which sequences are enriched and depleted and also magnifies their depletion and enrichment. This can be seen by sequences in the oligomer-rich (II) phase (indicated by orangish color code) that do not appear in either the oligomer-poor (I) or the mixed case. Consistently, the oligomer-poor phase and the mixed system at low total oligomer volume fraction enrich similar sequences such as *aabbabaa*.

The enrichment mediated by phase separation is most pronounced for small volumes of the oligomer-rich phase II. This property is apparent by comparing the equilibrium volume fraction averaged over both phases (Eq. (5.3)) with the mixed case where phase separation is suppressed, for varying the total volume fraction of

monomers and oligomers $\bar{\phi}_{\text{tot}}$; see Fig. 5.5c. Here, phase separation abruptly enriches the volume fraction of the alternating 8-mers, while depleting the homo-8-mers. The relative enrichment is the largest for small volumes of the oligomer-rich phase. This behavior is due to the volume fractions of long sequences increasing exponentially with the total mass (Eq. (5.9)), while the phase average increases linearly with the composition of the oligomer-rich phase (Eq. (5.3)). As a result, there is a relative enrichment of specific sequences up to a thousand-fold. For the mixed system, the two sequences depicted remain at approximately equal abundance until around $\phi_{\text{tot}} \simeq 0.1$, while phase separation is able to differentiate them at more dilute conditions. Phase separation thus allows for the enrichment of specific sequences through the local creation of a dense environment at conditions only slightly above the saturation volume fraction, ($\phi^{\text{I}}_{\text{tot}} \lesssim \bar{\phi}_{\text{tot}} \ll \phi^{\text{II}}_{\text{tot}}$), which would not be possible without phase separation. We note that weakly supersaturated conditions are common in experimental systems on oligonucleotide phase separation [220, 221, 222, 223].

For the most abundant 8-mers in a mixed system, displayed in Fig. 5.6, homo-sub-strands of *aa* and *bb* appear more often than in the oligomer-rich phase, as displayed in Fig. 5.5a. As $\bar{\phi}_{\text{tot}}$ increases, the most enriched sequences change from more similar to the dilute phase (bluish) to more similar to the oligomer-rich phase (orangish). When $\bar{\phi}_{\text{tot}}$ approaches the oligomer-poor or -rich phase, the sequence distribution becomes the same. The mixed volume fractions of the most abundant sequences are generally much smaller than their corresponding quantity for the phase average (indicated by the dashed green bars). The relative increase through phase separation is the largest for smaller $\bar{\phi}_{\text{tot}}$, as already observed in Fig. 5.5c. With an increasing value of $\bar{\phi}_{\text{tot}}$, the distributions become less flat, and a depletion through phase separation is possible (indicated by the black bar). The fact that the most enriched sequences in the oligomer-rich phase differ from those enriched when mixed indicates that phase separation is not merely a magnifier of volume fraction but also introduces a bias.

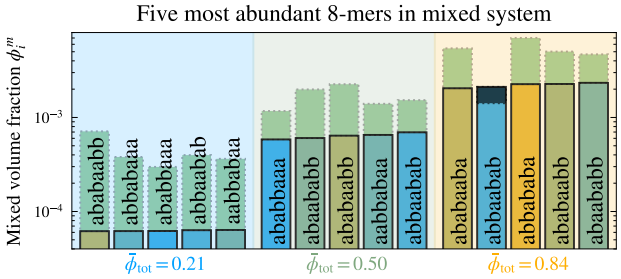


Figure 5.6: **Phase separation enriches specific sequences and alters enrichment bias:** In a mixed system, the most abundant 8-mers homo-sub-strands of *aa* and *bb* appear more often than in the oligomer-rich phase, as displayed in Fig. 5.5a. Phase separation generally increases the volume fractions (indicated by the dashed green bars), but for dense enough systems, depletion through phase separation is possible (indicated by the black bar). Phase separation is not solely a magnifier of sequences but creates a bias.

Sequence interactions collectively enrich similar sequence patterns in the respec-

tive phases. In the oligomer-poor phase, the oligomers are too dilute to affect each other's propensity through interactions. In the oligomer-rich phase, the oligomer-oligomer interactions determine the sequence distribution, as seen in Fig. 5.4b. Without interactions ($e_{ab} = 0$) or a too-weak interaction strength e_{ab}, the distribution remains approximately flat for each length, independent of the total volume fraction of monomers and sequences $\bar{\phi}_{\text{tot}}$. With stronger interactions (e_{ab} more negative), the volume fractions of the enriched/depleted sequences may enhance/deplete further (see Fig. 5.7a for 8-mers). However, a non-monotonous behavior can be observed for some sequences when increasing the interaction strength e_{ab}. For example, *abaababbb* is enriched at $e_{ab} = -4k_BT$, but depleted at $-5k_BT$. Another example is *aababababaa* being the most abundant sequence at $e_{ab} = -3k_BT$, but that is far from the case at $-5k_BT$. The non-monotonous dependence on the interaction strength e_{ab} originates from enrichment being a collective effect of the sequence distribution. For intermediate interaction strengths, the enrichment of *ab* is not strong enough for many repetitions of the *ab* pattern in the same sequence to minimize interaction energies; instead, a combination with *aa* and *bb* minimizes the interaction energies. For stronger interaction, the energetic cost of *aa* and *bb* patterns depletes the previously enriched sequences. The effect is collective because the entire sequence distribution determines the optimal patterns.

Sequence interactions and phase separation mediate a strong bias toward sequences with almost alternating sequence patterns at thermodynamic equilibrium (Fig. 5.5a Fig. 5.7a). This behavior is evident in the probability distribution for the proportion of *a*-*b* neighbors for 8-mers. The proportion of *a*-*b* neighbors takes the value of 0 for homo-oligomers and the value of 1 for a perfectly alternating oligomer (Fig. 5.7b, illustration left/right). Without interactions, the distribution is a binomial distribution symmetric around an *a*-*b* neighbor proportion of a half. With stronger interactions e_{ab}, the mean of the distribution shifts to a higher number of *a*-*b* neighbors, and the variance of the distribution decreases. The results support that oligomers close to homo-oligomers are statistically disfavoured as they have fewer complementary strand motifs where sequences can attach. Due to the volume fractions decaying with length (see Eq. (5.9)), the interactions with shorter sequences are the most relevant for enhancing motifs in longer sequences. Though alternating patterns are preferred, a perfect alternating pattern does not produce any complementary sub-strand for other non-perfectly alternating sequences and is, hence, not as enriched. The strongest enrichment occurs for approximately alternating sequences that deviate slightly from a perfect alternating pattern, allowing them to interact favorably with *aa* or *bb* segments of other sequences.

Phase separation and interactions among oligomers create long-range correlations

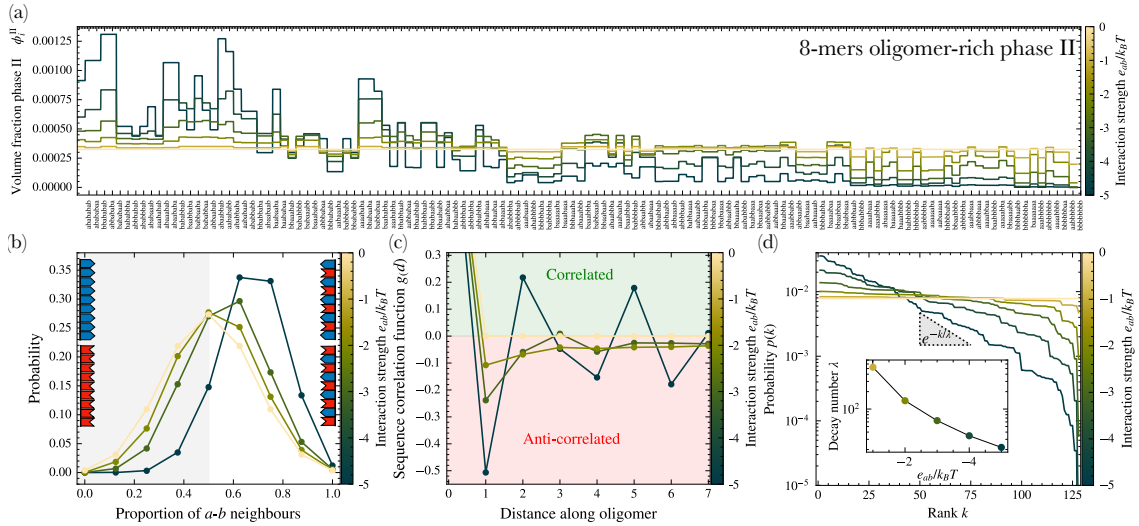


Figure 5.7: **Formation of alternating sequences requires phase separation and strong enough sequence interactions.** (a) There are strong variations between the equilibrium volume fraction of sequence in the distributions of 8-mers in the oligomer-rich phase II. The enrichment or depletion of specific sequences originates from sequence-sequence interactions e_{ab} (values indicated by color code). (b,c) Alternating sequences require strong enough attractive interactions, e_{ab} This trend is reflected in (b), showing the a-b neighbor distribution getting more skewed toward more a-b neighbors, and (c) with the sequence correlation function $g(d)$ (Eq. (5.12)) developing a clear oscillatory behavior with a wave number of about two monomers (i.e., units of ab). (d) Ordering the sequences by abundance gives the rank distribution $p(k)$, which decays exponentially, $p \propto e^{-k/\lambda}$. The characteristic rank scale λ decreases with more adhesive interactions e_{ab}. This behavior indicates that more monomers are shuffled to a smaller number of sequences to enhance their abundance.

along the oligomer sequence. The corresponding autocorrelation function

$$g(d,l) = \frac{\sum_{i\in\sigma} \delta_{n_i,l} \sum_{j=1}^{l} \phi_i \sum_{\pm} \begin{cases} 1 & \text{if } \Gamma_{i,j\pm d} = \Gamma_{i,j} \\ -1 & \text{if } \Gamma_{i,j\pm d} \neq \Gamma_{i,j} \end{cases}}{2l \sum_{i\in\sigma} \phi_i \delta_{n_i,l}}, \tag{5.12}$$

characterizes the occurrence of the same monomers along an oligomer of fixed length l. In the equation above, $\Gamma_{i,j}$ gives the monomer-type at position j of oligomer i. The auto-correlation function $g(d)$ shows pronounced oscillations for large enough interaction strength e_{ab} (Fig. 5.7c). Stronger interactions (more negative values of e_{ab}) give rise to stronger correlations, while no interactions ($e_{ab} = 0$) leave all sequences uncorrelated. The wave number of the correlation function $g(d)$ is approximately two, corresponding to the alternating pattern ab, which is statistically most common.

The stronger sequences interact (more negative e_{ab}), the more of the total monomer mass gets turned into enriched sequences. This trend can be quantified by the rank distribution $p(k)$ that orders the sequences by their abundance (rank k). We find

that the rank distribution for each length decays exponentially with a characteristic rank scale λ, such that $p(k) \propto \exp(-k/\lambda)$. The characteristic decay number λ decreases with the strength of interaction e_{ab}. Without interactions ($e_{ab} = 0$), the sequence distribution for a given length is flat, and so is the rank distribution $p(k)$ (yellow line shown in Fig. 5.7d). With increasing interaction strength, the more alternating sequences are enriched while most other sequences are depleted. The characteristic rank scale λ thus decreases with interaction strength e_{ab} (inset of Fig. 5.7d). As most sequences are depleted for large interaction strengths e_{ab}, stronger interactions reduce the size of the effectively occupied high-dimensional sequence space.

The preference towards alternating sequences originates from the necessity of all sequences to interact with both a's and b's. The preference for alternating sequences results from the interaction strength e_{ab} and the volume fractions of all sequences being large enough, which both affect the distribution in an exponential fashion (Eq. (5.10)).

While large sequences are dilute compared to short oligomers (dimers, trimers, ...), the dominant contributions in the exponential leading to the enrichment of alternating patterns in long oligomers (7-, 8-, 9-mers), stem from short oligomers (see inset Fig. 5.4b). The origin of short alternating sequences arises from them interacting more favorably with their environment, meaning their respective phases, i.e., they are more cooperative. In Fig. 5.3d and 5.4b for example, the dimer ab is much more abundant than both aa and bb, as it interacts favorably with both monomer types ($e_{aa,a} = 0$, $e_{aa,b} = e_{ab}$, while $e_{ab,a} = e_{ab,b} \approx e_{ab}$). But why are the most enriched sequences approximately one Hamming distance away from the perfectly alternating 8-mer $abababab$, and not the perfectly alternating 8-mer itself? This is because a single non-alternating segment, aa or bb, enhances the average interactions with the diverse sequence environment, further boosting a sequence's cooperativity. The number of such non-alternating segments in the enriched sequences decreases as the interaction strength is increased (see Fig. 5.7(a)).

5.5 Sequence selection by non-equilibrium fragmentation

To decipher the effects of phase separation in sequence selection away from equilibrium, we break detailed balance of the rates between the fragmentation and fusion pathways $i \rightharpoonup j + |i - j|$. In general, both pathways rely on the break-up of bonds along the sequence backbone, where sequence i fragments into the sub-oligomer/monomer j and $|i - j|$. An example is $aabab \rightharpoonup aa + bab$. Detailed balance

of the rates is broken by omitting the thermal fusion (ligation) pathway in the following, considering it to be slow compared to fragmentation: see Appendix J for further studies with ligation leading to qualitatively similar results. Thus, the system is maintained away from equilibrium. The rate of the fragmentation pathway for sequence i in phase α is written as

$$h^{\alpha}_{i\rightharpoonup j+|i-j|} = k_{\text{frag}} \frac{\phi^{\alpha}_{s}\phi^{\alpha}_{i}}{n_i - 1} \delta_{j\in\mathcal{S}_i} \,, \tag{5.13}$$

where k_{frag} denotes the fragmentation rate coefficient that we choose, for simplicity, to be phase-independent. The fragmentation rate $h^{\alpha}_{i\rightharpoonup j+|i-j|}$ is proportional to the local volume fraction ϕ^{α}_{i} ($\alpha = \text{I,II}$) of the fragmenting sequence. Each bond along the sequence backbone fragmentates equally, such that all possible length combinations of the two fragmentated sub-parts of i occur equally often. This is captured by the fragmentation rate being independent of j, as long as j is a continuous sub-string of i ($\delta_{j\in\mathcal{S}_i} = 1$), otherwise the rate is zero ($\delta_{j\notin\mathcal{S}_i} = 0$). The factor $(n_i - 1)^{-1}$, where n_i is the number of monomers in sequence i, normalizes the rate by the number of different fragmentation pathways. Inspired by the hydrolysis of biofilaments such as DNA and RNA, the fragmentation rate scales with the solvent volume fraction ϕ^{α}_{s}. In contrast to the studies discussed in Sect. 5.3, the system will not relax to thermodynamic equilibrium; rather, it settles in a non-equilibrium steady state, where generally there exists a net reaction r^{α}_{i} for all sequences in each phase α, which is balanced by the diffusive flux across the phase boundary. An emerging stationary sequence distribution is thus particularly distinct when the system is phase-separated from its mixed counterpart, where the oligomerization flux vanishes exactly at the non-equilibrium steady state.

Breaking detailed balance of the rates in the oligomerization kinetics gives rise to a selection pressure that strongly changes the sequence distribution compared to thermodynamic equilibrium. For the parameters studied, we always observe a non-equilibrium steady state with a stationary sequence distribution. For low fragmentation rate coefficients k_{frag}, alternating sequences are favored, and the sequence distribution remains almost unchanged compared to thermodynamic equilibrium, as shown by the yellow distribution in Fig. 5.8a. For intermediate values of k_{frag} (light green), the thermodynamic sequence bias is no longer evident, and the non-equilibrium sequence distribution at steady state is almost flat with approximately equal volume fractions for each sequence. For very large fragmentation rate coefficients k_{frag}, we find that the thermodynamic bias of sequences observed for vanishing k_{frag} inverts, i.e., sequences that are most depleted at thermodynamic equilibrium become most abundant in the non-equilibrium steady state, and vice versa.

The inverted thermodynamic bias for large fragmentation rate coefficients k_{frag}

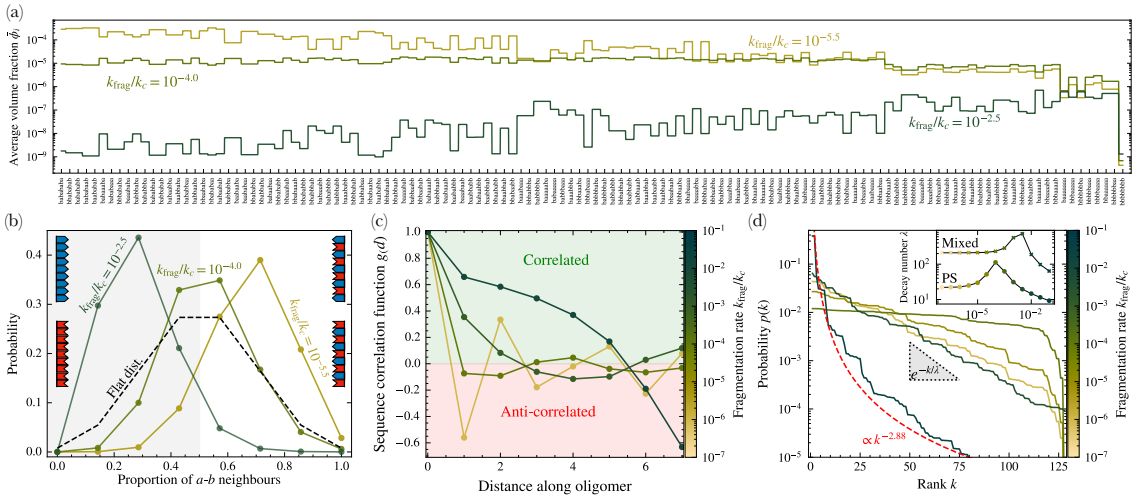


Figure 5.8: **Non-equilibrium fragmentation creates a selection pressure favoring sequences with longer blocks of the same monomer type:** (a,b) With increasing fragmentation rate coefficient $k_{\rm frag}$, the selection bias changes from alternating sequences to sequences with longer blocks of the same monomer type. For very large $k_{\rm frag}$, these blocks approach the maximal sequence length (8mers in (a)), and the most abundant sequences deviate from homopolymers within 1-2 Hamming distances. (b,c) The selection trend for increasing $k_{\rm frag}$ from alternating sequences to sequences with larger blocks is supported by (b) showing a pronounced shift toward fewer a-b-neighbors and a pronounced increase of the wave number of the sequence autocorrelation function $g(d,l)$ (see (c)). (d) The rank distribution $p(k) \propto \exp(-k/\lambda)$ decays exponentially for smaller fragmentation rate coefficients $k_{\rm frag}$. The characteristic rank scale λ (inset) has a non-monotonous behavior and is maximal for intermediate values of $k_{\rm frag}$ because the sequence distribution is approximately flat (a). For the very large fragmentation rate coefficients $k_{\rm frag}$, the rank distribution follows a power-law decay with $p(k) \propto k^{-\beta}$ illustrated by the exponent $\beta \simeq 2.88$ for $k_{\rm frag}/k_c = 10^{-2.5}$.

originates from the thermodynamically unfavorable oligomers having the largest oligomerization rates. A sequence i is at its steady state when its net oligomerization flux r_i and its fragmentation flux h_i balance, that is $r_i = \sum_k h_{i\to k+|i-k|}$. The flux r_i scales exponentially with the chemical potential $\bar{\mu}_j$ of sequences j that form i upon monomer-pickup $j + m = i$. Thus, sequences j that are more homopolymeric, increasing their chemical potential, thus balancing the fragmentation rate h_i at larger values of ϕ_i. The result is, following the derivation in Appendix K, a non-equilibrium steady state (NESS) sequence distribution in the oligomer-rich phase (II):

$$\phi_i^{\rm II,NESS} \simeq \frac{k_c}{k_{\rm frag}} \frac{\exp\left(\frac{\bar{\mu}_m}{k_B T}\right)}{\frac{V^{\rm II}}{V}\phi_s^{\rm II}} \sum_{j\in\mathcal{R}_i} \exp\left(\frac{\bar{\mu}_j}{k_B T}\right) , \tag{5.14}$$

where $\bar{\mu}_m$ is the monomer exchange chemical potential (Eq. (5.8)). We note that the effects of other sequences j on the steady state abundance of sequence i is described by the term $\sum_{j\in\mathcal{R}_i} \exp\left(\bar{\mu}_j/(k_B T)\right)$. Thus, a sequence i is more abundant if the set of sequences that oligomerize into i ($j \in \mathcal{R}_i$) have unfavorable interac-

tions with their environment, reflected in a larger value of their chemical potential $\bar{\mu}_j$. In other words, non-equilibrium fragmentation selects for sequences of lower cooperativity, which are homo-oligomer-like sequences. This selection mechanism is the opposite of the thermodynamic enrichment mechanism that favors sequences of higher cooperativity.

For increasing fragmentation rate coefficient k_{frag}, the most abundant sequences change from alternating oligomers to oligomers containing low complexity domains composed of the same monomeric unit. This trend is illustrated by the five most abundant sequences (Fig. 5.9) and by the probability distribution of *a*-*b*-neighbors (Fig. 5.8b). We see that the five most abundant sequences for weak non-equilibrium driving ($k_{\text{frag}}/k_c = 10^{-5.5}$) are alternating sequences with maximally two monomers of the same type neighboring each other. In this case, the *a*-*b*-neighbor distribution is skewed toward an *a*-*b*-proportion of 1, indicating that most *a*-monomers are neighbored by *b* monomers. Increasing k_{frag} to intermediate values enhances the length of domains composed of the same monomer type. For example, the five most abundant sequences contain sequences similar to *abbbaba*, containing both alternating and homo-oligomeric parts. The consequence is an *a*-*b*-neighbor distribution that peaks around a half, but is less spread than the flat distribution. When increasing the fragmentation rate coefficient k_{frag} further, domains composed of the same monomer type approach the maximum length, leading to almost homopolymeric sequences such as *abbbbbbb* or *abaaaaaa*. Consistently, the *a*-*b*-distribution is more skewed toward a small fraction of *a*-*b*-neighbors for increasing k_{frag}. To quantify the shift from alternating sequences to ones with extended domains of the same monomer type, we use the correlation function defined in Eq. (5.12). Fig. 5.8c confirms that faster fragmentation leads to an increase in the wave number of the correlation function, corresponding to longer correlated domains of the same monomer type.

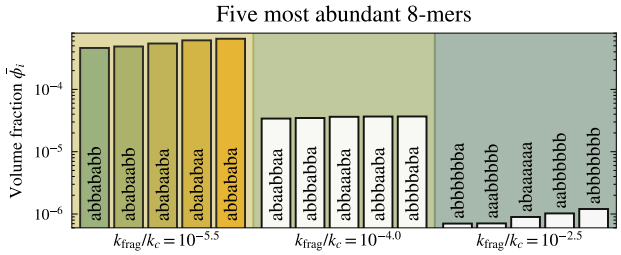


Figure 5.9: **Five most abundant sequences away from equilibrium when phase-separated:** With increasing fragmentation rate coefficient k_{frag}, the selection bias changes from alternating sequences to sequences with longer blocks of the same monomer type. For very large k_{frag}, these blocks approach the maximal sequence length, and the most abundant sequences deviate from homopolymers within 1-2 Hamming distances.

For weak non-equilibrium driving, the exponential decay of the rank distribution of 8-mers, $p(k) \simeq \exp\{(-k/\lambda)\}$ persists, as displayed in Fig. 5.8d. However, for intermediate fragmentation rates k_{frag}, the characteristic decay length λ increases

to a maximum before decreasing. At the maximum, the sequence distribution is almost flat. A steeper rank distribution (smaller λ) indicates a stronger selection pressure, where only a small subset of the sequences are selected. For very large fragmentation rate coefficients k_{frag}, the rank distribution approximately follows a power-law $p(k) \propto k^{-\beta}$ with $\beta \simeq 3$ for $k_{\text{frag}}/k_c = 10^{-1.5}$.

Power-laws in rank distributions were reported for many complex systems, such as city sizes, gene circuits, and self-replicating systems, and were shown to be a requirement for open-ended evolution [224]. Though here, the exponent is different from Zipf's law [225]. We note that without phase separation (mixed case), there is also a power-law behavior of the rank distribution for very large k_{frag} with an exponent close to Zipf's law. For details, see Fig. I.1 in Appendix I.

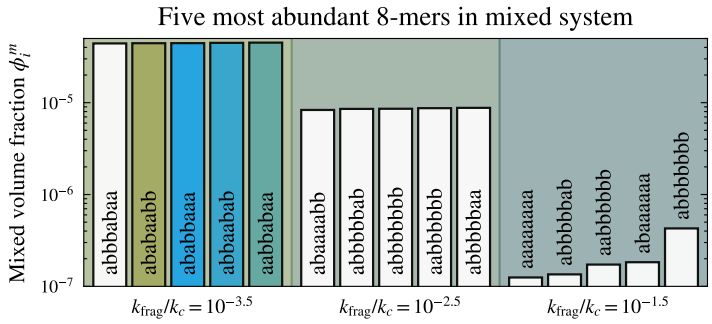


Figure 5.10: **Fragmentation enhances same sequences for mixed system, but less pronounced:** The most abundant 8-mers behave similarly when mixed compared to when phase separated (displayed in Fig. 5.8b), though a larger value of k_{frag} is necessary. Still, the same trend of changing the selection bias from alternating sequences to longer blocks of the same monomer type with increasing k_{frag} appears.

For a mixed system, a similar sequence selection as for phase-separated systems (displayed in Fig. 5.8) occurs with increasing fragmentation rate, as seen in Fig. 5.10. For slow fragmentation rates, the most enriched sequences are similar to mixed systems at chemical equilibrium (displayed in Fig. 5.5b). With increasing k_{frag}, the most enriched 8-mers have longer blocks of the same monomer type; this happens concomitantly with the sequences distribution becoming flatter (Fig. 5.8d). For large fragmentation rates, the homopolymeric blocks almost span the size of the oligomer, as was observed with phase separation. For a mixed system, a larger value of k_{frag}/k_c is necessary for the same type of sequence selection as for a phase-separated system.

Maintaining oligomerization away from equilibrium gives rise to a strong selection mechanism, where only a small subset of all possible sequences accumulates most of the total volume fraction for the given length. This effect is most pronounced for longer sequences. We characterize these observations by the quantity $\#_{90}(l)$ that is the smallest number of sequences corresponding to 90% volume fraction for a given length l (Fig. 5.11a):

$$\frac{\sum_{i\in\sigma}^{\#_{90}(l)} \phi_i \delta_{n_i,l}}{\sum_{i\in\sigma} \phi_i \delta_{n_i,l}} > 0.90\,. \tag{5.15}$$

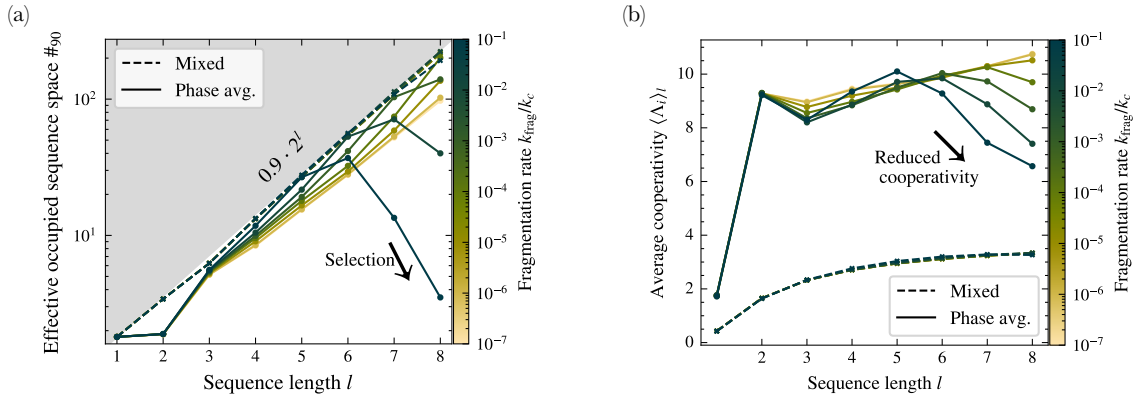


Figure 5.11: **Non-equilibrium fragmentation selects for less cooperative sequences:** (a) The number of sequences $\#_{90}(l)$ occupying 90% of the oligomer volume of a given length l (Eq. (5.15)) deviates from the exponential scaling as observed at thermodynamic equilibrium (vanishing values of k_{frag}). For the large fragmentation rate coefficients k_{frag}, $\#_{90}(l)$ even decreases for long enough sequences, indicating a strong selection pressure selecting a sub-exponential subset of the possible sequence for a given length l. (b) The cooperativity characterizes how well sequences of each length interact with their environment (Eq. (5.16)). Thermodynamics promotes cooperativity (low k_{frag}) and increases with sequence length l as the sequences have more freedom to adapt their structure to suit their environment. Increasing the fragmentation rate coefficient k_{frag} changes the selection pressure to promote oligomers with uncooperative substrates, reducing the average cooperativity at large l.

This quantity characterizes the effectively occupied sequence space. At thermodynamic equilibrium, and independent of the interaction strength e_{ab}, $\#_{90}(l)$ scales exponentially in sequence length l because longer length leads to exponentially more sequences (see Fig. 5.11a for low k_{frag}). Strikingly, the selection pressure, once driven away from equilibrium, gets so strong for fast enough fragmentation (large k_{frag}) such that $\#_{90}(l)$ scales even sub-exponentially (Fig. 5.11a). Though the sequence distribution is less diverse when maintaining oligomerization away from equilibrium, the selection pressure facilitates specific sequence motifs to win over the others. A relevant selection mechanism must contend with the exponentially growing sequence space. In this light, our findings of sub-exponential scalings highlight the potential of phase separation with oligomerization maintained away from equilibrium in selecting specific oligomer sequences already during their oligomerization kinetics.

Selection mediated by phase separation can be further corroborated by introducing a cooperativity measure. The cooperativity Λ_i per sequence length n_i characterizes how well sequence i interacts with its environment:

$$\Lambda_i = \gamma_i^{-n_i}\left(\{\phi_j\}\right) \exp\left\{-\frac{\bar{\mu}_i^0}{n_i\, k_B T}\right\}. \tag{5.16}$$

Here, the chemical potentials have been decomposed in terms of the composition-independent reference chemical potential $\bar{\mu}_i^0$, and the composition-dependent activ-

ity coefficient γ_i, such that $\bar{\mu}_i = \bar{\mu}_i^0 + k_B T \log(\phi_i \gamma_i)$, with the expressions given in Eq. (I.2). Chemical equilibrium enriches oligomers that interact well with sequences in their respective phase, corresponding to a larger average cooperativity (Fig. 5.11b). The longer an oligomer, the more cooperative segments exist at which other sequences can attach. We note that without phase separation (mixed case), sequences are relatively uncooperative and not well-adapted to their mixed environment. The increasing cooperativity trend with sequence length in the phase-separated case at thermodynamic equilibrium changes when oligomerization is maintained away from equilibrium. The cooperativity decreases as the selection pressure inverts the thermodynamic enrichment, leading to less alternating and more homopolymeric sequences. In other words, the faster the fragmentation, the more uncooperative sequences are selected. Cooperative sequences interact well with almost all other sequences. Selection of less cooperative sequences with longer repeating blocks of the same nucleotide can mediate more specific interactions and thereby provide specific functionalities for prebiotic systems [226, 227, 228], such as hairpins [229].

Finally, we ask whether the selected sequences store more or less information. We quantify the information content of each sequence S_i by the number of *a*-*b*-neighbors using Shannon's measure of information [230, 231]:

$$S_i = -\sum_{j \in \sigma} p(j|n_{a,b}^{(i)}) \log_2 \left(p(j|n_{a,b}^{(i)}) \right), \tag{5.17}$$

where $p(j|n_{a,b}^{(i)})$ is the probability of sequence j given the number of *a*-*b*-neighbors $n_{a,b}^{(i)}$ of sequence i, without prior knowledge about the sequence distribution. No prior knowledge corresponds to a flat sequence distribution. This means that each sequence of the given length with a specific $n_{a,b}^{(i)}$ value is equally likely. In this case, Eq. (5.17) gives the maximal value of

$$S_i^* = -\log_2 \{p(i|n_{a,b}^{(i)}))\}. \tag{5.18}$$

Thus, S_i^* quantifies the information gained in bits by knowing the exact sequence, with the only prior information being the number of *a*-*b* neighbors corresponding to sequence i. For example, there are only two sequences of length $l = 8$ that are homo-oligomers ($n_{a,b} = 0$). Thus, the probability of picking the correct homo-oligomer is one-half, resulting in an information gain of 1 bit upon determining the actual sequence. Similarly, the sequence *abbbaaba* has a probability of 1.4% among all sequences of four *a*-*b*-neighbors, yielding an information gain of 6.1 bits. Thus, S_i^* assigns the same value of information to the sequences with the same value of $n_{a,b}$. The average information of a sequence distribution is calculated as

$\langle S(l)\rangle = \sum_{i\in\sigma} \delta_{n_i,l}\bar{\phi}_i S_i^* / \sum_{i\in\sigma} \delta_{n_i,l}\bar{\phi}_i$.

Phase-separated systems away from equilibrium change the information content of the selected sequences. Fig. 5.11c shows the average sequence information of 8-mers, $\langle S\rangle$, at the non-equilibrium steady state for increasing non-equilibrium fragmentation rate coefficient k_{frag}. For phase-separated systems, $\langle S\rangle$ has a non-monotonous dependency on k_{frag}. Close to thermodynamic equilibrium (low k_{frag}), phase separation favors low-information sequences, i.e., alternating sequences. Increasing the fragmentation rate coefficient k_{frag} leads to an increase in the average sequence information, exceeding the corresponding average for a flat distribution of 8-mers. This is because there is an enrichment of sequences with 3 and 4 *a*-*b*-neighbors at the expense of homo-oligomers and alternating oligomers (Fig. 5.8b), which contain less information according to our definition. Note that no such local maximum is observed without phase separation, where a monotonic decrease in average information content with fragmentation rate k_{frag} is observed. The same behaviour is observed with alternative measures, such as Kolmogorov complexity [233], Lempel-Ziv [232], Lempel-Ziv-Welch [234], and the information content of the distribution [230]. Increasing k_{frag} further, in the case with phase separation, selects for longer blocks of the same monomer type, strongly decreasing the average information content. Such low-information sequences are particularly interesting because they have the potential to show rich functionalities, such as catalytic activity or folding, and the extended blocks can mediate specific, lock-like interactions [226, 227, 228, 229].

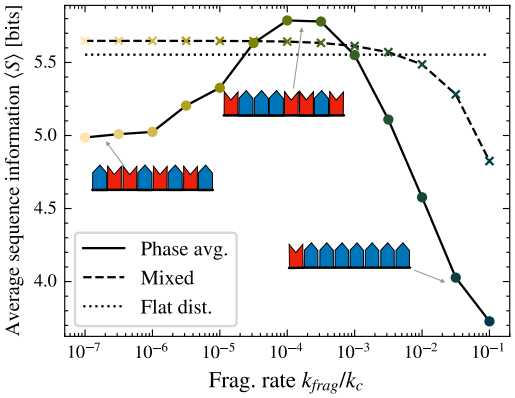


Figure 5.12: **Sequence information increases away from equilibrium:** The average sequence information over the distribution of 8-mers (5.17) has a local maximum with the fragmentation rate coefficient k_{frag}. In the limit of low and high fragmentation rates, the average information content of the sequence distribution decreases due to the selection of alternating sequences or sequences of larger homopolymeric blocks, respectively. Similar behavior is found for different measures of sequence information/complexity, such as the Kolmogorov complexity [224] and the Lempel–Ziv complexity [232]. The dotted horizontal line displays the average sequence information for a flat distribution.

5.6 Discussion and conclusion

In this chapter, we propose a theoretical framework for the oligomerization kinetics of sequences at non-dilute conditions. This framework builds on non-equilibrium

thermodynamics and can be used to explore high-dimensional sequence spaces. Using our framework, we show that the oligomerization kinetics trigger the phase separation of sequences, which in turn creates a very strong bias for the enrichment and depletion of specific sequences. A key finding is that, in sequence distributions at thermodynamic equilibrium, more cooperative sequences that interact well with their environment are favored, while non-equilibrium fragmentation selects for low-information and less cooperative sequences.

Our findings can be tested experimentally considering nucleotide mixtures capable of DNA or RNA polymerization [53, 235, 57]. For large concentrations of nucleotides, long enough sequences can form such that the mixing entropy is low enough to facilitate phase separation into oligomer-rich and oligomer-poor phases. A sufficiently high salt concentration screens the negative charges of the DNA or RNA backbone, enabling sequence-specific, base-pair-mediated interactions under non-dilute conditions. Such interactions were shown to form liquid phases [213, 214] and gels [209, 210]. Recent experiments show that RNA-based condensates are more robust and form over a broader range of conditions than peptide analogues [236]. Furthermore, in the presence of DNA, RNA-based condensates have a reduced viscosity, enabling the diffusion of reactive oligonucleotides and supporting RNA chemistry. Deep sequencing techniques combined with HPLC and mass spectroscopy can screen for more cooperative or less cooperative sequences using, for example, the cooperativity measure proposed in this work.

Our theoretical work is relevant for the molecular Origin of Life [204, 237], and for designing *de novo* life systems capable of undergoing open-ended evolution [66, 3]. In particular, our finding of a sub-exponential effective occupation of the possible sequence space indicates a very strong selection pressure by simply considering non-equilibrium fragmentation. Sequences of low cooperativity and low information are more likely to have exotic sequence patterns or motifs that may provide target-specific molecular functionalities, such as the catalysis of specific molecular components.

Chapter 6

The emergence of a high fidelity replicator

In this chapter, we employ the theory developed in the previous chapter, chapter 5, to investigate the effects of phase separation on replication fidelity. In order to handle the increased complexity of templated ligation and to understand the role of noise, we employ a stochastic approach by solving the master equation using the Gillespie algorithm [238]. We find that phase separation significantly enhances both sequence fitness and fidelity, allowing it to survive Eigen's error threshold, while a homogeneous system under the same conditions does not. Furthermore, we find a bistability in the sequence distribution with different condensate properties, providing a simple phenotype-genotype relationship. The presented work was conducted in collaboration with Christoph A. Weber, and a manuscript is currently in preparation. For reproducibility, figure parameters are listed in Appendix M.4. For reading convenience, we repeat parts of the necessary theory from chapter 2 and 5, as we introduce the model.

6.1 Introduction

A crucial step for the emergence of life is a reaction network capable of storing and replicating information. RNA is often considered pivotal in this transition, as it carries genetic information within its sequence and can catalyze replication [28, 29, 32, 31]. RNA thus serves a dual role, storing genetic information while also functioning as a catalytic molecule that enables both replication and enzymatic processes. Nevertheless, there are several challenges for a system to utilize these properties of RNA.

RNA is chemically less stable than DNA and other biopolymers. In aqueous environments, its backbone is susceptible to spontaneous cleavage, making it inher-

ently unstable, leading to fragmentation of RNA strands [39]. Moreover, the rate of fragmentation generally exceeds that of ligation [40, 41], strongly suppressing the accumulation of long RNA chains. To be able to compete with fragmentation, a highly efficient copying mechanism is necessary.

Not only must the sequence information be copied at a high rate, but it must also be copied with a high enough fidelity for the information to be maintained. Greater oligomer length requires increasingly accurate replication machinery. In modern biology, this challenge is overcome by specialized enzymes that detect and repair errors, enabling genomes of millions of base pairs to be maintained [42]. As highlighted by M. Eigen [43, 44], mutations impose an upper limit on the length of RNA where information can be preserved. Crucially, this limit falls well below the size of the simplest error-correcting enzymes under realistic mutation rates [44]. An alternative, less complex, mechanism of replication is templated ligation.

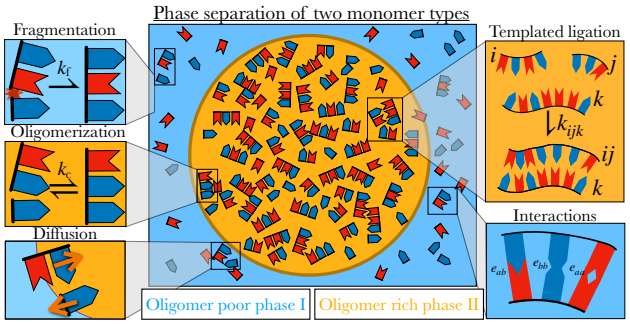


Figure 6.1: **Different processes of phase separating oligomers:** Two distinct monomeric units a and b reversibly oligomerize through monomer pickup and release with a rate coefficient k_c. Sequence-dependent interactions between oligomers can lead to phase separation, resulting in the coexistence of an oligomer-rich (II) and an oligomer-poor (I) phase, mediating the partitioning of monomers and sequences. Sequences interact through nucleotide base-specific interactions, with interaction parameters e_{ab}, e_{aa}, and e_{bb}. Phase separation creates non-dilute conditions and is mutually coupled to oligomerization. Additionally, the irreversible pathways of fragmentation, with rate coefficient k_f and template-aided ligation, with rate coefficient k_{ijk}, maintain the system away from thermodynamic equilibrium. Templated ligation favors products similar to the template, thus possibly serving as a copying mechanism.

Oligomers can grow through the process of ligation, where two oligomers join their ends to form a new, longer oligomer. As oligomers grow in length, the rate of ligation decreases rapidly [239]. More importantly, the process is independent of the sequences ligating and does not act as a replication mechanism. An alternative pathway is one where a third oligomer acts as a template to which two oligomers can hybridize. By the two oligomers hybridizing with the third oligomer, known as the template, they are brought into proximity with one another, facilitating their ligation. The newly ligated oligomer can separate from the template, resulting in the template and the product, allowing for the process to occur again. The process is depicted in Fig. 6.1 and 6.2(b). Due to strong base-pair interactions [240], which sequences hybridize on the template is strongly correlated with the sequence of the template [241].

While end-to-end ligation yields a longer oligomer, templated ligation can additionally act as a copying mechanism, as the template can create a copy of itself [241]. Ligation of two strands on an RNA template thus constitutes an autocatalytic step, with template–product base pairing providing the thermodynamic basis for its selection. The process is autocatalytic, as the template is not consumed in the process, and if the templated ligation accurately duplicates the template, its rate is enhanced further.

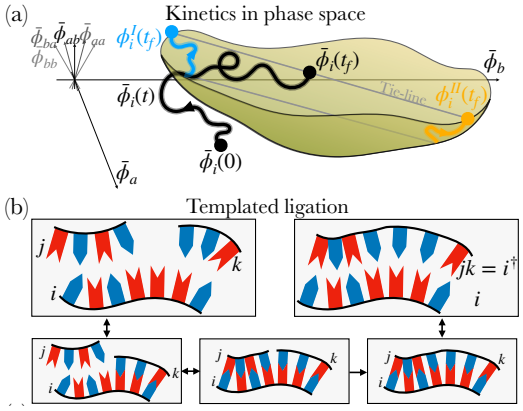


Figure 6.2: **Phase diagram kinetics and templated ligation:** (a) Due to different processes, the system evolves as a trajectory in the high-dimensional sequence space. Initializing the system in a homogeneous state with an average volume fraction $\bar{\phi}_i(t=0)$, the oligomerization trajectory $\bar{\phi}_i(t)$ (black) may intersect with the binodal manifold (green surface), leading to two coexisting phases. The phases I and II exhibit different compositions of oligomers (oligomer-rich and oligomer-poor) that are connected by a tie-line. Phase compositions $\phi_i^{\mathrm{I/II}}(t)$ move along the binodal manifold (blue and orange) and eventually reach a non-equilibrium steady state at time t_f. (b) A sequence i can act as a template by forming a duplex with two strings j and k simultaneously, increasing the chance that j and k ligate. The intermediates (lower row) are not treated explicitly. For the case depicted here, i acts as a template to create its own inverse $i^\dagger$, corresponding to a replication event.

Templated ligation has been shown to replicate sequence patterns, without enzymes [242], and produces non-monotonic length distributions [235], fundamentally different from the exponential decay due to end-end ligation [243, 244]. Sequence-resolved simulations have shown spontaneous symmetry breaking in sequence space to perfectly alternating or homopolymeric sequences [216]. Templated ligation could supply both the kinetics and selectivity needed for information copying without the development of sophisticated enzymatic RNA.

The same interactions that mediate templated ligation as a copying mechanism can cause phase separation between the oligomers and the solvent. Due to the strong attractive interactions between single-stranded RNA molecules, they can phase separate at rather dilute conditions on the order of μM [213, 214]. Furthermore, RNA condensates form under a broader range of conditions and are minimally viscous, allowing for diffusive and reactive activity [236]. Thermodynamics provides a natural link between the phenotype (through the sequence distribution) and the genotype (through the condensate properties). Furthermore, non-equilibrium thermodynamics predicts that chemical processes such as polymerization, ligation, and fragmentation of oligonucleotides are

Process	$\mathcal{W}$	Propensity	Stoichiometric change $(\nu_\alpha)_j$	D.B.
Diffusion I $\to$ II	$\mathcal{D}^{\mathrm{I}}$	$k_D \exp\{\mu_i^{\mathrm{I}}/(k_BT)\}$	$-\delta_{j,i^{\mathrm{I}}} + \delta_{j,i^{\mathrm{II}}}$	✓
Diffusion II $\to$ I	$\mathcal{D}^{\mathrm{II}}$	$k_D \exp\{\mu_i^{\mathrm{II}}/(k_BT)\}$	$\delta_{j,i^{\mathrm{I}}} - \delta_{j,i^{\mathrm{II}}}$	✓
Monomer pick-up	$\mathcal{M}^+$	$k_c \exp\{(\mu_i^\pi + \mu_m^\pi)/(k_BT)\}$	$-\delta_{j,i^\pi} - \delta_{j,m^\pi} + \delta_{j,im^\pi}$	✓
Monomer release	$\mathcal{M}^-$	$k_c \exp\{\mu_i^\pi/(k_BT)\}$	$-\delta_{j,i^\pi} + \delta_{j,\lvert i-m\rvert^\pi} + \delta_{j,m^\pi}$	✓
Fragmentation	$\mathcal{F}$	$k_f \phi_i^\pi \phi_s^\pi$	$-\delta_{j,i^\pi} + \delta_{j,\lvert i-k\rvert} + \delta_{j,k}$	×
Templated ligation	$\mathcal{L}$	$k_{ikl} \phi_i^\pi \phi_k^\pi \phi_l^\pi$	$-\delta_{j^\pi,i^\pi} - \delta_{j^\pi,k^\pi} + \delta_{j^\pi,(ik)^\pi}$	×

Table 6.1: **List of kinetic processes:** The six different processes $\mathcal{W}$ in the master equation (6.1) are listed with their respective propensities and stoichiometric changes. The superscript $\pi \in \{\mathrm{I}, \mathrm{II}\}$ gives the phase index of the component. The first four processes $(\mathcal{D}^{\mathrm{I,II}}, \mathcal{M}^\pm)$ obey detailed balance (D.B.) of the rates (6.2), while the last two $(\mathcal{F}, \mathcal{L})$ maintains the system away from thermodynamic equilibrium.

intrinsically coupled to their phase separation [113, 215], as we also saw in chapter 5. This coupling establishes feedback between the evolving sequence distribution and the tendency of these sequences to phase separate. How the fidelity and fitness of information copying are affected by this coupling remains elusive.

In this chapter, we investigate the interplay between templated ligation and phase separation, and compare it to templated ligation in a homogeneous system. While the replication rate of the homogeneous system remains well above Eigen's error threshold, the phase-separated system is able to survive the threshold, even for long sequence lengths. The error threshold is survived by the selection of a specific sequence pattern across all lengths. Surviving Eigen's error threshold is accompanied by a strong selection in sequence space, where only a small fraction of the exponentially large sequence space is explored. Furthermore, we find that stochastically, the system can transition between the enrichment of different sequence patterns, leading to different properties of the condensate. This coupling provides a simple phenotype-genotype relationship, necessary for Darwinian evolution. Together with the enhanced replication rate below Eigen's error threshold, this indicates condensed phases as an ideal localized environment for de novo life.

6.2 Modeling templated ligation in phase-separated systems

To describe the dynamics of sequences in phase-separated systems, we evolve the particle number, N_i^π, where π is the phase index $\pi \in \{\mathrm{I}, \mathrm{II}\}$. Here, the component label i includes the solvent (s), monomers (m), and oligomers, summarized in the set σ. For two monomeric building blocks, denoted a and b, the union of the set of monomers $\sigma_m = \{a, b\}$ and sequences $\sigma_o = \{aa, ab, ba, bb, \ldots\}$, and the solvent gives all components $\sigma = s \cup \sigma_m \cup \sigma_o$. From the particle numbers N_i^π, we can

calculate the phase volumes $V^\pi = \sum_i N_i^\pi v_i$, and volume fractions of each component, $\phi_i^\pi = v_i N_i^\pi / V^\pi$, where v_i is the molecular volume. For simplicity, we consider an incompressible system where molecular volumes v_i of each component are constants in time, phase-independent, and scale linearly with oligomer length n_i, such that $v_i = n_i \nu_1$ for $i \in \sigma_o$. The thermodynamic framework and components are identical to what was studied in chapter 5. This chapter deviates from the previous as we include the process of templated ligation and solve a stochastic kinetic equation.

The evolution of the probability $P(\boldsymbol{N}, t)$ for the state $\boldsymbol{N} = \{N_i^{\mathrm{I}}, N_i^{\mathrm{II}}\}$ at time t is the difference between the flow to all possible outgoing states from all possible incoming states, such that

$$\partial_t P(\boldsymbol{N}, t) = \sum_{\mathcal{W} \in \mathcal{P}} \sum_{\alpha \in \mathcal{T}(\mathcal{W})} \Big[-\mathcal{W}_\alpha(\boldsymbol{N}) P(\boldsymbol{N}, t) + \mathcal{W}_\alpha\left(\boldsymbol{N} - \boldsymbol{\nu}_\alpha\right) P\left(\boldsymbol{N} - \boldsymbol{\nu}_\alpha, t\right) \Big]. \quad (6.1)$$

Changes in the probability occur due to six different processes: diffusive exchange between phases ($\mathcal{D}^{\mathrm{I/II}}$), monomer-pickup and release ($\mathcal{R}^{\pm}$), fragmentation ($\mathcal{F}$), and template-aided ligation ($\mathcal{L}$), collected in the set $\mathcal{P} = \{\mathcal{D}^{\mathrm{I}}, \mathcal{D}^{\mathrm{II}}, \mathcal{R}^{+}, \mathcal{R}^{-}, \mathcal{F}, \mathcal{L}\}$. Each process $\mathcal{W}$ has a corresponding set of possible transitions $\alpha \in \mathcal{T}(\mathcal{W})$, with the vector $\boldsymbol{\nu}_\alpha$ denoting the stoichiometric change vector for the corresponding process. The processes and system are illustrated in Fig. 6.1, and a table of all processes, stoichiometric changes, and propensities is given in Table 6.1.

Diffusion and monomer pickup/release obey detailed balance of the rates, that is, the ratio of the forward ($\mathcal{W}_{\rightharpoonup}$) and backward ($\mathcal{W}_{\leftharpoondown}$) rates obey [118]

$$\frac{\mathcal{W}_{\rightharpoonup}}{\mathcal{W}_{\leftharpoondown}} = \exp\left\{ \frac{\Delta\mu}{k_B T} \right\}, \quad (6.2)$$

where $k_B T$ is the thermal energy and $\Delta\mu$ is the chemical potential difference of the process, quantifying the thermodynamic force. For diffusion of component i and monomer pickup/release, the thermodynamic force is respectively $\Delta_{\mathcal{D}} = \mu_i^{\mathrm{I}} - \mu_i^{\mathrm{II}}$ and $\Delta_{\mathcal{M}} = \mu_i + \mu_m - \mu_{|i+m|}$, where m is either monomer a or b. The notation $|i+m|$ refers to the index of a sequence equal to i, with a monomer m attached at the end, e.g. $i = aab$ with $|i+a| = aaba$ Thus, without fragmentation and templated ligation, the system would relax to a thermodynamic equilibrium state, as studied in section 5.3 and 5.4. Diffusion and oligomerization are driven by the same thermodynamic forces, the chemical potentials,

$$\mu_i^\pi = \left(\frac{\partial G}{\partial N_i^\pi} \right)_{N_{j \neq i}, T, p}, \quad (6.3)$$

which quantify the change in the Gibbs free energy G (defined in Eq. (3.2)) upon changes in particle number of the respective component. Diffusion between phases

equilibrates the chemical potential between the phases ($\mu_i^{\mathrm{I}} = \mu_i^{\mathrm{II}}$), and the monomer pickup/release equilibrates the chemical potential per length of all sequences, see chapter 5.

Fragmentation and templated ligation, on the other hand, do not obey detailed balance of the rates (Eq. (6.2)), thus maintaining the system away from thermodynamic equilibrium. Fragmentation irreversibly breaks the backbone along a sequence i, creating two sequences j and $|i-j|$ which are continuous sub-strings of the original sequence i. The rate of fragmentation is independent of the bond, but proportional to the volume fraction of the sequence ϕ_i^π times the volume fraction of the solvent ϕ_s^π

$$\mathcal{F}^\pi_{i\to j+|i-j|} = k_f \frac{\phi_i^\pi \phi_s^\pi}{n_i - 1}\,. \tag{6.4}$$

The irreversible ligation of two sequences i and j into the sequence ij can be aided by a template sequence k, that is $i + j + k \rightharpoonup ij + k$ with an associated ligation rate $\mathcal{L}^\pi_{ijk}$

$$\mathcal{L}^\pi_{i+j+k\to ij+k} = k_{ijk}\,\phi_i^\pi \phi_j^\pi \phi_k^\pi\,, \qquad\qquad k_{ijk} = k_{\mathrm{lig}} P_{ik} P_{jk}\,. \tag{6.5}$$

The probability P_{ik} represents the probability that sequence i and sequence k are forming a duplex, which is proportional to their interaction energy $P_{ik} \propto \exp\{-e_{ik}/(k_B T)\}$. This creates a sequence-specific ligation rate that favors the ligation of two sequences that both interact well with the template. Without templated ligation ($k_{\mathrm{lig}} = 0$), the kinetic pathways are identical to the one studied in section 5.5.

Templated ligation works as a copying mechanism, as the template k aids the formation of all sequences, including its inverse sequence $k^\dagger$. The accuracy of replication is quantified by the fidelity

$$Q_k^\pi = \frac{\sum_{i\in\sigma_o}\sum_{j\in\sigma_o}(\mathcal{L}^\pi_{i+j+k\to k^\dagger+k} + \mathcal{L}^\pi_{i+j+k\to k+k})}{\sum_{i\in\sigma_o}\sum_{j\in\sigma_o}\mathcal{L}^\pi_{i+j+k\to ij+k}}\,, \tag{6.6}$$

which quantifies the proportion of ligation events that produce an exact copy k or an inverse copy $k^\dagger$. By inverse, we mean all $a \to b$ and $b \to a$. The fitness of a sequence A_k^π is defined as its net total replication rate, such that

$$A_k^\pi = \frac{1}{\phi_k^\pi}\sum_{i\in\sigma_o}\sum_{j\in\sigma_o}\mathcal{L}^\pi_{i+j+k\to ij+k}\,. \tag{6.7}$$

The definition is given per volume fraction, and the sum over ligation pathways is independent of the product species, serving only to quantify the absolute replication rate. Following the work of Eigen [43, 44], we define the reproductive fitness of sequence k as the product of the fidelity Q_k^π and the fitness A_k^π. A large value of the reproductive fitness implies that a sequence can copy itself with good accuracy and

speed. Eigen showed that for the replication of a sequence to be below the Error threshold [43, 44], its reproductive fitness must be larger than the average fitness

$$\frac{A_k^\pi Q_k^\pi}{\sum_{i\neq k\in\sigma_o}\phi_i^\pi A_i^\pi} > 1\,, \tag{6.8}$$

where we set the threshold for each phase. To compare the reproductive fitness between a phase-separated and a homogeneous system, we compare the left-hand side of this equation averaged over the particle number of k in both phases. If a sequence is below the error threshold, its information content is maintained. Above the error threshold, information is lost because the population spreads randomly across many sequences. Note that our system is not an open system, as studied by Eigen, and the error threshold is therefore not identical to the one presented by Eigen in Ref. [43]. The oligomer with the largest reproductive fitness is defined as the master sequence $i_m = \max_i(Q_i A_i)$.

The interactions between sequences i and j are captured by second-order terms in the free energy of the form $e_{ij}\phi_i\phi_j$, where e_{ij} is the average interaction parameter between i and j. The quantity e_{ij} is calculated by performing a Boltzmann average over all possible combinations of nucleotide interactions (corresponding to base-pairs for DNA and RNA) between sequence i and j, as illustrated in Fig. 6.2. Inspired by RNA and DNA, we investigate cases where nucleotides interact favorably with their opposite, i.e., $e_{aa} = e_{bb} = 0$ and $e_{ab} < 0$ (attractive due to Watson-Crick base pairing). For more information about the interactions, see chapter 5 and appendix H. Importantly, the rate of templated ligation is determined by how the two sequences i and j interact with the template k. Motivated by hybridization rates of DNA and RNA [245], the ligation rate coefficient is at its maximum if both sequences interact well with the template, meaning $e_{ik}, e_{jk} \ll -k_B T$, as seen in Eq. (6.5).

In this chapter, we focus on the case of two coexisting phases and compare the results to the case where phase separation is suppressed (mixed case). During the phase-separated kinetics in the presence of two phases, the oligomer-rich (II) and oligomer-poor (I) phases are close to phase coexistence, but do not satisfy the conditions exactly due to noise and a finite exchange rate between the phases; see Fig. 6.2(a). Furthermore, all rate coefficients k_i, listed in table 6.1, are chosen to be phase-independent.

To understand the role of phase separation on templated ligation, we consider a mixed system as a reference. The mixed system is homogeneous, and the reaction kinetics follow the master equation (6.1), but without the diffusive process $\mathcal{D}^{\mathrm{I/II}}$ and phase index π. It is the same system as the phase-separated system, i.e., with the same physical parameters and total volume fractions, but where phase coexistence is suppressed.

In the following, we solve the master Equation (Eq. (6.1)) through the Gillespie algorithm [238]. That is, the time step τ at time t of reaction event α' is

$$\tau = -\frac{1}{\sum_{\beta\in\mathcal{S}} \mathcal{W}_\beta} \log\left(\xi_1\right), \qquad \alpha' = \min_\gamma \left(\sum_{\alpha\in\mathcal{S}}^{\gamma} \mathcal{W}_\alpha > \xi_2 \sum_{\beta\in\mathcal{S}} \mathcal{W}_\beta \right). \tag{6.9}$$

where $\sum_{\alpha\in\mathcal{S}} = \sum_{\mathcal{W}\in\mathcal{P}} \sum_{\alpha\in T(\mathcal{W})}$ and $\xi_{1/2}$ are two independent random numbers drawn from a uniform distribution between 0 and 1. The state and time is thus changes according to $t \to t+\tau$ and $\boldsymbol{N} \to \boldsymbol{N} + \boldsymbol{\nu}_{\alpha'}$. An advantage of this description is that it allows for a dynamic sequence space, without a predefined maximum length. For further discussion on the Gillespie algorithm, see Appendix L.

The average volume fractions are found from a weighted sum over both phases,

$$V\bar{\phi}_i = V^{\mathrm{I}}\phi_i^{\mathrm{I}} + V^{\mathrm{II}}\phi_i^{\mathrm{II}}, \tag{6.10}$$

where V is the total system volume $V = V^{\mathrm{I}} + V^{\mathrm{II}}$. For all studies in this work, we initialize the system at $t = 0$ with an equal amount of a and b monomers and no other non-solvent components. Specifically, $\bar{\phi}_a(t=0) = \bar{\phi}_b(t=0)$, and $\bar{\phi}_i(0) = 0$ $\forall i \in \sigma_o$, which further sets the total sequences volume fraction $\bar{\phi}_{\mathrm{tot}} = \sum_{i\in\sigma} \bar{\phi}_i$ which is conserved during the oligomerization kinetics.

6.3 bistability of sequence patterns

Before investigating the effects of templated ligation, we study the case without fragmentation and templated ligation, meaning $k_{\mathrm{f}} = 0$ and $k_{\mathrm{tal}} = 0$, in order to investigate the role of noise. Instead, the system is maintained away from equilibrium by a fuel with strength δ^π, as previously investigated in chapter 4. The fuel acts in the same way, where it partitions to the dense phase II, such that $\delta^{\mathrm{I}} = 0$ and $\delta^{\mathrm{II}} = \delta$. The fuel aids the pickup of monomer according to,

$$r^\pi_{i+m\rightleftharpoons|i+m|} = k^\pi_{imj}\left[\exp\left\{\frac{\mu_i^\pi + \mu_m^\pi + \delta^\pi}{k_B T}\right\} - \exp\left\{\frac{\mu^\pi_{|i+m|}}{k_B T}\right\}\right], \tag{6.11}$$

where π is the phase-index $\pi \in \{\mathrm{I}, \mathrm{II}\}$. For the same fueling scheme, we observed bistability with a single monomer type, as shown in section 4.3.2. With a single monomer type, we could solve the sequence distribution analytically to find bistability with one homogeneous stable steady state and one phase-separated steady state for the same total volume fraction. With the more intricate sequence-sequence interactions emerging for oligomers consisting of multiple types of monomeric units, the bistability might exist within the binodal. Due to the increase in complexity from the multiple monomeric units, the bistability is challenging to find visually from the

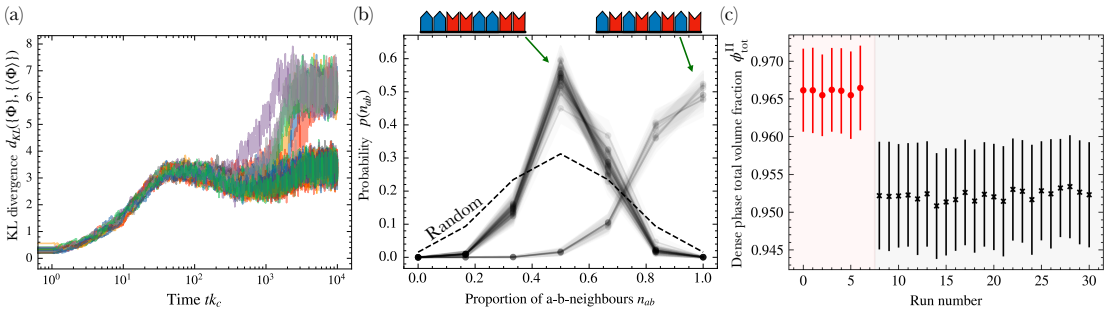


Figure 6.3: **Bistability of sequence patterns:** (a) Symmetric Kullback–Leibler divergence $d_{KL}(\{\Phi\},\{\langle\Phi\rangle\})$ (Eq. (6.12)) between each run and the ensemble average of thirty runs, with volume fractions rescaled to probabilities over sequences. A subset of trajectories deviates from the average at steady state, showing increased divergence. The different trajectories reach different states and stay there, indicating the system is bistable. (b) Sequence selection across runs: most trajectories stabilize into a repeating pattern of length four (*aabbaa*...), while a minority converge to length two (*abab*...), both strongly favored compared to random sequences (dotted line). (c) The two steady states differ in dense-phase composition $\phi^{\mathrm{II}}_{\mathrm{tot}}$ (Eq. (6.13)) and noise level, illustrating a direct genotype–phenotype relationship where selected sequence patterns determine condensate properties.

phase diagram, but can perhaps be found from the stochastic kinetic trajectories.

To quantify the difference between each run in the high-dimensional space, we use the symmetric Kullback–Leibler divergence (KL-divergence). The symmetric KL divergence d_{KL} between two probability distributions p and q over the same space σ is defined as

$$d_{KL}(\{p\},\{q\}) = \frac{1}{2}\sum_{i\in\sigma}\left(p_i \log\left(\frac{p_i}{q_i}\right) + q_i \log\left(\frac{q_i}{p_i}\right)\right) . \tag{6.12}$$

In Fig. 6.3a, the KL-divergence between each run and the ensemble average over thirty runs $d_{KL}(\{\Phi\},\{\langle\Phi\rangle\})$ is displayed. Note that we have rescaled the volume fractions in terms of probabilities among sequences and monomers (the set $\sigma_a = \sigma_m \cup \sigma_o$) according to $\Phi_i = \bar{\phi}_i / \sum_{i\in\sigma_a}\bar{\phi}_i$. All runs have the same parameters and initial conditions. Once a steady state is reached, some of the trajectories deviate from the average trajectory, as seen from an increase in the KL-divergence for a subset of the runs. Analyzing the same 30 runs in (b), we observe that the different runs lead to different sequence patterns being selected. The majority of the trajectories reached a steady state where a repeating alternating sequence pattern with length 4 (*aabbaa*...), while a few reached a pattern which alternates with length 2 (*abab*...). Furthermore, the selection patterns are strongly selected for, with probabilities above 50%, much higher than what is predicted for a random sequence distribution (dotted line).

Interestingly, the different selected sequence patterns result in different properties

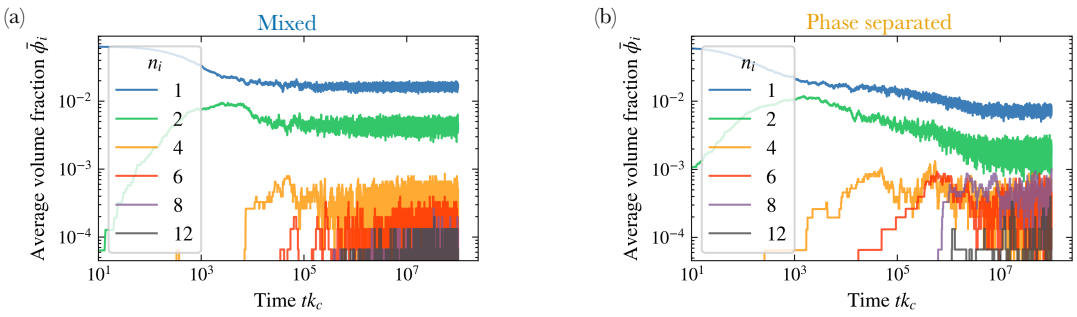


Figure 6.4: **Time evolution of alternating sequences:** The path to a non-equilibrium steady state is reached faster for mixed (a) compared to phase separated (b), but reaches a state with shorter average length and less noise.

of the condensate, as seen in Fig. 6.3c. Ordering the runs by their steady state, a clear distinction beyond the standard deviation can be seen in the dense phase total volume fraction

$$\phi_{\text{tot}}^{\text{II}} = \sum_{i \in \sigma_a} \phi_i^{\text{II}} \,. \tag{6.13}$$

Furthermore, the variation in $\phi_{\text{tot}}^{\text{II}}$ is different between the two steady states. The different condensate properties (phenotype) depend on the selected sequence pattern within the condensate (genotype). A simple phenotype-genotype relationship emerges from the sequence-specific interactions and the interplay between phase separation and the oligomers. A selection mechanism at the genotype level would then influence the sequence distribution.

6.4 Surviving the error threshold

From now on, we investigate the case with fragmentation and templated ligation without fuel ($k_{\text{lig}} \neq 0$, $k_f \neq 0$, $\delta^{\text{I}} = \delta^{\text{II}} = 0$). The evolution of a set of sequences in time is depicted in Fig. 6.4 for both a phase-separated (b) and a mixed system (a). Here, we have only depicted the phase-averaged volume fractions for alternating sequences of different length n_i. The initial condition for both sub-figures is the same, but the non-equilibrium steady state is distinct between the two cases. The phase-separated system reaches a higher average length, but uses a longer time to reach its non-equilibrium steady state. Furthermore, the noise level is larger when phase separated compared to the homogeneous mixed system. When the non-equilibrium steady state is reached, there is a balance between all processes, such that $\mathrm{d}_t \phi_i^{\pi} \simeq 0$ for all components i in both phases $\pi \in \{\text{I}, \text{II}\}$.

In the non-equilibrium steady state, we can compare the sequence fidelity Q_i and fitness A_i between the mixed and the phase-separated system. The average sequence fidelity Q_i (defined in Eq. (6.6)) in Fig. 6.5a decays exponentially for a

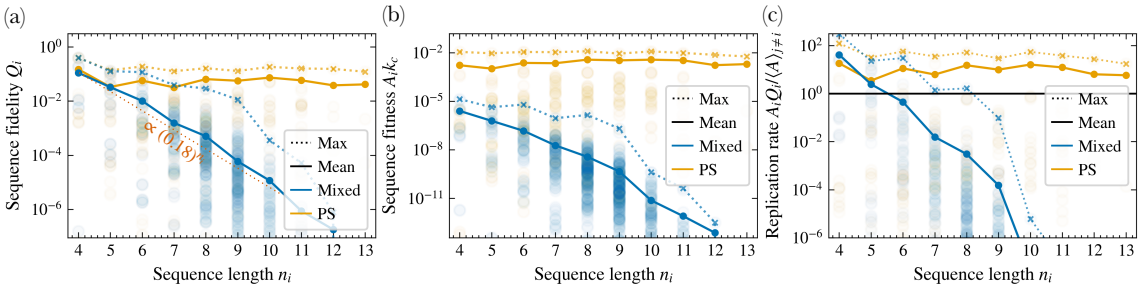


Figure 6.5: **Replication survives the error threshold when phase separated:** Both the average sequence fidelity Q_i in (a) and sequence fitness A_i in (b) decay exponentially with sequence length for homogeneous systems. The decay is similar for both the maximal quantity of the given length (dashed) and the volume-fraction weighted mean (solid). Conversely, for a phase-separated system, both quantities remain approximately flat. Thus, phase separation allows for replication rates, shown in (c), below Eigen's error threshold (solid black line).

mixed system. A replication accuracy of approximately 18% per nucleotide fits the decaying distribution.

For the phase-separated system, on the other hand, the sequence fidelity shows no sign of decay over the length distribution found. The fidelity of most sequences still decays with length, but due to their low concentration, it does not contribute significantly to the average. For the mixed system, even the maximal fidelity (dotted line) decays significantly with length. A similar behavior is observed in Fig. 6.5b, where the average sequence fitness $\langle A_i \rangle$ (defined in Eq. (6.7)) decreases exponentially for the mixed system, while approximately constant when phase separated.

With the fidelity and fitness, we compare how the reproductive rate $A_k Q_k$ compares to Eigen's error threshold as defined in Eq. (6.8). For the mixed system, the error threshold is only surpassed for the smallest templates capable of templated ligation. As the length of the template increases, both the average and maximum replication rate decrease exponentially when mixed. If the system is phase separated, the behavior is completely different. The average replication rate is approximately flat, with the maximal replication rate surviving the error threshold for all lengths.

Templated ligation is a very non-linear process. The ligation rate is cubically proportional to the volume fraction of the three oligomers partaking in the ligation event. Phase separation creates a local environment with a much larger total oligomer volume fraction $\phi_{\text{tot}}^{\text{II}} \gg \bar{\phi}_{\text{tot}}$ than in the mixed system, significantly enhancing the average sequence fitness (as seen in Fig. 6.5(b)). With a higher fitness, more ligation events occur, which biases the sequence distribution towards a specific sequence pattern. A dominating sequence pattern throughout the phase results in a higher fidelity (as seen in Fig. 6.5(a)). These two effects combined result in the replication rate below the error threshold when phase separated, but not when homogeneous.

In Fig. 6.6, the 16 sequences with the largest replication rate above a length of six are displayed and sorted by length. For the mixed system (a), most sequences are below the error threshold. The dominating sequence pattern is an alternating one, but other sequences do appear among the top 16. For the phase-separated system, the alternating sequence pattern almost completely spans the top 16 rank distribution (b). While the mixed system achieves its largest replication rate for an 8-mer, the phase-separated system achieves it for a 10-mer. Furthermore, most sequences are below the error threshold. For both cases, the volume fractions typically decay with oligomer length.

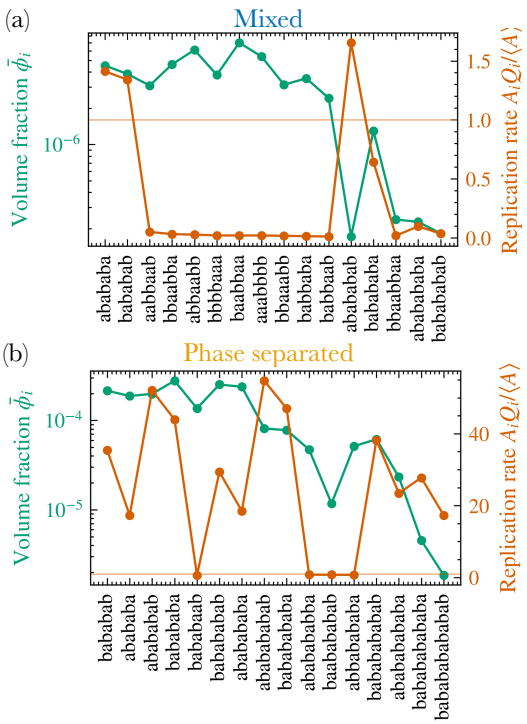


Figure 6.6: **Long and accurate replicators when phase-separated:** The replication rate for the best replicators in a phase-separated (b) compared to a mixed (a) system is orders of magnitude larger, and makes longer sequences more efficient replicators.

We define the master sequence i_m as the oligomer of length 6 or more with the largest replication rate. In Fig. 6.7, the total volume fraction and average replication rate are shown as a function of the Levenshtein distance d to the master sequence. The Levenshtein distance measures the minimum number of changes between two strings, when altering and removing characters [246]. For both systems, a strong decay of the volume fractions with the Levhenstein distance is observed, with a local maximum at a distance of two, $d(2)$. Above a distance of two, the total volume fraction decays exponentially. For the phase separated system, the average relative replication rate remains above 1 as the Levhenstein distance increases.

In addition to phase separation resulting in a sequence distribution closer to the master sequence, it results in much less of the total sequence space being occupied. This is displayed in Fig. 6.8, where a histogram of volume fraction sorted by length is depicted in (a), and the selectivity, defined as the minimal $\#_{90}$ satisfying

$$\frac{\sum_{i\in\sigma}^{\#_{90}(l)} \phi_i \delta_{n_i,l}}{\sum_{i\in\sigma} \phi_i \delta_{n_i,l}} > 0.90\,, \tag{6.14}$$

in (b). The selectivity quantifies the minimal number of sequences needed to account

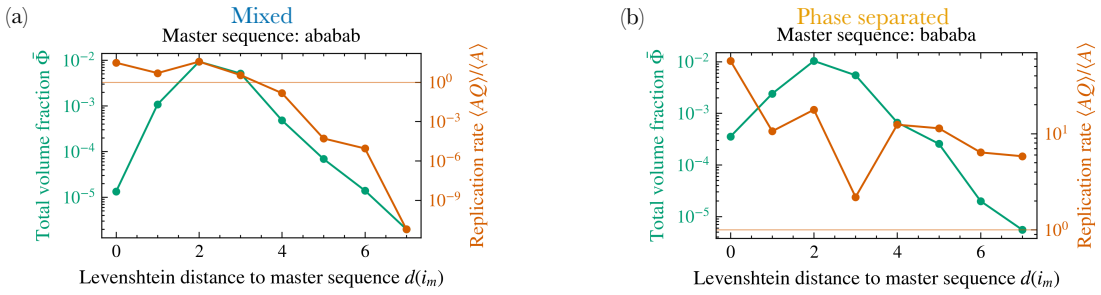


Figure 6.7: **replication decays with distance to master sequence:** The replication rate decays and the total volume fraction increases with the Levenshtein distance to the master sequence (i.e., the largest replication rate) for both mixed (a) and phase-separated (b). For the mixed system, the replication decays more than eight orders of magnitude, while the phase separated remains less than two orders. The master sequence in (a) is only of length four, while in (b) it is length 10, resulting in a different maximum for the total volume fraction. Furthermore, despite the larger average length, all sequences have a shorter distance to the master sequence when phase separated.

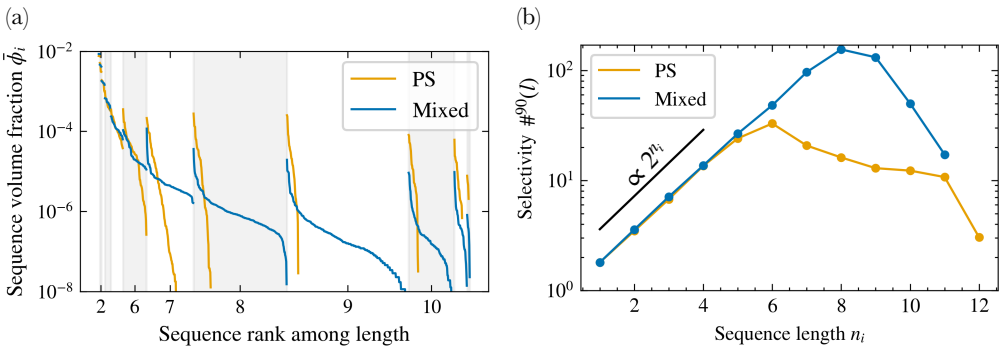


Figure 6.8: **Phase separation narrows the occupied sequence space:** (a) Histogram of volume fractions sorted by oligomer length. (b) The selectivity S defined in (6.14), measures the minimal number of sequences required to account for 90% of the volume fraction at a given length. Under phase separation, sequence abundances decay rapidly with rank (a), leading to smaller selectivity values compared to the mixed system.

for 90% of the volume fraction for the given length. Phase separation results in the volume fractions of each length being large, but decaying rapidly with rank, resulting in an intersection with the mixed system for small sequence ranks for that length. This results in the selectivity deviating from the expected exponential growth when phase separated. Not only does phase separation aid ligation in both its fidelity and fitness, but it also heavily selects a specific sequence pattern that only occupies a small part of the total sequence space. As the oligomer length increases, occupying a small sub-space becomes increasingly important to avoid the dilution of a larger and larger sequence space.

6.5 Discussion and conclusion

In Fig. 6.3, only two different attractors were found in phase-space. The dominant sequence patterns of the attractors were alternating, with the repeating chunk being of length two or four. Other attractors, with other chunk lengths or more complicated patterns, might exist, making the system multistable instead of simply bistable. In order to find other attractors in sequence space, the noise should be increased. Reducing the total number of particles in the system will lead to a larger noise. Alternatively, the average composition $\bar{\phi}_A$ and $\bar{\phi}_B$ can be reduced, in order to reduce the average volume of the condensed phase, leading to larger noise. Orthogonally, by increasing the number of nucleotides, the number of attractors in sequence space would likely increase, and allow for more complicated sequence patterns.

As seen in chapter 5, phase-separation is able to enhance the abundance of specific sequence types by multiple orders of magnitude at conditions just slightly above the saturation concentration. The local oligomer-rich environment of the condensed phase allowed for interactions among sequences to make specific sequence patterns energetically preferential. As a sequence pattern becomes preferential, the sequence distribution changes, and the preference becomes stronger still. This positive feedback led to highly correlated sequence patterns. In this chapter, the additional pathway of templated ligation utilized this correlated sequence pattern to bias the selection further. Templated ligation enhanced the alternating sequence pattern through the thermodynamic base-pair interactions, favoring the hybridization of complementary sequence pairs. Again, a positive feedback loop emerges, as the bias towards a sequence pattern enhances its ligation and hence fidelity. Not only did the condensed phase correlate the sequences through interactions to improve fidelity, but it also further enhanced fitness through increased rates. Templated ligation is a three-body interaction, making the rate proportional to the concentration of the partaking sequences cubed. The pure up-concentration through phase-separation thus enhances the rate of templated ligation more than other pathways. The combined effect of enhanced fidelity and fitness led to the replication rate surviving Eigen's error threshold.

In chapter 5, sequences with different sequence patterns could be enriched depending on how far the system was maintained from equilibrium. Here, the alternating sequence pattern was selected independent of the fragmentation rate for the parameters studied. More interesting sequence patterns might emerge for larger fragmentation rates, or perhaps by different forms of breaking detailed balance. A system that first reaches a non-equilibrium steady state of selected sequence patterns of higher complexity state through fragmentation should then be able to enhance this bias towards the complex sequence motifs through the templated ligation stud-

ied here. More theoretical work is needed to scrutinize this possibility.

A central obstacle for RNA-based replication is the problem of strand separation, which is not included in the presented model. Once complementary strands hybridize into a duplex, they are stabilized by base-pairing and stacking interactions, making spontaneous dissociation energetically unfavorable under mild conditions [204]. Various physical mechanisms have been proposed for strand separation, like temperature cycles [247] and wet-dry cycles [60], which are applicable to abiotic systems. More work is needed to include strand separation and understand its dependence and interplay with phase separation.

Note that Eigen's model [43, 44] is not identical to the one studied here. For one, Eigen considered a pool of sequences all with the same length, obeying a phenomenological kinetic equation under dilute conditions. The system was open, as substrate was continuously supplied to be at a constant excess value, maintaining the system out of equilibrium. Correspondingly, material was removed from the system proportionally to its abundance. Additionally, the mutation rate between each sequence was a chosen input to the model, which makes the master sequence itself an input. Equation (6.8) proposes a threshold for phase-separated mixtures with oligomers of different lengths. Its precise form warrants further theoretical development. Despite a more rigorous extension of Eigen's threshold being needed for complex systems, such as the one studied here, the approximately constant value of A_i, Q_i and $A_iQ_i/\langle A_jQ_j\rangle_{j\neq i}$ as a function of length, is drastically different to the exponential decay for a mixed system (Fig. 6.5). Despite a generalized error threshold being needed, the results still suggest condensates as a viable path to survive the error threshold.

To achieve open-ended evolution, the functionalities of different sequences should be accounted for. Finding possible secondary structure can be done using packages like ViennaRNA [248, 249], which uses a thermodynamic model of nearest-neighbor interactions to find folded RNA structure. Additionally, functions of specific RNA could be included through known tables, like the ones developed by The RNAcentral Consortium [250]. Once functions are included, the system could continue to evolve with time. The current system has a high enough fidelity to avoid information loss through destructive mutations, but it is still unclear whether or not it is adaptive enough to explore a wide chemical landscape. Surviving Eigen's error threshold through compartmentalization provides important steps to achieving open-ended evolution towards life.

Chapter 7

Outlook and conclusion

In the preceding three chapters, we presented the main results of this thesis. In the following, we outline potential directions for future work before concluding on the findings presented.

7.1 Sequence selection through wet-dry cycles

In chapter 3, we saw how wet-dry cycles could enhance the speed of selective reaction pathways depending on the molecular interactions of the substrate. Performing further studies with many components, a more complex reaction network, and higher-order reactions is necessary to scrutinize their effects on more complex chemistry.

One possible case would be to study the interplay between phase separation and oligomerization from chapter 5 and 6 under the influence of wet-dry cycles. In fact, the effects of wet-dry cycles on RNA oligomerization have recently been investigated experimentally [251]. Their work shows that RNA polymerize efficiently at room temperature in alkaline conditions when subject to wet-dry cycles, without the need for activators. Furthermore, different nucleotides are affected differently by the same wet-dry cycles. For example is the yield of guanine the highest, with a propensity to form repeated patterns of guanine. In their work, they explain this observation from guanine's self-assembly propensity increasing as the solvent concentration is reduced. This explanation agrees with the suggested mechanism in chapter 3, as a solvent-dependent internal energy appears as a second-order virial coefficient in the free energy. By performing a Taylor expansion of the internal energy $\omega_i(\phi_0) \approx \omega_i^0 - (1-\phi_0)\,\mathrm{d}\omega_i/\mathrm{d}\phi_0$, we see that its solvent dependence would enter the free energy in the second order. Thus, a solvent-dependent internal energy will affect the sequence kinetics equivalently to a χ_{ij}, which were found to be the important parameters to achieve speedup in chapter 3.

A theoretical study of how different sequence patterns can be selected for depending on the solvent interactions and oscillation frequency of the wet-dry cycles would give further insight. Additionally, the total amount of each nucleotide in the system could be varied in order to balance the enhanced reactivity of one of the nucleotides during cycling.

7.2 Excitable media without membranes

In extant life, excitability is used to transmit, amplify, and coordinate signals reliably across space and time. Excitable waves propagate information rapidly and over long distances without decaying, enabling many cells or tissues to act in unison. The textbook example is the neuron, where small inputs decay unless a threshold is crossed, in which case a full action potential is triggered, followed by a refractory period before the cell can fire again [252]. Excitability of neurons emerges from ion gradients across the plasma membrane, which are actively maintained by ATP-driven ion pumps [253].

By contrast, biomolecular condensates lack membranes, making the maintenance of potential differences challenging. Whether excitable droplets, such as those investigated in chapter 4, can infact behave similar to other excitable media remains an open question. One promising case may be the arrested ripening regime described in Section 4.7, where a non-equilibrium steady state could serve as an initial state for excitable behavior. In such a regime, a local perturbation might trigger the dissolution of a droplet, potentially initiating a traveling wave of droplet dissolution. This would allow membrane-less droplets to effectively "charge" a potential and release it upon a sufficiently strong perturbation, thereby transmitting information over long distances through a rapid, non-decaying wave. A traveling wave was not observed upon a perturbation in Fig. 4.13, but perhaps could occur for other perturbation mechanisms or different parameter regimes. Further theoretical work is needed to determine the parameter regimes in which such behavior could emerge.

Experimental investigation will be essential to test whether such excitable dynamics is possible through condensates. The general mechanism is a phase-dependent fuel aiding the creation of a product that phase separates stronger than the substrate. Possible candidates are post-translational modifications such as arginine methylation or phosphorylation [200], or by synthetic fuel-driven approaches including peptide phosphorylation and condensing reactions [201, 202, 96, 203]. The second reaction does not have to be identical to the one presented in chapter 4. Its main role is to break the conservation law of the bi-stable reaction and create an inert product. One example would be the transition from A and B being the unfolding of a polymer, as for example studied in Ref [97], while the second reaction

is the formation of two duplexes from the unfolded polymers $2B \rightleftharpoons C$. Additionally, the necessity of component B to phase separate stronger than A can be achieved by B having a larger molecular volume, as seen in section 4.3.2.

7.3 Advancing phase-separated sequence kinetics

In chapter 5 and 6, a Flory-Huggins free energy density was used to describe the oligomers. To improve the description of the free energy, one could go to higher-order virial coefficients proportional to $\phi_i\phi_j\phi_k$, thus accounting for third-order interactions. Third-order virial coefficients do not qualitatively change the behavior of the system, but can allow for multiphase coexistence [254]. An orthogonal approach is to refine the parameterization of the sequence–sequence interaction strength χ_{ij}.

One approach is to increase the accuracy of the sequence-sequence interactions through Wertheim theory [255, 256], which is a statistical mechanics framework to describe molecules with specific and directional binding sites, allowing it to accurately model sequence-specific interactions. The association between components is described by a Helmholtz free energy for interactions

$$\frac{F_{\text{inter}}}{V\, k_B T} = \sum_{i=0}^{M} n_i \sum_{\alpha \in i} \left(\ln X_{i\alpha} + \frac{1 - X_{i\alpha}}{2} \right) , \tag{7.1}$$

where i runs over the $M+1$ particles in the system of volume V, with k_BT being the thermal energy. The new quantity compared to our description is the quantity $X_{i\alpha}$, which is the fraction of site α along component i, which is unbounded. Wertheim derived an expression for the site fraction

$$X_{i\alpha}^{\text{equil}} = \frac{1}{1 + \sum_{j=0}^{M} n_j \sum_{\beta \in j} X_{j\beta} \Delta_{i\alpha,j\beta}} , \tag{7.2}$$

where $\Delta_{i\alpha,j\beta}$ denotes the strength of interaction between site α on component i and site β on component j [257]. This coupled set of equations can be solved for all $X_{i\alpha}$ self-consistently. Note that the above expression is for the equilibrium site fraction. For relaxation kinetics or non-equilibrium systems, the site fractions can be dynamic quantities obeying equations similar to the law of mass action, as has been done for patchy particles [258].

The free energy of interactions in Eq. (7.1) can be combined with the mixing entropy and internal energies of the Flory-Huggins free energy in Eq. (2.12), as in Ref. [259], to study the interplay between phase separation and accurately account for oligomer-oligomer interactions. The increased complexity of the interactions might allow for more accurate predictions when compared with experimental

measurements.

Another avenue of theoretical improvements is to efficiently deal with complexes. For a system of N components, the number of duplexes is N^2, and if accounting for triplexes, the number is even larger. A possible method for efficiently treating complexes is to regard them as dilute clients that do not lead to phase separation. With this approximation, the Client-Scaffold framework [120] can be employed. In this framework, components are separated into clients and scaffolds depending on their abundance and interaction strengths. Scaffolds make up the condensed phase, while clients only partition without any effective interactions. Extending their framework to allow for transitions between clients and scaffolds could prove a powerful and numerically efficient way to treat weakly interacting components like duplexes.

The study of polymer behavior has a history approaching a century [260], and in recent years, advances in computational power and efficient algorithms have fundamentally transformed approaches in the field, which could be applied in order to enhance the accuracy of sequence evolution and sequence distribution. Mainly, the rudimentary method we used for calculating sequence-sequence interactions should be improved, and secondary structure should be accounted for.

The Nucleic acid package (NUPAC) is based on a thermodynamic model with experimentally measured values for neighbor and nearest-neighbor energy rules for RNA/DNA hybridization, allowing it to predict minimum free energy structures and equilibrium properties of multiple strands. Finding possible secondary structure can be done using packages like ViennaRNA [248, 249], which uses a thermodynamic model of nearest-neighbor interactions for RNA folding. While also being able to predict secondary structure, Multistrand [261, 262] gives insight into how base pairs form and break over time, while accounting for other quantities like salt dependence. The aforementioned packages work by stochastically evolving thermodynamical models fit to experimental measurements. Other orthogonal approaches exist, such as coarse-grained molecular dynamics in OxDNA [263], or the growing approach of AI methods trained on massive experimental data-sets such as EternaFold [264] or SPOT-RNA [265]. Using these packages to determine secondary structure and interaction strengths, χ_{ij}, would greatly enhance the application of our framework to real systems.

In previous work on the interplay between phase separation and sequence evolution [266, 267], each sequence is not typically individually accounted for, as we have done in Chapter 5 and 6. Accounting for each sequence comes with clear limitations, as the number of components for n different nucleotide-types scales with the max-length l as n^{l+1}. The stochastic approach in chapter 6 dealt with this issue by investigating a system of few particles, resulting in most long-length sequences

never existing. An alternative approach is to account for all sequences efficiently through machine learning.

It has been shown that many-body quantum systems, with exponentially high-dimensional wave functions, can be effectively solved using artificial neural networks [268]. A promising approach to high-dimensional phase spaces of sequences might be made by utilizing the fact that the phase space of a set of n-mers of m nucleotides is equivalent to the phase space of n quantum particles of spin $(m-1)/2$. The method in Refs. [269, 270] to evolve the quantum spin-system in such a phase space is based on three key principles. The first is that the state of the system can be represented by an artificial neural network that uses far fewer parameters than the total number of variables in the system. The second is the time-dependent variational principle, which allows one to derive dynamical equations for the artificial neural network parameters from the system's kinetic equations [271]. The third replaces all sums over the system variables with Monte Carlo averages. This last step provides a large speedup in integration compared to solving the full set of coupled dynamical equations. These key principles likely could be applied to our system of sequences. An implementation might be possible using the jVMC library [272]. Machine learning has proven to be a powerful tool for addressing high-dimensional quantum systems [273, 274, 275, 268]. An open question is whether these approaches can also be extended to high-dimensional sequence systems.

A successful implementation of this method could be combined with the improved interaction matrix mentioned before, allowing the framework to accurately deal with phase-separating oligomers in a sequence space of size $\sim 10^{30}$. Accounting for secondary structure requires further theoretical understanding of the interplay between the secondary structure and the condensed phase [276, 277].

7.4 Further theoretical steps towards de novo life

In chapter 5, we saw how condensed phases create localized environments with correlated sequence motifs. The environments allow for a strong selection of specific sequence patterns compared to a homogeneous mixture. As the degree to which the system was maintained away from equilibrium was changed, the enriched sequence pattern changed with it. The selection led to only a small subset of the entire sequence space being occupied, even at overall dilute conditions. In chapter 6, we continued with the system, studying in part the role of stochasticity. Through the stochastic evolution of the oligomers, different non-equilibrium steady states were reached from the same initial condition. The different steady states selected for different sequence patterns, which in turn led to different properties of the mesoscopic condensate. This relationship between sequence distributions within the condensate

and droplet properties provides a simple phenotype-genotype relationship, where a selection mechanism on the genotypic level affects which sequences survive. Furthermore, we found that the replication rate of sequences through templated ligation was orders of magnitude more accurate than for a homogeneous system. The enhanced replication rate was large enough to create a set of long sequences that survived Eigen's error threshold.

Many theoretical challenges remain to understand the emergence of de novo life. An essential remaining question is whether the selected oligomers found in chapter 5 and 6 have distinct chemical properties like catalysis. In order to investigate this, the possible secondary structure of sequences and their functionalities should be known. By using packages as mentioned in the previous section (section 7.3), a table of functional sequences and their function can be included in the model. If such a sequence is produced throughout the stochastic trajectories, its functions can be kinetically implemented in the algorithm. Over time, this could allow for a continuous evolution in sequence space, with the creation of new sequences and functions over time. Importantly, such a framework could not rely on keeping track of all possible states simultaneously. A Gillespie-type framework of a dynamic sequence space should be employed in order for it to be a numerically feasible problem. The inclusion of folding, sequence function, duplexes, and their interplay with phase separation requires further theoretical modeling.

Every coarse-grained theory rests on approximations and assumptions about chemical pathways, kinetic rates, interaction strengths, and possible phase transitions that must ultimately be tested against experiments. Robust predictions about a system's long-term evolution can only emerge from repeated feedback between theory and experiment, with theory helping to steer experiments toward the conditions most suitable for the emergence of functions such as replication.

7.5 Conclusion

Being able to create life de novo would be a revolutionary event in science. It would grant a new understanding of living matter and help to find general features and functions of living systems, with potentially revolutionary solutions to challenging problems. Furthermore, it would create a unique platform to study how complexity and evolution emerge from first principles, bridging physics, chemistry, and biology.

A central challenge in the creation of de novo life is to understand how complex collective behaviors can emerge from the reactions and interactions among simple molecular components. In their path towards life, chemical systems must be capable of overcoming key challenges. To mention some, chemical systems need to accelerate reaction pathways that are otherwise thermodynamically unfavorable, regulate

their dynamics in a way that ensures robust responses to external perturbations, enrich and stabilize oligomers against degradation, and maintain sequence information with efficient copying. Addressing how these behaviors could emerge requires understanding the physical mechanisms and non-equilibrium conditions that support them. In this thesis, we suggest that the interplay between chemical reactions and phase separation in condensates, when driven away from equilibrium, provides a promising environment for meeting these challenges. By exploring how condensates affect reaction kinetics, transport, and information storage, we have aimed to understand how emergent behavior toward de novo life can emerge from simple mixtures.

To explore this idea, we employed non-equilibrium thermodynamics and statistical mechanics to make models of non-dilute, interacting mixtures away from equilibrium. Within these models, we investigated how the presence of condensed phases modifies reaction and transport kinetics, and whether condensates can facilitate emergent behavior.

In an abiotic setting, enzymes are not present to aid the selection and acceleration of thermodynamically slow reaction pathways. In chapter 3, we developed a theoretical framework for non-dilute chemical reactions subject to nonequilibrium conditions of evaporation and condensation, finding that exchanging material with an external environment could both speed up and select reaction pathways. The speedup originates from interactions between the components of the mixture, where the solvent interaction was found to be the dominating parameter to determine the degree of speedup. With repeated cycling, the system entered an orbit of continuous chemical turnover, where a specific oscillation frequency maximizes the turnover. The existence of an optimal cycle frequency was robust to cycles where components directly participated in the chemical reaction were exchanged. Surprisingly, for fast exchange with the reservoir, wet-dry cycles could be even more effective at harvesting energy from the environment than when exchanging reactive components. The ability of external reservoirs to accelerate specific reaction pathways suggests a potential mechanism to initiate chemical evolution in abiotic systems.

In living cells, cycles of formation and dissolution of liquid droplets can mediate biological functions such as DNA repair. In chapter 4, we presented a simple model composed of only two independent chemical components with their diffusive and chemical fluxes governed by non-equilibrium thermodynamics. Through the turnover of a fuel, which maintains a chemical reaction away from equilibrium, the droplet becomes active. We find that a single active droplet undergoes a saddle-node bifurcation in the droplet volume upon increasing the fueling strength. The requirement of bistability is that the fuel aids the production of a component that phase separates stronger than the substrate. We demonstrated this bistability for a mini-

mal system and for a system of polymers of different lengths. Strikingly, the active droplet becomes excitable upon adding a further chemical reaction, allowing for an excited trajectory upon a perturbation. For sufficient fueling, the system undergoes self-sustained oscillations cycling between droplet formation and dissolution. Lastly, we showed that the self-sustaining cycles also exist for an emulsion of droplets, for a particular parameter regime. In an emulsion, the coarsening of the droplets was either arrested or led to oscillations that could synchronize in space. The minimal nature of our model suggests self-sustained active droplets as functional modules for de novo life.

Non-equilibrium selection pressures have been proposed for the formation of oligonucleotides with rich functionalities encoded in their sequences, such as catalysis. Since phase separation was shown to direct various chemical processes, we asked in chapter 5 whether condensed phases can provide mechanisms for sequence selection. To answer this question, we used non-equilibrium thermodynamics and describe the reversible oligomerization at non-dilute conditions prone to phase separation. We find that when sequences oligomerize, their interactions give rise to phase separation, leading to enrichment and depletion of specific sequences. Our key result is that phase separation gives rise to a selection pressure for the oligomerization of specific sequence patterns when fragmentation maintains the system away from equilibrium. Specifically, slow fragmentation favors alternating sequences that interact well with their environment, while fast fragmentation selects sequences with extended motifs. Our findings highlight that out-of-equilibrium condensed phases could provide localization and strong selection pressures, both relevant for de novo life.

A necessity for evolution is an accurate copying mechanism of genotypic information. In chapter 6, we investigated sequence evolution in the presence of noise in phase-separated systems. Through noise, the system could move between different attractors in phase space from the same initial condition. The different steady states led to different droplet properties, providing a simple genotype-phenotype relationship. Through templated ligation, we find that a phase-separated system is able to copy sequences with an accuracy and a rate below Eigen's error threshold. If the same system were to remain mixed, both the fidelity and fitness are orders of magnitude lower than their phase-separated counterpart, not enabling information copying below Eigen's error threshold.

Taken together, these findings point to condensation as a fundamental physical mechanism that can enable emergent behaviors, such as excitability, selectivity, and information maintenance, which may be crucial on the path toward de novo life.

Chapter 8

Appendices

A Flory-Huggins free energy

An incompressible binary mixture of particles A and B can be described as particles on a lattice. Each lattice site i is either occupied by a particle of type A, $\sigma_i = 1$, or of type B, $\sigma_i = 0$. A particular arrangement of particles $\{\sigma\}$ has a corresponding energy captured by the Hamiltonian $H(\{\sigma\})$. If energy can be exchanged between the mixture and an external heat bath of temperature T, thermodynamics tells us that the system is dictated by its partition function

$$Z = \sum_{\Sigma} \exp\left\{\left(-\frac{H(\{\sigma\})}{k_B T}\right)\right\}, \tag{A.1}$$

where k_B denotes Boltzmann's constant and Σ is the space of all possible lattice configurations, assuming each particle is indistinguishable from other particles of the same type. The energy of the system can be captured by interaction energies e_{ij} between particles on neighbouring lattice nodes:

$$H(\{\sigma\}) = \frac{1}{2}\sum_{i}^{M}\sum_{j\in n(i)} \left(e_{AA}\sigma_i\sigma_j + e_{BB}\left(1-\sigma_i\right)\left(1-\sigma_j\right)\right) + e_{AB}\left[\sigma_i\left(1-\sigma_j\right) + \sigma_j\left(1-\sigma_i\right)\right], \tag{A.2}$$

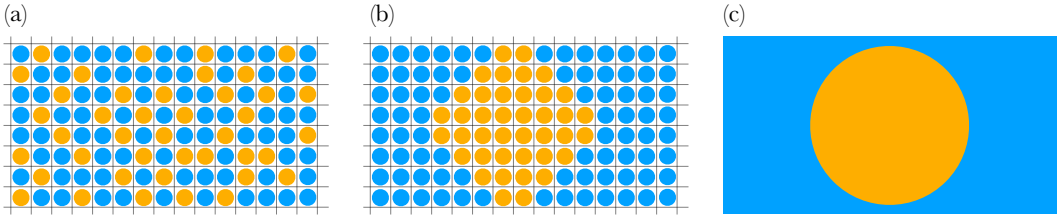


Figure A.1: **Coarse-graining of a lattice model:** An illustration of a lattice model in a homogeneous state (a), and a phase-separated state (b). Performing the continuum mean-field limit results in a profile similar to the one in (c).

where $n(i)$ is set of neighbours of lattice site i. The interaction strengths e_{ij} can originate from van der Waals, dipolar, or screened electrostatic interactions. For this Hamiltonian, the partition function in Eq. (A.1) is difficult to find. Neglecting spatial correlations between molecules, by assuming that the probability ϕ_A of an arbitrary lattice site is of particle type A is equal to the average composition of A in the system, i.e. $\phi_A = N_A/(N_A + N_M)$, the Hamiltonian simplifies to

$$H(\{\sigma\}) = \frac{zM}{2}\left(e_{AA}\phi_A^2 + e_{BB}\left(1-\phi_A\right)\left(1-\phi_A\right) + 2e_{AB}\phi_A\left(1-\phi_A\right)\right) , \tag{A.3}$$

where M is the total number of lattice sites $M = N_A + N_B$ and z is the number of neighbours per lattice site. The number of arrangements Ω for each combination of N_A and N_B is equal to M choose N_A, which sets the entropy $S = k_B \log|\Omega|$.

The Helmholtz free energy F, quantifying the difference between energy E and entropy TS, $F = E - TS$, is the negative logarithm of the partition function, such that

$$\begin{aligned} F = {} & k_B T\left[\phi\log\phi + (1-\phi)\log\left(1-\phi\right)\right] \\ & + \frac{z}{2}\left(e_{AA}\phi_A^2 + e_{BB}\left(1-\phi_A\right)\left(1-\phi_A\right) + 2e_{AB}\phi_A\left(1-\phi_A\right)\right) . \end{aligned} \tag{A.4}$$

The two terms represent the mixing entropy, while the latter term represents interaction energies between the components. By defining ν as the volume of a lattice site, we find the Helmholtz free energy density $f = F/V$, where the volume $V = M\nu$, to be

$$f = \frac{k_B T}{\nu}\left[\phi\log\phi + (1-\phi)\log\left(1-\phi\right) + \chi\phi\left(1-\phi\right) + \omega\phi\right] . \tag{A.5}$$

This way of writing the free energy density emphasizes the competition between entropic mixing through the logarithms and the interactions captured by the Flory-Huggins interaction parameter χ. All linear terms have been absorbed in the internal energy ω, and we have defined the Flory-Huggins interaction parameter, as defined in Eq. (2.13).

Generalizing to N components [87], the Flory-Huggins free energy takes the form given in Eq. (2.12). The ρ_i in the dominator of the argument of the logarithm is optional, as it can equivalently be included in the internal energies ω_i. As this derivation is with a fixed number of lattice sites and particles, it describes an incompressible system in the Helmholtz ensemble.

B Deriving the exchange flux in the fast diffusion limit

Here, we follow the work of Bauermann et al. [113] to derive the expression for the diffusive flux between the two phases, $j_i^{\mathrm{I/II}}$ in Eq. (2.60).

For two phases to remain at phase coexistence, the chemical potentials of all $M+1$ components in the two phases evolve equally in time[1],

$$\partial_t \mu_i^{\mathrm{I}} = \partial_t \mu_i^{\mathrm{II}} \,, \tag{B.1}$$

Using the chain rule, the time derivative of a chemical potential can be written as

$$\partial_t \mu_i^{\alpha} = \sum_{k=1}^{M} \dot{\phi}_k^{\alpha} \frac{\partial \mu_i^{\alpha}}{\partial \phi_k} \,, \tag{B.2}$$

where α is the phase index; $\alpha \in \{\mathrm{I}, \mathrm{II}\}$. Combining Eq. (2.60a) with the expression for the volume change of the phases (Eq. (2.60b)) for volume conserving chemical reaction ($\sum_{i=1}^{M} r_i^{\mathrm{I/II}} = 0$), the chemical potential change in time from Eq. (B.2) can be written as

$$\partial_t \mu_i^{\alpha} = \sum_{k=1}^{M} \left(r_k^{\alpha} - j_k^{\alpha} + \phi_k^{\alpha} \sum_{l=1}^{M} j_l^{\alpha} \right) \frac{\partial \mu_i^{\alpha}}{\partial \phi_k} \,. \tag{B.3}$$

The condition that the change in chemical potentials are equal in both phases (Eq. (B.1)) thus reads

$$\begin{aligned} \sum_{k=1}^{M} j_k^{\mathrm{I}} \Bigg\{ \sum_{l=1}^{M} \phi_l^{\mathrm{I}} \frac{\partial \mu_i^{\mathrm{I}}}{\partial \phi_l^{\mathrm{I}}} - \frac{\partial \mu_i^{\mathrm{I}}}{\partial \phi_k^{\mathrm{I}}} + \frac{V^{\mathrm{I}}}{V^{\mathrm{II}}} \left(\sum_{l=1}^{M} \phi_l^{\mathrm{II}} \frac{\partial \mu_i^{\mathrm{II}}}{\partial \phi_l^{\mathrm{II}}} - \frac{\partial \mu_i^{\mathrm{II}}}{\partial \phi_k^{\mathrm{II}}} \right) \Bigg\} \\ = - \sum_{k=1}^{M} \frac{\partial \mu_i^{\mathrm{I}}}{\partial \phi_k^{\mathrm{I}}} r_k^{\mathrm{I}} + \sum_{k=1}^{M} \frac{\partial \mu_i^{\mathrm{II}}}{\partial \phi_k^{\mathrm{II}}} r_k^{\mathrm{II}} \,, \end{aligned} \tag{B.4}$$

where we have used that particle exchange between phases conserves the number of particles,

$$j_i^{\mathrm{I}} V^{\mathrm{I}} = -j_i^{\mathrm{II}} V^{\mathrm{II}} \,. \tag{B.5}$$

This set of equations (B.4) can be written as a matrix-vector equation

$$A_{ik} j_k^{\mathrm{I}} = s_i, \tag{B.6}$$

which can be solved for j_k^{I}. The reaction rates in each phase act as source terms

[1] In the Helmholtz ensemble, it is the exchange chemical potential of the M-non solvent components and the osmotic pressure that evolves equally in both phases, that is $\partial_t \mu_i^{\mathrm{I}} = \mu_i^{\mathrm{II}}$ for $i = 1, \ldots, M$ and $\partial_t \Pi^{\mathrm{I}} = \partial_t \Pi^{\mathrm{II}}$. The derivation will be similar, see for example Ref. [1, 4]

in s_i, while A_{ik} specifies how the composition changes in each phase will affect the chemical potentials, constraining the diffusive fluxes between the phases to equalize the chemical potential change for all components in both phases.

C Reaction scheme with an explicit fuel component

Detailed balance is broken by a fuel $\tilde{\delta}$ according to Eq. (4.2). Breaking detailed balance in this fashion is equivalent to the combined effect of the two reactions:

$$A \overset{k_u}{\rightleftharpoons} B\,, \qquad A + F \overset{k_\mathrm{f}}{\rightharpoonup} B\,, \tag{C.1}$$

where the former reaction obeys detailed balance of the rates, while the latter breaks it. The reaction rates of the unfueled (r_u) and fueled reaction (r_f) are

$$r_\mathrm{u} = k_u \left[\exp\left\{\frac{\mu_A}{k_B T}\right\} - \exp\left\{\frac{\mu_B}{k_B T}\right\}\right], \tag{C.2}$$

$$r_\mathrm{f} = k_\mathrm{f} \exp\left\{\frac{\mu_A + \mu_F}{k_B T}\right\}, \tag{C.3}$$

such that $r_B = -r_A = r_\mathrm{u} + r_\mathrm{f}$. Performing the mapping between these rates and the coarse-grained rate studied (Eq. (4.3)), we find

$$\tilde{\delta} = \log\left(1 + \frac{k_\mathrm{f}}{k_u} \exp\left\{\frac{\mu_F}{k_B T}\right\}\right), \tag{C.4}$$

where μ_F is the chemical potential of the fuel. Solving for the reaction rate coefficient of the fueled reaction k_f, we find

$$k_\mathrm{f} = k_u \exp\left\{-\frac{\mu_F}{k_B T}\right\}\left(\exp\left\{\frac{\tilde{\delta}}{k_B T}\right\} - 1\right), \tag{C.5}$$

with $k_u = k_1$. Thus, the fuel's chemical potential can be regarded as phase-independent, meaning it satisfies the conditions of phase coexistence, and the phase-dependent fuel strength originates from the composition dependence of the rate coefficient k_f. For the studied case of $\tilde{\delta}^\mathrm{I} = 0$ and $\tilde{\delta}^\mathrm{II} = \delta$, this is equivalent to the fueled reaction having a reaction rate of zero in the dilute phase, and a non-zero value according to Eq. (C.5) in the dense phase.

D Deriving the normal for of bistable droplets

For the single reaction r_1, defined in Eq. (4.2), the evolution of the system is efficiently described by the dynamics of two fields; the reaction extent $\bar{\xi} \equiv (\bar{\phi}_B - \bar{\phi}_A)/2$ and the total volume fraction $\bar{\psi} \equiv (\bar{\phi}_A - \bar{\phi}_B)/2$. With this choice of coordinates, $\bar{\psi}$ is a constant $\bar{\psi}_0$, which is set by the initial conditions, while the change in $\bar{\xi}$ is equal to the phase-averaged reaction rate

$$\mathrm{d}_t \bar{\xi} = \bar{r}_1(\bar{\xi}, \bar{\psi}_0)\,. \tag{D.1}$$

Here, we will expand the reaction rate at two different ξ-values, one where the system is homogeneous and one where it is phase-separated. By connecting the two cases, we find a piece-wise function that captures the normal form of bistable droplets given in Eq. (4.13), and displayed in Fig. 4.3(b).

Let us start with the homogeneous case. The fluxes around the homogeneous chemical equilibrium point, denoted $(\bar{\xi}_0, \bar{\psi}_0)$, is approximately linear,

$$\mathrm{d}_t \bar{\xi} = -\left(\bar{\xi} - \bar{\xi}_0\right) h_1(\bar{\xi}_0, \bar{\psi}_0)\,, \tag{D.2}$$

where $h_1 > 0$ makes $\bar{\xi}_0$ a stable steady state. The first order coefficient of the Taylor expansion of $\bar{r}_1$, denoted h_1, takes the form

$$h_1 = r_1^{\rightharpoonup} \left[\frac{2\bar{\psi}_0}{\bar{\psi}_0^2 - \bar{\xi}_0^2} - \frac{\partial}{\partial \bar{\xi}} \ln\left(\frac{\gamma_A}{\gamma_B}\right)\bigg|_{\bar{\xi}_0, \bar{\psi}_0} \right]\,, \tag{D.3}$$

where we have decomposed the chemical potential in terms of the reference chemical potentials μ_i^0 and the activity coefficients γ_i, as defined through Eq. (2.16). The chemical equilibrium point $(\bar{\xi}_0, \bar{\psi}_0)$ around which we have expanded, defined by $\mu_A = \mu_B$, is given by

$$\frac{\bar{\xi}_0}{\bar{\psi}_0} = 1 - \frac{2}{1 + \frac{\gamma_A}{\gamma_B} \exp\left\{ \frac{\mu_A^0 + \mu_B^0}{k_B T} \right\}}\,. \tag{D.4}$$

The chemical forward and backward rates as used in Eq. (D.3), $r_1^{\rightharpoonup}$ and $r_1^{\leftharpoondown}$ respectively, are by definition positive

$$r_1^{\rightharpoonup} \equiv k_1 \exp\left\{ \frac{\mu_A}{k_B T} \right\}, \qquad r_1^{\leftharpoondown} \equiv k_1 \exp\left\{ \frac{\mu_B}{k_B T} \right\}. \tag{D.5}$$

Furthermore, the first term in the square bracket of Eq. (D.3) is, by the definitions of $\bar{\psi}$ and $\bar{\xi}$, positive, while the second term can take both signs. Using our form of

the chemical potential

$$\mu_i = k_B T \left(\log \phi_i + 1\right) + \omega_i + \nu_i \left(p - \Gamma\right) + \frac{\nu_i}{\nu_0} \sum_j \chi_{ij} \phi_j, \tag{D.6}$$

the full expression simplifies to

$$h_1 = r_1^{\rightharpoonup} \left[\frac{2\bar{\psi}_0}{\bar{\psi}_0^2 - \bar{\xi}_0^2} - 2\frac{\chi_{AB}}{k_B T} \right], \tag{D.7}$$

independently of the number of components.

We now perform a Taylor expansion of $\bar{r}_1$ for the case where the system is phase-separated. For phase-separated systems, changes in $\bar{\xi}$ now result in changes in the phase volumes and the composition of the phases. The local maxima in the phase-separated domain in Fig. 4.3(a), denoted as $(\bar{\xi}_2, \bar{\psi}_0)$, can be approximated as

$$\partial_t \bar{\xi} = h_0 - \left(\bar{\xi} - \bar{\xi}_2\right)^2 h_2(\bar{\xi}_2, \bar{\psi}_0), \tag{D.8}$$

where h_0 is a constant given by $h_0 = \bar{r}_1(\bar{\xi}_2, \bar{\psi}_0)$, and h_2 is the negative of the second order coefficient of the Taylor expansion, i.e., $h_2 = -\partial_{\bar{\xi}}^2 \bar{r_1}\big|_{\bar{\xi}_2, \bar{\psi}_0}/2$, is given by

$$2h_2 = -\frac{\partial^2}{\partial \bar{\xi}^2} \left(r_1^{\rightharpoonup} - r_1^{\leftharpoondown}\right) - \left[\exp\left\{\frac{\delta}{k_B T}\right\} - 1\right] \frac{\partial^2}{\partial \bar{\xi}^2} \left(\frac{V^{\mathrm{II}} r_1^{\rightharpoonup}}{V}\right). \tag{D.9}$$

Notably, the composition in each phase ($\phi_i^{\mathrm{I/II}}$) have a non-linear transcendental dependence on the phase-averaged quantities $(\bar{\xi}, \bar{\psi})$ which must be known to perform the derivatives in Eq. (D.9). Further note that the conditions of phase coexistence (Eq. (2.29)) allow us to omit the phase index of the forward and backward reaction rates (defined in Eq. (D.5)). All non-equilibrium effects are captured by the second line of Eq. (D.9), where volume changes are irrelevant for equilibrium systems ($\delta = 0$).

For bistability to occur h_0, as defined in Eq. (4.14), must be positive. Remember that both the forward and backward reaction rates are positive by definition (Eq. (D.5)). For a sufficiently strong fuel strength δ, the coefficient h_0 can become positive, as we see in Eq. (4.14).

By ensuring continuity of $\partial_t \bar{\xi}$ at the binodal $\bar{\xi} = \xi^{\mathrm{I}}(\bar{\psi}_0)$, we can find an expression for h_2 in terms of h_0 and h_1, yielding

$$h_2 = \frac{h_1 \left(\xi^{\mathrm{I}} - \xi_0\right) + h_0}{\left(\xi^{\mathrm{I}} - \bar{\xi}_2\right)^2}. \tag{D.10}$$

This alters the curvature at the expansion point $(\bar{\xi}_2, \bar{\psi}_2)$ while keeping the constant

offset fixed. The choice of expressing h_2 in terms of the other coefficients is done to reduce the complexity of the normal form, while still capturing the essential role of the sign of h_0.

From the two expressions, we can write a piece-wise function for $\partial_t\bar{\xi}$ as given in the main text in Eq. (4.13). Usually, the normal form of bistable chemical reactions in homogeneous systems originates from a cubic relationship between the reaction extent and its time derivative [184]. In contrast, for phase-separated systems, the minimal prerequisite of bistability is a quadratic relationship between ξ and its time derivative. By utilizing the different behavior of homogeneous and phase-separated systems, bistability can occur for simpler reaction schemes. In the following Appendix, we perform the same expansion for systems where $\bar{\psi}$ is dynamic.

E Deriving the normal form of excited droplets

We follow a similar procedure as in Appendix D to describe an excitable system. The key difference is a further chemical reaction with the rate r_2, defined in Eq. (4.28), which makes $\bar{\psi}$ change in time according to

$$\mathrm{d}_t\bar{\psi} = \bar{r}_2\,, \tag{E.1}$$

while the kinetic equation of $\bar{\xi}$ in Eq. (D.1) remains unchanged. The procedure of using piece-wise functions by treating the homogeneous and phase-separated cases separately will still be followed. We begin with the homogeneous case.

As the additional reaction $\bar{r}_2$ obeys detailed balance with phase-independent rate coefficients ($k_2^{\mathrm{I}} = k_2^{\mathrm{II}}$), the reaction rate r_2 is equal in both phases. This makes changes in phase volumes not affect the phase-averaged reaction rates, such that the change in phase-averaged reaction rate $\partial_{\bar{\psi}}\bar{r}_2$ is equal to the reaction rate change in either phase $\partial_{\bar{\psi}}\bar{r}_2 = \partial_{\bar{\psi}}r_2^{\mathrm{I}} = \partial_{\bar{\psi}}r_2^{\mathrm{II}}$. As a consequence, its expansion can approximately be made independently of phase separation, where we find

$$\mathrm{d}_t\bar{\psi} = -\left(\bar{\xi}-\bar{\xi}_1\right) f_\xi\left(\bar{\psi}_1,\bar{\xi}_1\right) - \left(\bar{\psi}-\bar{\psi}_1\right) f_\psi(\bar{\psi}_1,\bar{\xi}_1)\,. \tag{E.2}$$

The expansion point $(\bar{\psi}_1,\bar{\xi}_1)$ is chosen as the unstable steady state point in Fig. 4.7(a), allowing us to set the zeroth order contribution to zero. Furthermore, the expansion is stable to changes in $\bar{\psi}$ ($f_\psi > 0$), where we have

$$\begin{aligned}
\frac{f_\psi}{2} &= \vec{\bar{r}_2}\left(\frac{2\bar{\xi}_1^2-\bar{\psi}_1}{\left(\bar{\psi}_1^2-\bar{\xi}_1^2\right)\left(2\bar{\psi}_1-1\right)} + \frac{\chi_{AB}-2\chi_{BC}-2\chi_{AC}}{k_BT}\right),\\
\frac{f_\xi}{2} &= \vec{\bar{r}_2}\left(\frac{\bar{\xi}_1}{\bar{\xi}_1^2-\bar{\psi}_1^2} + \frac{\chi_{AC}-\chi_{BC}}{k_BT}\right),
\end{aligned} \tag{E.3}$$

where the first term of f_ψ is always positive. The sign of f_ξ can be both positive and negative. Performing a Taylor expansion of $\bar{r}_1$ now requires an expansion in both variables $\bar{\xi}$ and $\bar{\psi}$. The homogeneous expansion becomes

$$\partial_t \bar{\xi} = -\left(\bar{\xi} - \bar{\xi}_0\right) h_\xi - \left(\bar{\psi} - \bar{\psi}_0\right) h_\psi \,. \tag{E.4}$$

Here, h_ξ takes the same form as in Eq. (D.7), and h_ψ can take both signs

$$h_\psi = 2 r_1^{\rightharpoonup} \left(\frac{-\bar{\xi}_0}{\bar{\psi}_0^2 - \bar{\xi}_0^2} + \frac{\chi_{AC} - \chi_{BC}}{k_B T} \right) . \tag{E.5}$$

For the phase-separated expansion at $(\bar{\xi}_2, \bar{\psi}_1)$, the only addition in comparison to the derivation in Appendix D, which granted the normal form in Eq. (4.13), is the expansion in $\bar{\psi}$, which yields a first-order coefficient

$$\tilde{h}_\psi = \frac{\partial}{\partial \bar{\psi}} \left(r_1^{\leftharpoondown} - r_1^{\rightharpoonup} \right) . \tag{E.6}$$

Such that the normal form for $\bar{\xi}$ becomes

$$\mathrm{d}_t \bar{\xi} = \begin{cases} -\left(\bar{\psi} - \bar{\psi}_0\right) h_\psi - \left(\bar{\xi} - \bar{\xi}_0\right) h_\xi, & \bar{\xi} \leq \xi^{\mathrm{I}}, \\ \bar{r}_1 - \left(\bar{\psi} - \bar{\psi}_1\right) \tilde{h}_\psi - h_{\xi\xi} \left(\bar{\xi} - \bar{\xi}_2\right)^2, & \bar{\xi} > \xi^{\mathrm{I}} . \end{cases} \tag{E.7}$$

The system follows the two coupled equations for the phase-averaged total volume fraction $\bar{\psi}$ and for the phase-averaged reaction extent $\bar{\xi}$ given in Eq. (4.30). The coefficients f_ξ, h_ψ, and $\tilde{h}_\psi$ induces a bilateral coupling between $\bar{\psi}$ and $\bar{\xi}$. Similar to other oscillating chemical reactions [184, 278], the kinetic equation for $\bar{\xi}$ is a sum of a coupling between the two quantities and a bistability. The combination of bistability and a negative feedback between $\bar{\psi}$ and $\bar{\xi}$ induces self-sustained stable chemical oscillations. The negative feedback emerges when large $\bar{\psi}$-values ($\bar{\psi} > \bar{\psi}_1$) leads to a decrease in $\bar{\xi}$ through $\tilde{h}_\psi > 0$, while small $\bar{\psi}$-values ($\bar{\psi} < \bar{\psi}_0$) leads to an increase in $\bar{\xi}$ through $h_\psi < 0$. When this is combined with $\bar{\xi} > \bar{\xi}_1$ decreasing $\bar{\psi}$ and $\bar{\xi} < \bar{\xi}_1$ increasing it, i.e. $f_\xi > 0$, self-sustained cycles in $\bar{\xi}$ and $\bar{\psi}$ appear.

F Deriving the the critical fuel strength

Here, we derive an approximate expression for the critical fuel strength δ^*, above which bistability occurs. To this end, we use the expression for the volume of the dense phase V^{II} at the phase-separated nullcline given in Eq. (4.10), and insert it into the definition of the phase-averaged total volume fraction $\bar{\psi}$ using Eq. (4.6). Finding the conditions for bistability is equivalent to finding non-equilibrium steady

states with $\bar{\psi} < \psi_0^{\mathrm{I}}$, where $\psi_0^{\mathrm{I}} = \phi_A^0 + \phi_B^0$ is the total volume fraction where the dilute branch of the binodal intersects the nullcline of r_1 (Eq. (4.2)). Solving for δ^* gives an expression for the critical fuel strength necessary for bistability as a function of $\Delta\psi = \bar{\psi} - \psi^{\mathrm{I}}$:

$$e^{\frac{\delta^*}{k_B T}} = 1 + (1 - e^{\frac{-\Delta\mu}{k_B T}}) \left(\frac{k_1^{\mathrm{I}}}{k^{\mathrm{II}}} \left(1 + \frac{\psi^{\mathrm{II}} - \psi^{\mathrm{I}}}{\bar{\psi} - \psi^{\mathrm{I}}} \right) - 1 \right) , \tag{F.1}$$

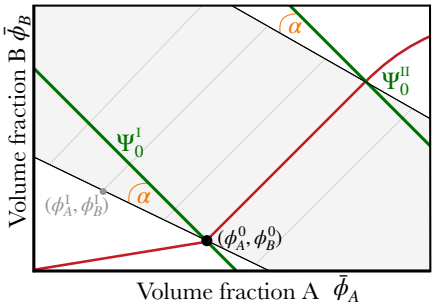


Figure F.1: **Approximate phase diagram:** To approximate the chemical potential around the chemical equilibrium point of the dilute binodal branch (ϕ_A^0, ϕ_B^0), we approximate the binodal as a straight line with angle α relative to the conserved line $\psi_0^{\mathrm{I}} = \phi_A^0 + \phi_B^0$. With this approximation, a point along the binodal can be written as $(\phi_A^0 - \epsilon \sin(\alpha + \pi/4), \phi_B^0 - \epsilon \sin(\alpha - \pi/4))$, leading to the expression for the chemical potential in Eq. (F.3).

where $\Delta\mu = \mu_A - \mu_B$ is the chemical potential difference between A and B. As the fuel converts A to B, the chemical potential difference is negative; $\Delta\mu < 0$. So far, no approximations have been made.

Due to the non-linearity of Eq. (F.1), we perform a Taylor expansion at the point where the r_1 nullcline intersects the dilute branch of the binodal, which we denote (ϕ_A^0, ϕ_B^0). We investigate small deviations around this point, i.e., close to chemical equilibrium where $|\Delta\mu| \ll k_B T$. Furthermore, we follow Ref. [98] and approximate the tie-lines to be orthogonal to the binodal. This implies that $(\psi^{\mathrm{II}} - \psi^{\mathrm{I}})$ is a constant along the binodal and can be set equal to the difference at the chemical equilibrium tie-line $(\psi_0^{\mathrm{II}} - \psi_0^{\mathrm{I}})$. With these assumptions, we find

$$\begin{aligned} e^{\frac{\delta^*}{k_B T}} = 1 - \frac{\Delta\mu}{k_B T} \left[\frac{k_1^{\mathrm{I}}}{k_1^{\mathrm{II}}} \left(1 + \frac{\psi_0^{\mathrm{II}} - \psi_0^{\mathrm{I}}}{\Delta\psi} \right) - 1 \right] \\ + \frac{1}{2} \left(\frac{\Delta\mu}{k_B T} \right)^2 \left[\frac{k_1^{\mathrm{I}}}{k_1^{\mathrm{II}}} \left(1 + \frac{\psi_0^{\mathrm{II}} - \psi_0^{\mathrm{I}}}{\Delta\psi} \right) - 1 \right] , \end{aligned} \tag{F.2}$$

where the Taylor expansion is truncated above second order, ignoring $\mathcal{O}(\Delta\mu/k_B T)^3$. We continue by introducing the angle α of the dilute branch of the binodal at (ϕ_A^0, ϕ_B^0), relative to the conserved line of $\psi_0^{\mathrm{I}} = \phi_A^0 + \phi_B^0$ at the same point. This definition of the angle is such that $\alpha = 0$ gives a binodal parallel to the conserved line, where the dilute branch of the binodal can be approximated as $(\phi_A^0 - \epsilon \sin(\alpha + \pi/4), \phi_B^0 - \epsilon \sin(\alpha - \pi/4))$, where $\epsilon \simeq \Delta\psi/(\sqrt{2} \sin(\alpha))$. An illustra-

tion is given in Fig. F.1. Performing this Taylor expansion in the chemical potential difference, we find $\Delta\mu \approx \Delta\psi\,\theta$, where

$$\theta = \frac{\cos\alpha\,(\phi_A^0)^{-1} + \sin\alpha\,(\phi_B^0)^{-1}}{\sqrt{2}\sin\alpha} + \frac{\chi_{BC} - \chi_{AC}}{k_B T} + \frac{1}{\tan\alpha}\frac{\chi_{AB}}{k_B T}\,. \tag{F.3}$$

Using this approximate expression for the chemical potential in Eq. (F.2), we find

$$\begin{aligned}\exp\left\{\frac{\delta^*}{k_B T}\right\} = 1 + \frac{k_1^{\mathrm{I}}}{k_1^{\mathrm{II}}}\theta\left(\psi_0^{\mathrm{II}} - \psi_0^{\mathrm{I}}\right) - \Delta\psi\,\theta\left(1 - \frac{k_1^{\mathrm{I}}}{k_1^{\mathrm{II}}}\left(1 + \theta\frac{\psi_0^{\mathrm{II}} - \psi_0^{\mathrm{I}}}{2}\right)\right) \\ + \left(\Delta\psi\right)^2\frac{\theta^2}{2}\left(1 - \frac{k_1^{\mathrm{I}}}{k_1^{\mathrm{II}}}\right)\,.\end{aligned} \tag{F.4}$$

The terms independent of $\Delta\psi$ give the expression for the minimal fuel strength needed for bifurcation, as shown in Eq. (4.12) and Fig. 4.3(e), where we have separated the thermodynamic parameters into $\Theta \equiv \left(\psi_0^{\mathrm{II}} - \psi_0^{\mathrm{I}}\right)\theta$, such that

$$\Theta = \left(\psi_0^{\mathrm{II}} - \psi_0^{\mathrm{I}}\right)\left(\frac{\chi_{BC} - \chi_{AC}}{k_B T} + \frac{1}{\tan\alpha}\frac{\chi_{AB}}{k_B T} + \frac{\cos\alpha\,(\phi_A^0)^{-1} + \sin\alpha\,(\phi_B^0)^{-1}}{\sqrt{2}\sin\alpha}\right)\,. \tag{F.5}$$

Notably, the expression diverges at an angle $\alpha = 0$, and any angle $\alpha < 0$ gives a large negative value of Θ, which does not yield physical solutions of δ^* in Eq. (4.12). This can be understood geometrically, as the r_1 nullcline has to lie within the binodal for multiple intersections with ψ to occur, and therefore, ψ has to also lie within the binodal. Note that α is in general negative when $\chi_{AC} > \chi_{BC}$, and positive when $\chi_{AC} < \chi_{BC}$, making the latter a necessity for bistability. The minimal value of Θ, which occurs for $\alpha = \pi/2$, is finite. Consequently, the critical fuel strength δ^* never approaches zero, and a finite fuel strength is always necessary for bistability. Note further that large values of α, which lowers the critical fuel strength, occur when $\chi_{BC} \gg \chi_{AC}$, thereby increasing the constant offset of Θ. It is, therefore, not trivial to minimize the expression for Θ. The predicted critical fuel strength is compared with the analytic expression in Fig. 4.3(e), which agrees with the numerical solution within 10%. As $\Delta\psi$ increases, the deviation from chemical equilibrium increases, and higher order terms in $\Delta\mu$ are needed for an agreement, creating a growing discrepancy between the analytic and numerical results. For the case where the reaction rate coefficient is equal in both phases, the expression simplifies to

$$e^{\frac{\delta^*}{k_B T}} = 1 + \left(\psi_0^{\mathrm{II}} - \psi_0^{\mathrm{I}}\right)\theta + \Delta\psi\left(\psi_0^{\mathrm{II}} - \psi_0^{\mathrm{I}}\right)\frac{\theta^2}{2}\,, \tag{F.6}$$

as the second order contribution of $\Delta\psi$ in Eq. (F.4) cancels.

G Full phase diagrams of excitable and excited droplets

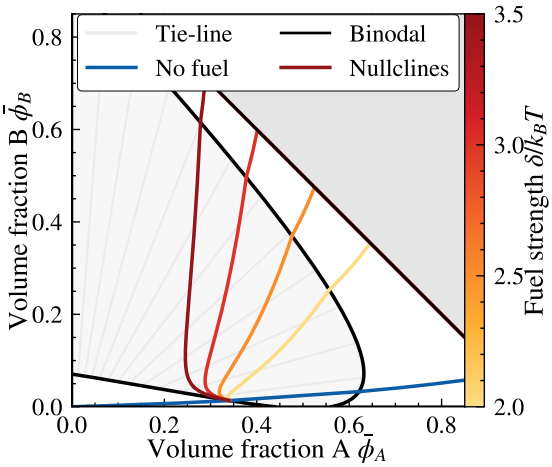


Figure G.1: **Full phase diagram for bistability:** The full phase space, instead of the subset displayed in Fig. 4.2, shows the entire binodal and how the unfueled system selects a single tie-line.

The phase diagrams in Fig. 4.3(a), Fig. 4.6, and Fig. 4.7(a) only show a subset of the whole phase space. Therefore, the entire phase diagrams are shown in Fig. G.1 and Fig. G.2. For the former, we see how the unfueled case (blue) selects a specific tie-line, while the fueled nullclines cross different tie-lines with different phase compositions. Furthermore, the full binodal makes it apparent how A-C only barely phase separates, while B-C is strongly repulsive. These parameters are chosen to achieve the conditions for bistability, as discussed in Appendix F, by making the angle α in Eq. (F.5) large.

The full phase space for an excitable droplet, as depicted in Fig. 4.6 and Fig. 4.7(a), is shown in Fig. G.2(a,b). Here we see that the system is a bistable excitable medium, as there are two stable state points (black dots), and one unstable steady state (white dot). Even though there are two stable steady state points, only the lower one is excitable, as the flow lines show. The separatrix between the two equilibria is approximately given by the conserved line $\bar{\psi}$ of the unstable steady state. All initial conditions with $\bar{\psi}$ smaller than this value will create an excitable system.

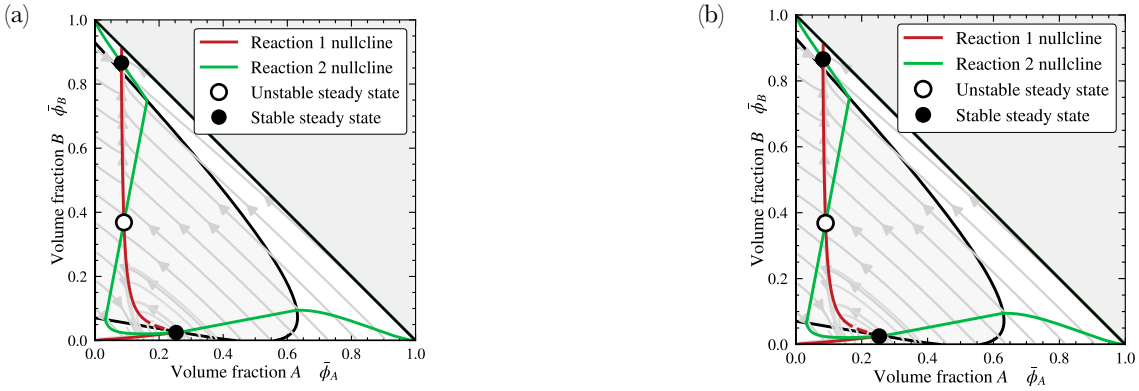


Figure G.2: **Full phase diagram of excitable systems:** The full phase space, instead of the subset displayed in Fig. 4.6a and Fig. 4.7a, shown in (a) and (b) respectively, shows that there are two stable steady state points. Different initial conditions lead to different attractors, roughly separated by the total volume fraction ψ of the unstable steady state point.

H Determining sequence-sequence interaction strengths

The chemical potentials $\mu_i(\{\phi_j\}_{j\in\sigma}) = \nu_i \partial f/\partial\phi_i|_{\phi_{j\neq i}}$ determine the oligomerization rates (Eq. (5.1d)) and are calculated from the free energy density f. We consider a free energy with a mixing entropy and interaction up to second order in the volume fraction ϕ_i, often referred to as the Flory-Huggins free energy [113, 120]:

$$f = \frac{k_B T}{\nu_s}\left[\sum_{i\in\sigma}\left\{\frac{\phi_i}{\rho_i}\log\left(\frac{\phi_i}{\rho_i}\right) + \frac{\omega_i\phi_i}{k_B T}\right\} + \frac{\omega_s\phi_s}{k_B T}\right. \tag{H.1}$$
$$\left. + \phi_s\log\left(\phi_s\right) + \frac{1}{2}\sum_{i\in\sigma}\sum_{j\in\sigma}\frac{\chi_{ij}\phi_i\phi_j}{k_B T} + \phi_s\sum_{i\in\sigma}\frac{\chi_{is}\phi_i}{k_B T}\right].$$

For two monomer types a and b, the number of oligomers grows exponentially with the maximal length L; $|\sigma| = 2(2^L - 1)$. We assume the volumes of the two monomeric units to be equal and time-independent ($\nu_a = \nu_b = \nu_1$ with $\partial_t \nu_i = 0$), and the volume of oligomers to increase linearly with the number of constituting monomers n_i, as $\nu_i = \nu_1 n_i$. Consequently, the relative oligomer molecular volumes $\rho_i = \nu_i/\nu_s$ also scale linearly with the number of monomers n_i. The first four terms in equation (H.1) represent the mixing entropy and internal energy per length ω_i of the oligomers and the solvent. Whereas the last two terms describe the mean field Flory-Huggins interactions between two oligomers i and j as χ_{ij}, and between each oligomer and the solvent s as χ_{is}, relative to the thermal energy $k_B T$. As the sum over all volume fractions must add to one, the solvent volume fraction ϕ_s^α is set by the total volume fractions $\phi_s^\alpha = 1 - \phi_{\text{tot}}^\alpha$.

The Flory-Huggins interactions parameter between oligomers i and j is defined as $\chi_{ij} = e_{ij} - (e_{ii} + e_{jj})/2$ [87], where e_{ij} gives the effective interaction energy between

the two oligomers. The value for e_{ij} is determined by performing a Boltzmann average over all possible interaction energies $\langle e_{ij}^{(\zeta)} \rangle$ when sliding one oligomer along the other. The interaction strength for a specific configuration ζ can be written as (Fig. H.1):

$$e_{ij}^{\zeta,\pm} = \frac{1}{\underline{n_{ij}}} \sum_{\beta=1}^{n_{ij}} \begin{cases} 0 & \text{if } \Gamma_{i,\beta} = 0 \text{ or } \Gamma_{j,\beta\pm\zeta} = 0 \\ e_{aa} & \text{if } \Gamma_{i,\beta} = \Gamma_{j,\beta\pm\zeta} = a \\ e_{ab} & \text{if } \Gamma_{i,\beta} \neq \Gamma_{j,\beta\pm\zeta} \neq 0 \\ e_{bb} & \text{if } \Gamma_{i,\beta} = \Gamma_{j,\beta\pm\zeta} = b \end{cases} , \tag{H.2}$$

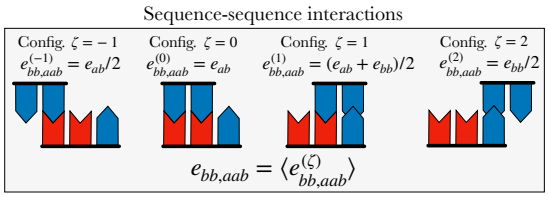


Figure H.1: **Sequence interactions is calculated through a sliding average:** To calculate the interaction strength e_{ij} between two sequences, here $i = bb$ and $j = aab$, we perform an average sliding configurations ζ. Each sliding configuration has an associated interaction energy $e_{ij}^{(\zeta)}$ given by the total energy of the overlapping base-pairs interactions (e_{aa}, e_{ab} e_{bb}) for that given configuration (H.2). The average over all sliding configurations $\langle e_{ij}^{(\zeta)} \rangle$ is a Boltzmann average (H.3), such that favorable configurations (here $\zeta = 0$) contribute more to the average. Both orientations are considered, though not depicted here, due to the inversion symmetry of bb.

where $\Gamma_{i,\beta}$ picks the monomer-type at position β of sequence i. If there is an overhang, and therefore no unit to pair up with, $\Gamma_{i,\beta}$ takes a value of 0. To get the average interaction per length, the sum is normalized by the upper limit of the number of base-pair interactions, which is the length of the shortest of the two sequences, denoted $\underline{n_{ij}} = \min(n_i, n_j)$. The sum thus adds up the total interaction energy for a given configuration, assuming pairwise interactions; these interactions are similar to Watson-Crick base pair interactions for the case of DNA and RNA. The pairwise interaction strengths e_{aa}, e_{ab}, and e_{bb}, are input parameters populating the full interaction matrix. This matrix can be calculated prior to the numerical studies for a given maximal length L, setting the number of possible sequences $|\sigma|$. An illustration of the interaction energy of a specific configuration k is displayed in Fig. 5.1. Averaging over all possible pair-wise configurations, the effective interaction energy is found from the Boltzmann average:

$$e_{ij} = \frac{\sum_{\pm} \sum_{\zeta=-\underline{n_{ij}}+1}^{\overline{n_{ij}}-1} e_{ij}^{\zeta,\pm} \exp\left\{-\frac{e_{ij}^{\zeta,\pm}}{k_B T}\right\}}{\sum_{\pm} \sum_{\zeta=-\underline{n_{ij}}+1}^{\overline{n_{ij}}-1} \exp\left\{-\frac{e_{ij}^{\zeta,\pm}}{k_B T}\right\}}, \tag{H.3}$$

where the sum over $(\pm)$ takes both orientations into account, and the second sum is

over the $(n_i + n_j - 1)$ possible sliding configurations per orientation, where $\overline{n_{ij}}$ gives the largest length of the two sequences $\overline{n_{ij}} = \max(n_i, n_j)$.

The solvent interaction is assumed to be equal to the solvent interaction of its constituting monomeric units, $e_{is} = (a_i e_{as} + b_i e_{bs})/n_i$, entering the Flory-Huggins interactions parameter as $\chi_{is} = e_{is} - (e_{ii} + e_{ss})/2$, where a_i and b_i are the number of monomers of type a and b in sequence i.

A key contribution to the free energy (H.1) is the internal energies per length ω_i. We assign to each monomer type an intrinsic internal energy Ω_a and Ω_b, and let each monomer in the oligomer interact with its neighbors through the back-bone bond energies ϵ_{aa}, ϵ_{bb}, ϵ_{ab}, such that

$$\omega_i = \frac{a_i \Omega_a + b_i \Omega_b}{n_i} + \frac{e_{ii}}{2} + \frac{n_{aa}^{(i)} \epsilon_{aa} + n_{bb}^{(i)} \epsilon_{bb} + n_{ab}^{(i)} \epsilon_{ab}}{n_i}, \tag{H.4}$$

where $n_{jk}^{(i)}$ is the number of (jk) neighbors along oligomer i. The self-interaction term $e_{ii}/2$ is included to retrieve an unbiased sequence distribution, i.e., flat, in the dilute limit, where $\bar{\phi}_s \to 1$. In chapter 5, we treat all backbones equally (sequence-independent), such that $\epsilon_{aa} = \epsilon_{bb} = \epsilon_{ab} = \epsilon$. This way, the backbone has no preference towards specific sequence patterns, making all bias shown exclusively arise from interactions among sequences. Decreasing the value of ϵ will result in a larger average length. More parameters for next neighbor correlation could be introduced to better model sequence distributions in peptides [218, 219].

I Details on cooperativity and rank distributions

The chemical potentials of each sequence can be expressed in terms of the reference chemical potential $\bar{\mu}_i^0$, which is composition independent, and the activity coefficient γ_i, containing the contributions from the interactions among sequences:

$$\bar{\mu}_i = \bar{\mu}_i^0(T, V) + k_B T \log\left(\phi_i \, \gamma_i\{\{\phi_i\}, T, V\}\right). \tag{I.1}$$

Through these quantities, one can define a cooperativity Λ_i, measuring how well adapted a specific sequence is to its environment, as defined in Eq. (5.16). Using the expression for the chemical potentials in Eq. (5.8), one finds that

$$\begin{aligned} \bar{\mu}_i^0 &= 1 - \rho_i - \log \rho_i + \rho_i \frac{\omega_i - \omega_s}{k_B T}, \\ \gamma_i^\alpha &= \left(\frac{1}{\phi_s^\alpha}\right)^{\rho_i} \exp\left\{\rho_i \frac{\sum_{j \in \sigma} \phi_j^\alpha \left(\chi_{ij} - \chi_{js}\right) + \phi_s^\alpha \chi_{is}}{k_B T}\right\}. \end{aligned} \tag{I.2}$$

The ratio of the activity coefficients of each phase sets the partitioning coefficient, $\phi_i^{\mathrm{I}}/\phi_i^{\mathrm{II}} = \gamma_i^{\mathrm{II}}/\gamma_i^{\mathrm{I}}$ [113]. At thermodynamic equilibrium, it is the component with the largest cooperativity that becomes the most abundant.

In fig. 5.8d, the rank distribution for 8-mers was plotted in a lin-lin scale. For large fragmentation rates, the curves in the figure transitioned from an exponential decay on the form $p(k) \propto e^{-k/\lambda}$ to a power law distribution. For completeness, we display the same data in Fig. I.1, but with log-log axis. To verify the transition from an exponential decay in the rank distribution, as displayed in Fig. 5.8d, to a power-law decay, the same data is plotted in a log-log plot in Fig. I.1. Once sufficiently far from equilibrium, the rank distribution approaches a straight line on a log-log representation, indicating a power-law decay. The best-fit exponent for the power law is displayed for different fragmentation rates in the inset, where the upper curve is phase-separated, and the lower is mixed. Note that the inset includes the best-fit power-law exponent for distributions with only a weak tendency of a power-law. A power law exponent equal to one, as suggested by Zipf's law [224], is indicated as a horizontal line in the inset.

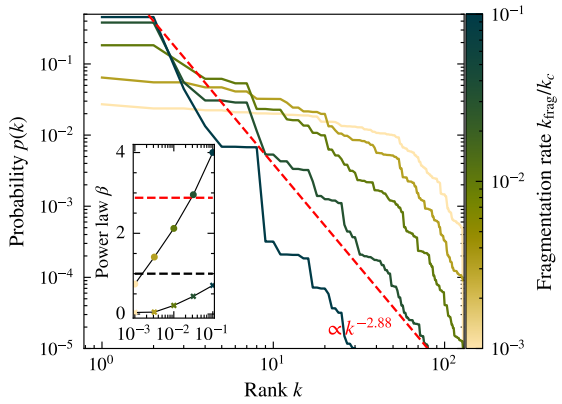


Figure I.1: **The rank distribution transitions from an exponential to a power-law with fragmentation rate:** For a low fragmentation rate, the rank distribution of 8-mers follows an exponential decay. With large fragmentation rates, on the other hand, the distribution transitions to a power-law behavior $p(k) \propto k^{-\beta}$. The exponent β increases strongly with the fragmentation rate k_{frag} (inset) if the system is phase separated (upper curve). The example exponent of 2.88 is displayed in the inset together with Zipf's law exponent of 1 [224]. For a mixed system (lower curve inset), $\beta = 1$ is not reached.

For non-equilibrium steady states, the constraint of chemical equilibrium is lifted, and Gibbs phase rule no longer holds. To estimate the number of coexisting phases, we can instead deploy the recent framework of Qiang and Zwicker [279, 280], by which we can calculate the number of phases and their composition given a set of volume fractions, their molecular volumes, and their interactions. Their framework does not account for components that react among each other, which would couple the volume fraction distribution to phase-coexistence [113]. We apply their framework to the non-equilibrium steady state sequence distribution in Fig. 5.8. We search for the possibility whether the system's steady state could reduce its free energy by forming additional phases. Using their numerical package `flory` [280], we find two phase coexistence, thus supporting the a priori assumptions of two phases.

J Effects of ligation on sequence distribution away from equilibrium

In chapter 5, we initially considered monomer pickup obeying detailed balance of the rates (see Eq. (5.1c)), and maintained the system away from equilibrium via fragmentation reaction steps that break detailed balance of the rates (see Eq. (5.13)). In this appendix, we briefly investigate the effects of ligation reactions to probe the generality of the key result of chapter 5, namely that phase separation away from equilibrium creates a strong selection pressure during sequence oligomerization.

Ligation refers to the joining or fusion of the ends of two oligomers, $i + j \rightleftharpoons k$, with the reaction rate given by:

$$r_i^{\text{lig}} = k_c \left[\exp\left\{ \frac{\mu_i + \mu_j}{k_B T} \right\} - \exp\left\{ \frac{\mu_k}{k_B T} \right\} \right] . \tag{J.1}$$

For simplicity, we take k_c to be the same for ligation and monomer pickup. An example reaction is $aba + bbbb \rightleftharpoons ababbbb$. The chemical equilibrium condition in Eq. (5.5) remains unchanged by the inclusion of ligation implying that the equilibrium distributions shown in Figs. 5.3 to 5.8 are unaffected upon adding the ligation. However, the non-equilibrium steady-state distribution will be affected.

In Fig. J.1, the distribution of *ab*-neighbors as a function of the relative fragmentation rate k_{frag}/k_c shows that the preference for longer homopolymeric sequences observed in Fig. 5.11 still occurs when ligation is included. However, the distribution becomes broader, and the corresponding peaks are slightly reduced. Moreover, a higher fragmentation rate is required to achieve selection of sequences similar to those seen without ligation (Fig. 5.11(b)). The proposed selection mechanism for monomer pickup away from equilibrium is based on thermodynamic sequence–sequence interactions, which remain relevant for ligation. While the additional reaction pathways introduced by ligation affect the properties of the distribution in Fig. J.1, the leftward shift of the mean corresponding to the selection of sequences with longer

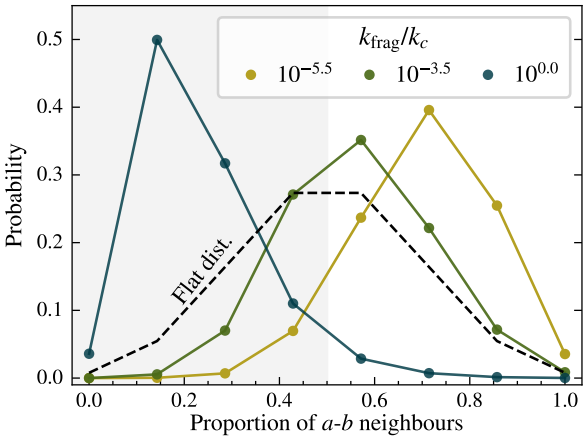


Figure J.1: **Homo-polymeric enrichment bias away from equilibrium persists with ligation:** In addition to monomer-pickup and release, the non-equilibrium steady state distribution when accounting for ligation (Eq. (J.1)) shifts to more homopolymeric sequences, as it did without ligation in Fig. 5.11b.

segment of the same monomeric types is robust.

K Net oligomerization rate with non-equilibrium fragmentation

The net rate $\tilde{r}_i^\alpha$ contains the monomer-pickup/release rate r_i (Eq. (5.1c)) that fulfills detailed balance of the rates (Eq. (5.1e)), and the non-equilibrium fragmentation rate $h^\alpha_{i\rightharpoonup j+|i-j|}$ (Eq. (5.13)):

$$\tilde{r}_i^\alpha = r_i^\alpha - \sum_{j\in\mathcal{S}_i} h^\alpha_{i\rightharpoonup j+|i-j|} + \sum_{j\in\sigma} \frac{n_i}{n_j} h^\alpha_{j\rightharpoonup i+|i-j|}(1+\delta_{i,|i-j|})\delta_{i\in\mathcal{S}_j}\,, \tag{K.1}$$

where $\mathcal{S}_i$ is the set of the (n_i-1) sub-oligomers of sequence i resulting from a single fragmentation event, and $\delta_{j\in\mathcal{S}_i}$, takes a value of 1 if j is a member of $\mathcal{S}_i$, and zero otherwise, and $\delta_{i,|i-j|}$ is the Kronecker-delta.

One can find and approximate expression (5.14) for the sequence distribution of long sequences at the non-equilibrium steady state (NESS) by using that all time derivatives of Eq. (5.1a) is zero. This results in the condition that

$$\tilde{r}_i^\alpha - j_i^\alpha = 0 \tag{K.2}$$

concomitantly in both phases. Multiplying this condition by the respect phase volume yields, by utilizing Eq. (B.5), the simpler condition of

$$0 = \tilde{r}_i^{\mathrm{I}} V^{\mathrm{I}} + \tilde{r}_i^{\mathrm{II}} V^{\mathrm{II}}\,. \tag{K.3}$$

For olgiomers of the maximal length L, the gain term from fragmentation in Eq. (K.1) is zero, such that Eq. (K.3) becomes

$$0 = V^{\mathrm{I}}\left(r_i - \sum_{j\in\mathcal{S}_i} h^{\mathrm{I}}_{i\rightharpoonup j+|i-j|}\right) + V^{\mathrm{II}}\left(r_i - \sum_{j\in\mathcal{S}_i} h^{\mathrm{II}}_{i\rightharpoonup j+|i-j|}\right),$$

where the phase condition of the oligomerization rate has been omitted due to phase coexistence. From the diluteness of long sequences in the oligomer-poor phase (II), approximately no fragmentation of these sequences is occurring, such that

$$\frac{V}{V^{\mathrm{II}}} r_i = \sum_{j\in\mathcal{S}_i} h^{\mathrm{II}}_{i\rightharpoonup j+|i-j|}\,. \tag{K.4}$$

Using the expression for $h^\alpha_{i\rightharpoonup j+|i-j|}$ in Eq. (5.13), and the definition of the reaction

rate r_i (5.1d), where the reverse pathway has been omitted,

$$\frac{Vk_c}{V^{\mathrm{II}}} \sum_{j \in \mathcal{R}_i} \exp\left\{ \frac{\mu_j + \mu_m}{k_B T} \right\} = k_{\mathrm{frag}} \phi_s^{\mathrm{II}} \phi_i^{\mathrm{II}} \,. \tag{K.5}$$

Here, $\mathcal{R}_i$ is the set of sequences that can oligomerize into sequence i. Solving for ϕ_i^{II} yields the final expression of Eq. (5.14).

L Gillespie-algorithm for sequence kinetics

The Gillespie-algorithm integrates the master equation (Eq. (6.1)) exactly by stochastically choosing different processes and updating particle numbers. Given a system state at time t, it asks two questions *what is the next event?* and *when does it occur?*. The question is answered by drawing a random number ξ_2, and choosing the small integer value α satisfying

$$\alpha = \min_\gamma \left(\sum_{\alpha \in \mathcal{S}}^{\gamma} \mathcal{W}_\alpha > \xi_2 \sum_{\beta \in \mathcal{S}} \mathcal{W}_\beta \right) . \tag{L.1}$$

Here ξ_2 is a random number between 0 and 1 drawn from a uniform probability distribution. The sum $\sum_{\beta \in \mathcal{S}} = \sum_{\mathcal{W} \in \mathcal{P}} \sum_{\alpha \in T(\mathcal{W})}$ is a sum over all possible reaction pathways, making the β-sum the total rate of all pathways from the current state of the system. This integration scheme is equivalent to choosing each possible pathway proportionally to their rate, where the sum over all pathways acts as a normalization condition. Once a pathway is found, the time is increased by the logarithm of random number ξ_1, multiplied by the net rate of all processes

$$\tau = -\frac{1}{\sum_{\beta \in \mathcal{S}} \mathcal{W}_\beta} \log\left(\xi_1\right) . \tag{L.2}$$

The state and time is thus changes according to $t \to t + \tau$ and $\boldsymbol{N} \to \boldsymbol{N} + \boldsymbol{\nu}_{\alpha'}$. Given a system in a state $\boldsymbol{N}$ at time t, the probability, that the system's next reaction will occur in the infinitesimal time interval $[t+\tau, t+\tau+\mathrm{d}\tau]$ and will be of stoichiometry corresponding to the process α is given by

$$p(\tau, \alpha | \boldsymbol{N}, t) = \mathcal{W}_\alpha(\boldsymbol{N}, t) \exp\left\{ -\tau \sum_{\beta \in \mathcal{S}} \mathcal{W}_\beta(\boldsymbol{N}, t) \right\} . \tag{L.3}$$

The integration scheme in Eq. (L.1) and Eq. (L.2) exactly samples this probability distribution. Not that the rates $\mathcal{W}_\beta$ are not rate densities, but rate for particle numbers. They are therefore not equal to the reaction rates in Eq. (2.38), but

must be scaled with their respective phase volumes. This distinction is irrelevant for systems with constant volumes, but is vital for the case of coexisting phases which exchange volume.

There are several advantages of employing the Gillespie algorithm in contrast to integrating Eq. (2.60) with a stochastic noise obeying the fluctuation dissipation theorem. When working with particle numbers, there is no risk of the noise resulting in negative particle numbers or volume fractions larger than 1. In the Gillespie algorithm, the rate of a process with occupation number of zero, will also be zero. The divergence of chemical potential as $\phi_i \to 0$ is not a true divergence, but an artifact from the Sterling's approximation of the factorials in the derivation of the Flory-Huggins free energy. With the Gillespie algorithm, Sterling's approximation does not have to be made. Furthermore, for a large composition space, as studied in chapter 5, the used integration scheme must keep track of all possible states at all times. To make matters worse, most states for long sequences are very dilute, meaning we must accurately integrate volume fractions on the order of 10^{-7} and smaller, as the system relaxes. When working with particle numbers, many of these states will not be occupied, even for large total particle numbers $\bar{N}_{\text{tot}}$. In chapter 6, this fact was exploited by working with a dynamic sequence space. At each point in time, the rates and interactions were only calculated between occupied states, and if a process created a sequence n not yet encountered, their internal energy ω_n and interaction strengths χ_{in} with all other components were calculated and included in the free energy density. This allowed us to never enforce an upper length limit, and working with much larger sequences spaces than in chapter 5, where all possible states were defined together with the initial condition.

The Gillespie algorithm becomes less numerically efficient as the total number of particles in the system increases, and scales linearly with the number of processes $|\mathcal{S}|$. There exists modifications to the algorithm in order to enhance efficiency. One such modification is a method called τ leaping [281], where instead of doing a single sample α for each state $\boldsymbol{N}$, one does multiple before updating a rescaled time change τ and the state vector. An issue with tau-leaping for our system is that due to the dimensionality of $\boldsymbol{N}$, many states will have low occupancy. With low occupancy, leaps can produce negative occupations or large relative changes in the occupation of the state.

An alternative procedure to enhance efficiency is the Next Reaction Method [282], which makes the numerical cost scale logarithmically with the number of reactions, instead of linearly. The speed-up is achieved by only re-calculating the new rate of the process just sampled, and keep the rest of the rates identical to the ones in the previous time-step. This speed-up does not result produce the issues of tau-leaping, but in general does not hold for non-dilute chemistry. In our case, the

change of composition of one component will effect the rate of all processes. The Next Reaction Method is therefore no longer an exact integration method. Though, since the changes in reaction rates are small, it could perhaps be combined with tau-leaping. This method would consist of only updating the rates for reactions where the reactant's occupancy was changed by the previous event α, but only for a fixed number of times before recalculating all reaction rates. This procedure would be a combination of tau-leaping and the Next Reaction Method, which could speed-up systems with many reactions (as studied in chapter 6) without the aforementioned issues of tau-leaping (negative occupations for large state spaces) and the Next Reaction Method (one occupation change affecting the rate of all reactions).

Fig.	χ_{AC}	χ_{BC}	χ_{AB}	ω_A	ω_B	μ_C^r
3.2a	3.0	0.6	-0.6	-0.2	0.1	(-0.2, -0.5, -0.9)
3.2b i	3.0	0.6	-0.6	-0.2	0.1	-0.2
3.2b ii	2.84	2.04	-0.6	-0.2	0.1	-0.15
3.2b iii	1.72	1.72	-2.19	-0.1	0.4	-0.2
3.2b iv	-0.60	0.60	3.0	-0.2	0.1	-1.40
3.3 a,b	3.0	0.6	-0.6	-0.2	0.1	-0.5
3.4 a,b	3	-3	2	-10	0.0	-10
3.4c	–	–	–	0.0	-6.0	$\Delta 2$
3.4 d,e	2	3.5	-2	-10	0.0	2.77
3.4f	–	–	–	-6.0	0.0	$-\Delta 5$
3.5 a,b,c	3	-0.6	0.6	0	-6.0	$\langle -0.3 \rangle,\ \mu_C^{\text{amp}} = 0.6$
3.5 d,e,f	3	-0.6	0.6	0.4	1.8	$\langle \mu_C^r \rangle = -0.538,\ \mu_C^{\text{amp}} = 0.6$
3.7 solvent	3.0	-0.6	0.6	0.4	1.8	$\langle \mu_C^r \rangle = -0.35,\ \mu_C^{\text{amp}} = 0.35$
3.7 reactant	3.0	-0.6	0.6	0.4	1.8	$\langle \mu_A^r \rangle = -0.08,\ \mu_A^{\text{amp}} = 0.35$

Table M.1: The interaction strengths χ_{ij}, internal energies ω_i, and reservoir chemical potential values μ_C^r used for all figures is specified here. The Δ in the reservoir chemical potential means change relative to the initial chemical potential. For all figures in chapter 3, we have chosen $\nu_i = 1$, $k_c = 1$, $V(t = 0) = 1$, and $\omega_C = 0$. All energies are measured in terms of $k_B T$.

M Parameter tables for figures:

M.1 Parameters for chapter 3

The parameters used to generate all the figures are given in Table M.1. Our finding that turnover of solvophilic substrates is enhanced during drying agrees with the interactions of substrates used in phosphorylation reactions [56, 60], where speed-up through wet-dry cycles also have been observed. The interaction parameters used for Fig. 3.2, 3.3, 3.4 are motivated by systems with similar phase behavior [55, 79, 56, 96]. Moreover, also purified proteins show similar interactions [283].

We note that evaporating a major fraction of the solvent can strongly decrease the diffusion coefficient, potentially violating the conditions of fast-diffusion (Eq. (2.59)), as reacting compounds can become more crowded. However, removing the solvent also decreases the system's size, potentially maintaining the validity of the condition throughout the evaporation process.

Figure	χ_{AB}	χ_{AC}	χ_{BC}	ω_A	ω_B	ω_C	k_2/k_1	δ	κ	ψ_δ	$\bar{\psi}$
4.3(a, b, c), G.1	0.0	2.01	3	-2.5	0.1	-2.25	–	—	0.0005	0.25	0.16
4.3(d)	0.0	2.01	3	-2.5	0.1	-2.25	–	4	0.0005	0.25	–
4.3(e)	0.0	2.01	3	-2.5	0.1	-2.25	–	4	0.0005	0.25	–
4.6	0.0	2.01	3	-1.5	0.1	-1.68	5	4	0.0005	0.25	–
4.7(a,b),4.8(a,b)	0.0	2.01	3	-2.5	0.1	-2.25	1.0	4	0.0005	0.25	–
4.7(b)	0.0	2.01	3	-2.5	0.1	-2.25	–	–	0.0005	0.25	–
4.9	0.6	3.0	-0.6	-2.5	0.5	-1.35	1.0	2.0	0.04	0.225	–
4.10	0.0	2.01	3	-2.5	0.1	-2.25	1.0	—	0.0005	0.25	–
G.2	0.0	2.01	3	-2.5	0.1	-2.25	1.0	4	0.0005	0.25	–

Table M.2: **Parameter choices for each figure in fast-diffusion limit:** Common for all figures in chapter 4 that the molecular volumes are equal, $\nu_i = 1$. All energy values in the table are given in units of k_BT (χ_{ij}/k_BT, ω_i/k_BT and δ/k_BT).

Figure	L/ℓ	H/ℓ	$\kappa_i/\mathcal{K}$	δt	N_x	N_y	U_0	δ	k_2/k_1
4.8	0.15	0.15	$1.8 \cdot 10^{-5}$	10^{-7}	250	250	-1.5	4	10
4.11&4.13	2	2	$1.8 \cdot 10^{-5}$	10^{-6}	250	250	0	–	–
4.12&4.14a	2	2	$1.8 \cdot 10^{-5}$	10^{-6}	250	250	0	–	–
4.14b,c	10	0.5	$1.8 \cdot 10^{-5}$	10^{-6}	1000	1000	0	4	–
4.15	10, 5	0.5, 0.5	$1.8 \cdot 10^{-5}$	10^{-6}	1000, 500	50, 50	0	4	–

Table M.3: **Parameter choices for continuum equations:** Common for all figures in chapter 4 is that the molecular volumes are equal, $\nu_i = 1$, and $\chi_{AB} = 0$, $\chi_{AC} = 2.01 k_BT$, $\chi_{BC} = 3k_BT$, $\omega_A = -2.5$, $\omega_B = 0.1$, $\omega_C = -2.25$. We have further defined the reaction-diffusion length scale as $\ell = \sqrt{\Lambda_i^{(0)}/(k_BTk_1)}$ and $\mathcal{K} = k_BT\ell^2$.

M.2 Parameters for chapter 4

For reproducibility, the parameters used to generate all figures in the fast-diffusion limit are given in Table M.2, and the parameters for the continuum equations are given in Table M.3. We have assumed that the reaction rate coefficients are composition-independent $k_1^{\mathrm{I}} = k_1^{\mathrm{II}} = k_1$ and $k_2^{\mathrm{I}} = k_2^{\mathrm{II}} = k_2$. In order to retrieve the diffusion equation in the dilute limit ($\chi_{ij} = 0$), the mobilities scale linearly with the volume fractions according to $\Lambda_i = \Lambda_i^{(0)}\phi_i\phi_C$.

To solve the full spatial equation (Eq. (4.1)) needed to generate the data, the spectral method is employed [284]. The spectral method solves the continuous dynamic Eqs. (4.1) in (spatial) Fourier-space. Periodic boundary conditions were used. Lastly, the external potential is $U = U_0\delta_{i,N_x/2}\delta_{j,N_y/2}$, where i and j are the spatial indices in the x and y direction respectively, and $\delta_{\alpha,\beta}$ is the Kronecker delta.

For Fig. 4.5, the parameters were slightly different from the ones used in the rest of the chapter, and are therefore listed here instead of in the table: $\chi = 1.19k_BT$, $\bar{\phi}_{\mathrm{tot}} = 0.25$, $\Delta\omega = -k_BT$, $\omega_\infty = 0$, $\omega_s = 0$, $\delta = 2k_BT$, $\rho_1 = 1$.

Figure	e_{ii}	e_{AB}	e_{AS}	e_{BS}	e_{SS}	ϵ	Ω_A	$\bar{\phi}_A$	$\bar{\phi}_B$	L
5.3	0	-5	-0.92	-0.92	-3.35	-1.0	2	0.03	0.03	5
5.5	0	-5	-0.92	-0.92	-3.35	-1.0	1	–	–	8
5.4	0	e_{ab}	$2.005 + \frac{e_{ab}}{4}$	$2.005 + \frac{e_{ab}}{4}$	-3.35	-0.6	0.6	–	–	9
5.6	0	-5	-0.92	-0.92	-3.35	-1.0	1.0	0.11	0.11	8
5.7	0	e_{ab}	$2.005 + \frac{e_{ab}}{4}$	$2.005 + \frac{e_{ab}}{4}$	-3.35	-0.6	0.6	–	–	9
5.9	0	-5	-0.92	-0.92	-3.35	-1.0	1.0	0.11	0.11	8
5.8	0	-5	-0.92	-0.92	-3.35	-1.0	1.0	0.11	0.11	8
I.1	0	-5	-0.92	-0.92	-3.35	-1.0	1.0	0.11	0.11	8
5.10	0	-5	-0.92	-0.92	-3.35	-1.0	1.0	–	–	8
5.11 & 5.12	0	-5	-0.92	-0.92	-3.35	-1.0	1.0	0.11	0.11	8

Table M.4: **Parameters used for figures in chapter 5:** Common for all figures in chapter 5 is that $\rho_1 = 1$, $k_c = 1$, $\omega_S = 0$, and $k_BT = 1$. All energies in this table are measured in units of k_BT. The i subscript denotes both A and B.

M.3 Parameters for chapter 5

In chapter 5, we considered the case where phase equilibrium holds at each time during the oligomerization kinetics. As a result, each thermodynamic quantity is homogeneous in each phase. This simplified case is satisfied if diffusion rates of reacting components are fast compared to the time scales of the chemical rates and fragmentation rate coefficients:

$$k_c,\ k_{\text{frag}} \ll \frac{D_{\min}}{l^2}\,, \tag{M.1}$$

where l is the characteristic system size and $D_{\min}$ is the smallest diffusion coefficient of all the components. The validity of this approximation can be tested through kinetic measurements of diffusion coefficients and chemical reaction rates. Using experimentally measured values of the diffusion coefficient of single stranded RNA and DNA [285, 286] and typical reaction time-scales of oligomerization [287], the system size should be much smaller than 100μm, which holds for many biologically relevant systems [209, 214] and for investigated experimental settings related the origins of Life [213, 237]. For larger system sizes, weak gradients are expected to occur, which can be theoretically described using a sharp interface method [120], where our model at phase equilibrium describes the dynamics of sequences at the condensate interface. The parameters used are listed in Table M.4.

Figure	e_{ii}	e_{AB}	e_{iS}	e_{SS}	ϵ	Ω_i	$\bar{\phi}_i$	k_D	k_f	k_{tal}	$\delta^{\text{I/II}}$
6.3	0	-4	-1.175	-3.35	-2.5	2.5	0.032	5	0	0	$0,4$
6.4,6.5,6.6,6.7,6.8	0	-4	-1.175	-3.35	-2.5	2.5	0.032	5	10^{-3}	1	$0,0$

Table M.5: **Parameter used for all figures:** Common for all figures in chapter 6 is that $\rho_1 = 1$, $k_c = 1$, $\omega_S = -3k_BT$. All energies in this table are measured in units of k_BT, and all time-scales in units of $1/k_c$. The index i here runs over the monomer types: $i \in \{A, B\}$.

M.4 Parameters for chapter 6

Compared to previous work, we are no longer exactly at phase equilibrium. Instead, we allow deviations away from phase equilibrium, with fluxes back to phase equilibrium proportional to the exponential of the chemical potential (see table 6.1). The parameters used are listed in Table M.5. All rate coefficients are chosen to be phase-independent, with a factor proportional to the relative phase volume. The factor of the phase volume is needed as the rates are not rate densities, but govern the particle numbers N_i.

The reaction rate of ligation in Eq. (6.5), is proportional to the product of P_{ik} and P_{jk}, meant to capture the probability of finding sequence k hybridized with sequence i and j respectively. The prefactor of this probability is chosen such that its product is at a maximum 1, granting

$$P_{ik}P_{jk} = \exp\left\{\frac{2e_{ab}}{k_BT} - \frac{e_{ik} + e_{jk}}{k_BT}\right\}. \tag{M.2}$$

The more favorable both sequences i and j interact with k, the larger the ligation rate becomes. Note that e_{ij} is the interaction strength per length, meaning its maximal value is e_{ab}.

References

[1] Ivar Svalheim Haugerud, Pranay Jaiswal, and Christoph A. Weber. Nonequilibrium Wet–Dry Cycling Acts as a Catalyst for Chemical Reactions. *The Journal of Physical Chemistry B*, 128(7):1724–1736, 2024.

[2] Giacomo Bartolucci, Ivar S. Haugerud, Thomas C. T. Michaels, and Christoph A. Weber. The interplay between biomolecular assembly and phase separation. *eLife*, 13, 2024.

[3] Christine ME Kriebisch, Olga Bantysh, Lorena Baranda Pellejero, Andrea Belluati, Eva Bertosin, Kun Dai, Maria de Roy, Hailin Fu, Nicola Galvanetto, Julianne M Gibbs, et al. A roadmap toward the synthesis of life. *Chem*, 11(3), 2025.

[4] Ivar S Haugerud, Giacomo Bartolucci, Dieter Braun, and Christoph A Weber. Theory for sequence selection via phase separation and oligomerization. *arXiv preprint arXiv:2410.08778*, 2024.

[5] Ivar S Haugerud, Hidde D Vuijk, Job Boekhoven, and Christoph A Weber. Excitability and oscillations of active droplets. *arXiv preprint arXiv:2503.11604*, 2025.

[6] Pranay Jaiswal, Ivar S. Haugerud, Hidde D. Vuijk, and Christoph A. Weber. Harvesting chemical power from cyclic environments. *arXiv preprint arXiv: 2510.10266*, 2025.

[7] Jean Baptiste van Helmont. Ortus medicinae; id est, initia physicae inaudita. *Progressus medicinae novus, in morborum ultionem, ad vitam longam*, orig. 1648, reprint 1966.

[8] Emily C Parke. Flies from meat and wasps from trees: Reevaluating francesco redi's spontaneous generation experiments. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 45:34–42, 2014.

[9] G Brent Dalrymple. The age of the earth in the twentieth century: a problem (mostly) solved. *Geological Society, London, Special Publications*, 2001.

[10] AE Ringwood. Some aspects of the thermal evolution of the earth. *Geochimica et Cosmochimica Acta*, 20(3-4):241–259, 1960.

[11] Robin M Canup and Erik Asphaug. Origin of the moon in a giant impact near the end of the earth's formation. *Nature*, 412(6848):708–712, 2001.

[12] Kent C Condie and Victoria Pease. *When did plate tectonics begin on planet Earth?*, volume 440. Geological Society of America, 2008.

[13] AG Cairns-Smith, P Ingram, and GL Walker. Formose production by minerals: possible relevance to the origin of life. *Journal of Theoretical Biology*, 35(3):601–604, 1972.

[14] F Albarède and Janne Blichert-Toft. The terrestrial craddle of life. *Origins of Life: Self-Organization and/or Biological Evolution?*, pages 1–12, 2009.

[15] Nicholas T Arndt and Euan G Nisbet. Processes on the young earth and the habitats of early life. *Annual Review of Earth and Planetary Sciences*, 40(1):521–549, 2012.

[16] Ben KD Pearce, Andrew S Tupper, Ralph E Pudritz, and Paul G Higgs. Constraining the time interval for the origin of life on earth. *Astrobiology*, 18(3):343–364, 2018.

[17] Jun Korenaga. Was there land on the early earth? *Life*, 11(11):1142, 2021.

[18] GJ Cooke, DR Marsh, Catherine Walsh, Benjamin Black, and J-F Lamarque. A revised lower estimate of ozone columns during earth's oxygenated history. *Royal Society Open Science*, 9(1):211165, 2022.

[19] Matthew S Dodd, Dominic Papineau, Tor Grenne, John F Slack, Martin Rittner, Franco Pirajno, Jonathan O'Neil, and Crispin TS Little. Evidence for early life in earth's oldest hydrothermal vent precipitates. *Nature*, 543(7643):60–64, 2017.

[20] Elizabeth A Bell, Patrick Boehnke, T Mark Harrison, and Wendy L Mao. Potentially biogenic carbon preserved in a 4.1 billion-year-old zircon. *Proceedings of the National Academy of Sciences*, 112(47):14518–14521, 2015.

[21] Edmund RR Moody, Sandra Álvarez-Carretero, Tara A Mahendrarajah, James W Clark, Holly C Betts, Nina Dombrowski, Lénárd L Szánthó, Richard A Boyle, Stuart Daines, Xi Chen, et al. The nature of the last universal common ancestor and its impact on the early earth system. *Nature Ecology & Evolution*, 8(9):1654–1666, 2024.

[22] Madeline C Weiss, Filipa L Sousa, Natalia Mrnjavac, Sinje Neukirchen, Mayo Roettger, Shijulal Nelson-Sathi, and William F Martin. The physiology and habitat of the last universal common ancestor. *Nature microbiology*, 1(9):1–8, 2016.

[23] Fouad El Baidouri, Chris Venditti, Sei Suzuki, Andrew Meade, and Stuart Humphries. Phenotypic reconstruction of the last universal common ancestor reveals a complex cell. *bioRxiv*, pages 2020–08, 2020.

[24] Gerald F Joyce. The antiquity of rna-based evolution. *nature*, 418(6894):214–221, 2002.

[25] Francis H Crick. On protein synthesis. In *Symp soc exp biol*, volume 12, page 8, 1958.

[26] James D Watson. *Molecular biology of the gene*. Pearson Education India, 2004.

[27] Harold S Bernhardt. The rna world hypothesis: the worst theory of the early evolution of life (except for all the others). *Biology direct*, 7(1):23, 2012.

[28] Michael P. Robertson and Gerald F. Joyce. The Origins of the RNA World. *Cold Spring Harbor Perspectives in Biology*, 4(5):a003608, 2012.

[29] K. Kruger, P. J. Grabowski, A. J. Zaug, J. Sands, D. E. Gottschling, and T. R. Cech. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell*, 31(1):147–157, 1982.

[30] Leslie E Orgel and FH Crick. Anticipating an rna world. some past speculations on the origin of life: where are they today? *The FASEB journal: official publication of the Federation of American Societies for Experimental Biology*, 7(ISSN: 0892-6638), 1993.

[31] Walter Gilbert. Origin of life: The RNA world. *Nature*, 319:618, 1986.

[32] Martha J. Fedor and James R. Williamson. The catalytic diversity of RNAs. *Nature Reviews. Molecular Cell Biology*, 6(5):399–412, 2005.

[33] Jan M Hollis, Frank J Lovas, Philip R Jewell, and LH Coudert. Interstellar antifreeze: ethylene glycol. *The Astrophysical Journal*, 571(1):L59, 2002.

[34] John D Sutherland. The origin of life—out of the blue. *Angewandte Chemie International Edition*, 55(1):104–121, 2016.

[35] Ronald Breslow. On the mechanism of the formose reaction. *Tetrahedron Letters*, 1(21):22–26, 1959.

[36] Rosa Larralde, Michael P Robertson, and Stanley L Miller. Rates of decomposition of ribose and other sugars: implications for chemical evolution. *Proceedings of the National Academy of Sciences*, 92(18):8158–8160, 1995.

[37] Robert Shapiro. Prebiotic ribose synthesis: a critical analysis. *Origins of Life and Evolution of the Biosphere*, 18(1):71–85, 1988.

[38] Igor V Kornienko, Olga Yu Aramova, Anna A Tishchenko, Dmitriy V Rudoy, and Michael Leonidas Chikindas. Rna stability: a review of the role of structural features and environmental conditions. *Molecules*, 29(24):5978, 2024.

[39] Kristian Le Vay and Hannes Mutschler. The difficult case of an rna-only origin of life. *Emerging Topics in Life Sciences*, 3(5):469–475, 2019.

[40] Zhonggang Hou, Peng Jiang, Scott A Swanson, Angela L Elwell, Bao Kim S Nguyen, Jennifer M Bolin, Ron Stewart, and James A Thomson. A cost-effective rna sequencing protocol for large-scale gene expression studies. *Scientific reports*, 5(1):9570, 2015.

[41] Huajuan Shi, Ying Zhou, Erteng Jia, Min Pan, Yunfei Bai, and Qinyu Ge. Bias in rna-seq library preparation: Current challenges and solutions. *BioMed research international*, 2021(1):6647597, 2021.

[42] Bruce Alberts, Dennis Bray, Karen Hopkin, Alexander D Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter. *Essential cell biology*. Garland Science, 2015.

[43] Manfred Eigen. Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften*, 58(10):465–523, 1971.

[44] Manfred Eigen, John McCaskill, and Peter Schuster. The molecular quasi-species. *Advances in chemical physics*, 75:149–263, 1989.

[45] Maxine J McCall, Philip Hendry, and Philip A Jennings. Minimal sequence requirements for ribozyme activity. *Proceedings of the National Academy of Sciences*, 89(13):5710–5714, 1992.

[46] Michael Yarus. *Life from an RNA world: the ancestor within*. Harvard University Press, 2010.

[47] K. R. Birikh, P. A. Heaton, and F. Eckstein. The structure, function and application of the hammerhead ribozyme. *European Journal of Biochemistry*, 245(1):1–16, 1997.

[48] Christopher Deck, Mario Jauker, and Clemens Richert. Efficient enzyme-free copying of all four nucleobases templated by immobilized RNA. *Nature Chemistry*, 3(8):603–608, 2011.

[49] Angelica Mariani, David A. Russell, Thomas Javelle, and John D. Sutherland. A Light-Releasable Potentially Prebiotic Nucleotide Activating Agent. *Journal of the American Chemical Society*, 140(28):8657–8661, 2018.

[50] Travis Walton, Wen Zhang, Li Li, Chun Pong Tam, and Jack W. Szostak. The Mechanism of Nonenzymatic Template Copying with Imidazole-Activated Nucleotides. *Angewandte Chemie (International Ed. in English)*, 58(32):10812–10819, 2019.

[51] Sreekar Wunnava, Christina F. Dirscherl, Jakub Výravský, Aleš Kovařík, Roman Matyášek, Jiří Šponer, Dieter Braun, and Judit E. Šponer. Acid-Catalyzed RNA-Oligomerization from 3',5'-cGMP. *Chemistry (Weinheim an Der Bergstrasse, Germany)*, 27(70):17581–17585, 2021.

[52] Stanley L Miller. A production of amino acids under possible primitive earth conditions. *Science*, 117(3046):528–529, 1953.

[53] Christof B. Mast, Severin Schink, Ulrich Gerland, and Dieter Braun. Escalation of polymerization in a thermal gradient. *Proceedings of the National Academy of Sciences*, 110(20):8030–8035, 2013.

[54] Thomas Matreux, Paula Aikkila, Bettina Scheu, Dieter Braun, and Christof B. Mast. Heat flows enrich prebiotic building blocks and enhance their reactivity. *Nature*, 628(8006):110–116, 2024.

[55] Hadi M. Fares, Alexander E. Marras, Jeffrey M. Ting, Matthew V. Tirrell, and Christine D. Keating. Impact of wet-dry cycling on the phase behavior and compartmentalization properties of complex coacervates. *Nature Communications*, 11(1):5423, 2020.

[56] Oliver R. Maguire, Iris B. A. Smokers, and Wilhelm T. S. Huck. A physicochemical orthophosphate cycle via a kinetically stable thermodynamically activated intermediate enables mild prebiotic phosphorylations. *Nature Communications*, 12(1):5517, 2021.

[57] Hannes Mutschler, Aniela Wochner, and Philipp Holliger. Freeze–thaw cycles as drivers of complex ribozyme assembly. *Nature Chemistry*, 7(6):502–508, 2015.

[58] Marcel J. Eleveld, Yannick Geiger, Juntian Wu, Armin Kiani, Gaël Schaeffer, and Sijbren Otto. Competitive Exclusion among Self-Replicating Molecules Curtails the Tendency of Chemistry to Diversify, 2024.

[59] Kai Liu, Alex Blokhuis, Chris van Ewijk, Armin Kiani, Juntian Wu, Wouter H. Roos, and Sijbren Otto. Light-driven eco-evolutionary dynamics in a synthetic replicator system. *Nature Chemistry*, 16(1):79–88, 2024.

[60] Matthias Morasch, Jonathan Liu, Christina F. Dirscherl, Alan Ianeselli, Alexandra Kühnlein, Kristian Le Vay, Philipp Schwintek, Saidul Islam, Mérina K. Corpinot, Bettina Scheu, Donald B. Dingwell, Petra Schwille, Hannes Mutschler, Matthew W. Powner, Christof B. Mast, and Dieter Braun. Heated gas bubbles enrich, crystallize, dry, phosphorylate and encapsulate prebiotic molecules. *Nature Chemistry*, 11(9):779–788, 2019.

[61] Ryo Mizuuchi, Alex Blokhuis, Lena Vincent, Philippe Nghe, Niles Lehman, and David Baum. Mineral surfaces select for longer RNA molecules. *Chemical Communications*, 55(14):2090–2093, 2019.

[62] Maria Kalambokidis and Michael Travisano. The eco-evolutionary origins of life. *Evolution*, 78(1):1–12, 2024.

[63] Carlos Mariscal, Ana Barahona, Nathanael Aubert-Kato, Arsev Umur Aydinoglu, Stuart Bartlett, María Luz Cárdenas, Kuhan Chandru, Carol Cleland, Benjamin T Cocanougher, Nathaniel Comfort, et al. Hidden concepts in the history and philosophy of origins-of-life studies: A workshop report. *Origins of Life and Evolution of Biospheres*, 49(3):111–145, 2019.

[64] Richard P Feynman. *The meaning of it all: Thoughts of a citizen-scientist*. Hachette UK, 2009.

[65] Petra Schwille, Joachim Spatz, Katharina Landfester, Eberhard Bodenschatz, Stephan Herminghaus, Victor Sourjik, Tobias J Erb, Philippe Bastiaens, Reinhard Lipowsky, Anthony Hyman, et al. Maxsynbio: avenues towards creating cells from the bottom up. *Angewandte Chemie International Edition*, 57(41):13382–13392, 2018.

[66] Sijbren Otto. An Approach to the De Novo Synthesis of Life. *Accounts of Chemical Research*, 55(2):145–155, 2022.

[67] Chenwang Tang, Lin Wang, Jing Sun, Guangda Chen, Junfeng Shen, Liang Wang, Ying Han, Jiren Luo, Zhiying Li, Pei Zhang, et al. Degradable living plastics programmed by engineered spores. *Nature Chemical Biology*, 21(7):1006–1011, 2025.

[68] Peng-Fei Xia, Guo-Chang Zhang, Berkley Walker, Seung-Oh Seo, Suryang Kwak, Jing-Jing Liu, Heejin Kim, Donald R Ort, Shu-Guang Wang, and Yong-Su Jin. Recycling carbon dioxide during xylose fermentation by engineered saccharomyces cerevisiae. *ACS synthetic biology*, 6(2):276–283, 2017.

[69] Paul Adamski, Marcel Eleveld, Ankush Sood, Ádám Kun, András Szilágyi, Tamás Czárán, Eörs Szathmáry, and Sijbren Otto. From self-replication to replicator systems en route to de novo life. *Nature Reviews Chemistry*, 4(8):386–403, 2020.

[70] Alyssa Adams, Hector Zenil, Paul CW Davies, and Sara Imari Walker. Formal definitions of unbounded evolution and innovation reveal universal mechanisms for open-ended evolution in dynamical systems. *Scientific reports*, 7(1):997, 2017.

[71] Jack W Szostak, David P Bartel, and P Luigi Luisi. Synthesizing life. *Nature*, 409(6818):387–390, 2001.

[72] Paulien Hogeweg and Nobuto Takeuchi. Multilevel selection in models of prebiotic evolution: compartments and spatial self-organization. *Origins of Life and Evolution of the Biosphere*, 33(4):375–403, 2003.

[73] Meniz Altay, Yigit Altay, and Sijbren Otto. Parasitic behavior of self-replicating molecules. *Angewandte Chemie International Edition*, 57(33):10564–10568, 2018.

[74] Shigeyoshi Matsumura, Ádám Kun, Michael Ryckelynck, Faith Coldren, András Szilágyi, Fabrice Jossinet, Christian Rick, Philippe Nghe, Eörs Szathmáry, and Andrew D Griffiths. Transient compartmentalization of rna replicators prevents extinction due to parasites. *Science*, 354(6317):1293–1296, 2016.

[75] Yohsuke Bansho, Taro Furubayashi, Norikazu Ichihashi, and Tetsuya Yomo. Host–parasite oscillation dynamics and evolution in a compartmentalized rna replication system. *Proceedings of the National Academy of Sciences*, 113(15):4045–4050, 2016.

[76] Yei-Chen Lai, Ziwei Liu, and Irene A Chen. Encapsulation of ribozymes inside model protocells leads to faster evolutionary adaptation. *Proceedings of the National Academy of Sciences*, 118(21):e2025054118, 2021.

[77] Siowling Soh, Michal Banaszak, Kristiana Kandere-Grzybowska, and Bartosz A Grzybowski. Why cells are microscopic: a transport-time perspective. *The Journal of Physical Chemistry Letters*, 4(6):861–865, 2013.

[78] Clifford P Brangwynne, Christian R Eckmann, David S Courson, Agata Rybarska, Carsten Hoege, Jöbin Gharakhani, Frank Jülicher, and Anthony A Hyman. Germline p granules are liquid droplets that localize by controlled dissolution/condensation. *Science*, 324(5935):1729–1732, 2009.

[79] Anatol W. Fritsch, Andrés F. Diaz-Delgadillo, Omar Adame-Arana, Carsten Hoege, Matthäus Mittasch, Moritz Kreysing, Mark Leaver, Anthony A. Hyman, Frank Jülicher, and Christoph A. Weber. Local thermodynamics govern formation and dissolution of caenorhabditis elegans p granule condensates. *Proceedings of the National Academy of Sciences*, 118(37):e2102772118, 2021.

[80] Anthony A. Hyman, Christoph A. Weber, and Frank Jülicher. Liquid-Liquid Phase Separation in Biology. *Annual Review of Cell and Developmental Biology*, 30(1):39–58, 2014.

[81] Salman F. Banani, Hyun O. Lee, Anthony A. Hyman, and Michael K. Rosen. Biomolecular condensates: Organizers of cellular biochemistry. *Nature Reviews Molecular Cell Biology*, 18(5):285–298, 2017.

[82] Simon Alberti, Amy Gladfelter, and Tanja Mittag. Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates. *Cell*, 176(3):419–434, 2019.

[83] Simon Alberti and Dorothee Dormann. Liquid–liquid phase separation in disease. *Annual review of genetics*, 53(1):171–194, 2019.

[84] Simon Alberti and Anthony A. Hyman. Biomolecular condensates at the nexus of cellular stress, protein aggregation disease and ageing. *Nature Reviews Molecular Cell Biology*, 22(3):196–213, 2021.

[85] Benedikt Rieß, Raphael Kurt Grötsch, and Job Boekhoven. The design of dissipative molecular assemblies driven by chemical reaction cycles. *Chem*, 6(3):552–578, 2020.

[86] David Zwicker. Chemically active droplets. *arXiv preprint arXiv:2407.09859*, 2024.

[87] Christoph A. Weber, David Zwicker, Frank Jülicher, and Chiu Fan Lee. Physics of active emulsions. *Reports on Progress in Physics*, 82(6):064601, 2019.

[88] Frank Jülicher and Christoph A. Weber. Droplet Physics and Intracellular Phase Separation. *Annual Review of Condensed Matter Physics*, 15:237–261, 2024.

[89] Marta Tena-Solsona, Jacqueline Janssen, Caren Wanzke, Fabian Schnitter, Hansol Park, Benedikt Rieß, Julianne M Gibbs, Christoph A Weber, and Job Boekhoven. Accelerated ripening in chemically fueled emulsions. *ChemSystemsChem*, 3(2):e2000034, 2021.

[90] Leonardo Demarchi, Andriy Goychuk, Ivan Maryshev, and Erwin Frey. Enzyme-enriched condensates show self-propulsion, positioning, and coexistence. *Physical Review Letters*, 130(12):128401, 2023.

[91] Jean David Wurtz and Chiu Fan Lee. Chemical-Reaction-Controlled Phase Separated Drops: Formation, Size Selection, and Coarsening. *Physical Review Letters*, 120(7):078102, 2018.

[92] David Zwicker, Rabea Seyboldt, Christoph A. Weber, Anthony A. Hyman, and Frank Jülicher. Growth and division of active droplets provides a model for protocells. *Nature Physics*, 13(4):408–413, 2017.

[93] David Zwicker, Anthony A. Hyman, and Frank Jülicher. Suppression of Ostwald ripening in active emulsions. *Physical Review E*, 92(1):012317, 2015.

[94] Antonia Winter, Yuhao Liu, Alexander Ziepke, George Dadunashvili, and Erwin Frey. Phase separation on deformable membranes: Interplay of mechanical coupling and dynamic surface geometry. *Physical Review E*, 111(4):044405, 2025.

[95] Florian Raßhofer, Simon Bauer, Alexander Ziepke, Ivan Maryshev, and Erwin Frey. Capillary wave formation in conserved active emulsions. *arXiv preprint arXiv:2505.20028*, 2025.

[96] Alexander M. Bergmann, Jonathan Bauermann, Giacomo Bartolucci, Carsten Donau, Michele Stasi, Anna-Lena Holtmannspötter, Frank Jülicher, Christoph A. Weber, and Job Boekhoven. Liquid spherical shells are a non-equilibrium steady state of active droplets. *Nature Communications*, 14(1):6552, 2023.

[97] Giacomo Bartolucci, Omar Adame-Arana, Xueping Zhao, and Christoph A. Weber. Controlling composition of coexisting phases via molecular transitions. *Biophysical Journal*, 120(21):4682–4697, 2021.

[98] Jonathan Bauermann, Giacomo Bartolucci, Job Boekhoven, Frank Jülicher, and Christoph A Weber. Critical transition between intensive and extensive active droplets. *arXiv preprint arXiv:2409.03629*, 2024.

[99] Pablo Zambrano, Xiaoyao Chen, Christine Kriebisch, Brigitte Kriebisch, Oleksii Zozulia, and Job Boekhoven. Chemically driven self-division in protocells models. *ChemRxiv.*, 2024.

[100] Aleksandr Ivanovich Oparin. *The Origin of Life*. Dover Publications Inc., New York, NY, 2 edition, 1938.

[101] N Amy Yewdall, Alain AM André, Tiemei Lu, and Evan Spruijt. Coacervates as models of membraneless organelles. *Current Opinion in Colloid & Interface Science*, 52:101416, 2021.

[102] Philip W Anderson. More is different: broken symmetry and the nature of the hierarchical structure of science. *Science*, 177(4047):393–396, 1972.

[103] Robert H Swendsen. *An introduction to statistical mechanics and thermodynamics*. Oxford University Press, 2020.

[104] Rana A Fine and Frank J Millero. Compressibility of water as a function of temperature and pressure. *The Journal of Chemical Physics*, 59(10):5529–5536, 1973.

[105] Mehran Kardar. *Statistical physics of particles*. Cambridge University Press, 2007.

[106] Paul J Flory. Thermodynamics of high polymer solutions. *The Journal of Chemical Physics*, 10(1):51–61, 1942.

[107] Maurice L Huggins. The viscosity of dilute solutions of long-chain molecules. iv. dependence on concentration. *Journal of the American Chemical Society*, 64(11):2716–2718, 1942.

[108] Pierre C Hohenberg and Bertrand I Halperin. Theory of dynamic critical phenomena. *Reviews of Modern Physics*, 49(3):435, 1977.

[109] Lars Onsager. Reciprocal relations in irreversible processes. i. *Physical Review*, 37(4):405, 1931.

[110] Lev Davidovich Landau and Evgenii Mikhailovich Lifshitz. *Statistical Physics: Volume 5*. Elsevier, 2013.

[111] Sybren Ruurds De Groot and Peter Mazur. *Non-equilibrium thermodynamics*. Courier Corporation, 2013.

[112] Ilya M Lifshitz and Vitaly V Slyozov. The kinetics of precipitation from supersaturated solid solutions. *Journal of physics and chemistry of solids*, 19(1-2):35–50, 1961.

[113] Jonathan Bauermann, Sudarshana Laha, Patrick M. McCall, Frank Jülicher, and Christoph A. Weber. Chemical Kinetics and Mass Action in Coexisting Phases. *Journal of the American Chemical Society*, 144(42):19294–19304, 2022.

[114] Michael Rubinstein and Ralph H Colby. *Polymer physics*. Oxford University Press, 2003.

[115] Jonathan Bauermann. *Organization of chemical reactions by phase separation*. PhD thesis, Technische Universität Dresden, 2022.

[116] Francesco Avanzini, Emanuele Penocchio, Gianmaria Falasco, and Massimiliano Esposito. Nonequilibrium thermodynamics of non-ideal chemical reaction networks. *The Journal of Chemical Physics*, 154(9):094114, 2021.

[117] NG Van Kampen. Nonlinear irreversible processes. *Physica*, 67(1):1–22, 1973.

[118] Frank Jülicher, Armand Ajdari, and Jacques Prost. Modeling molecular motors. *Reviews of Modern Physics*, 69(4):1269–1282, 1997.

[119] Jonathan Bauermann, Christoph A. Weber, and Frank Jülicher. Energy and Matter Supply for Active Droplets. *Annalen der Physik*, 534(9):2200132, 2022.

[120] Sudarshana Laha, Jonathan Bauermann, Frank Jülicher, Thomas C. T. Michaels, and Christoph A. Weber. Chemical reactions regulated by phase-separated condensates. *Physical Review Research*, 6(4):043092, 2024.

[121] Andriy Goychuk, Leonardo Demarchi, Ivan Maryshev, and Erwin Frey. Self-consistent sharp interface theory of active condensate dynamics. *Physical Review Research*, 6(3):033082, 2024.

[122] Rob Phillips, Jane Kondev, Julie Theriot, and Hernan Garcia. *Physical biology of the cell*. Garland Science, 2012.

[123] Sidney Becker, Christina Schneider, Hidenori Okamura, Antony Crisp, Tynchtyk Amatov, Milan Dejmek, and Thomas Carell. Wet-dry cycles enable the parallel origin of canonical and non-canonical nucleosides by continuous synthesis. *Nature Communications*, 9(1):163, 2018.

[124] Emre Tekin, Annalena Salditt, Philipp Schwintek, Sreekar Wunnava, Juliette Langlais, James Saenz, Dora Tang, Petra Schwille, Christof Mast, and Dieter Braun. Prebiotic Foam Environments to Oligomerize and Accumulate RNA. *ChemBioChem*, 23(24):e202200423, 2022.

[125] Avinash Vicholous Dass, Sreekar Wunnava, Juliette Langlais, Beatriz von der Esch, Maik Krusche, Lennard Ufer, Nico Chrisam, Romeo C. A. Dubini, Florian Gartner, and Severin Angerpointner, et al. RNA Oligomerisation without Added Catalyst from 2,3-Cyclic Nucleotides by Drying at Air-Water Interfaces. *ChemSystemsChem*, 5(1):e202200026, 2023.

[126] Richard I. Masel. *Chemical Kinetics and Catalysis*. Wiley, 2001.

[127] Sebastian Werner, Marco Haumann, and Peter Wasserscheid. Ionic Liquids in Chemical Engineering. *Annual Review of Chemical and Biomolecular Engineering*, 1(1):203–230, 2010.

[128] Hans-Ulrich Blaser and Martin Studer. The role of catalysis for the clean production of fine chemicals. *Applied Catalysis A: General*, 189(2):191–204, 1999.

[129] Paul A. Srere. Complexes of Sequential Metabolic Enzymes. *Annual Review of Biochemistry*, 56(1):89–124, 1987.

[130] Manfred Eigen. New looks and outlooks on physical enzymology. *Quarterly Reviews of Biophysics*, 1(1):3–33, 1968.

[131] Aleksandr Ivanovich Oparin. The Origin of Life and the Origin of Enzymes. *Advances in enzymology and related areas of molecular biology*, 27:347–80, 1965.

[132] Sam Hay, Linus O. Johannissen, Michael J. Sutcliffe, and Nigel S. Scrutton. Barrier Compression and Its Contribution to Both Classical and Quantum Mechanical Aspects of Enzyme Catalysis. *Biophysical Journal*, 98(1):121–128, 2010.

[133] Philipp Baaske, Franz M. Weinert, Stefan Duhr, Kono H. Lemke, Michael J. Russell, and Dieter Braun. Extreme accumulation of nucleotides in simulated hydrothermal pore systems. *Proceedings of the National Academy of Sciences*, 104(22):9346–9351, 2007.

[134] Giacomo Bartolucci, Adriana Calaça Serrão, Philipp Schwintek, Alexandra Kühnlein, Yash Rana, Philipp Janto, Dorothea Hofer, Christof B. Mast, Dieter Braun, and Christoph A. Weber. Sequence self-selection by cyclic phase separation. *Proceedings of the National Academy of Sciences*, 120(43):e2218876120, 2023.

[135] G. Nicolis. Dissipative systems. *Rep. Prog. Phys.*, 49(8):873, 1986.

[136] Paul G. Higgs. The effect of limited diffusion and wet–dry cycling on reversible polymerization reactions: Implications for prebiotic synthesis of nucleic acids. *Life*, 6(2):24, 2016.

[137] Jay G. Forsythe, Sheng-Sheng Yu, Irena Mamajanov, Martha A. Grover, Ramanarayanan Krishnamurthy, Facundo M. Fernández, and Nicholas V. Hud. Ester-Mediated Amide Bond Formation Driven by Wet–Dry Cycles: A Possible Path to Polypeptides on the Prebiotic Earth. *Angewandte Chemie International Edition*, 54(34):9871–9875, 2015.

[138] Irena Mamajanov, Patrick J. MacDonald, Jingya Ying, Daniel M. Duncanson, Garrett R. Dowdy, Chelsea A. Walker, Aaron E. Engelhart, Facundo M. Fernández, Martha A. Grover, and V. Hud, Nicholas et al. Ester Formation and Hydrolysis

during Wet–Dry Cycles: Generation of Far-from-Equilibrium Polymers in a Model Prebiotic Reaction. *Macromolecules*, 47(4):1334–1343, 2014.

[139] Arthur L. Weber. Thermal synthesis and hydrolysis of polyglyceric acid. *Origins of Life and Evolution of the Biosphere*, 19(1):7–19, 1989.

[140] Marc Rodriguez-Garcia, Andrew J. Surman, Geoffrey J. T. Cooper, Irene Suárez-Marina, Zied Hosni, Michael P. Lee, and Leroy Cronin. Formation of oligopeptides in high yield under simple programmable conditions. *Nature Communications*, 6(1):8385, 2015.

[141] Laura Da Silva, Marie-Christine Maurel, and David Deamer. Salt-Promoted Synthesis of RNA-like Molecules in Simulated Hydrothermal Conditions. *Journal of Molecular Evolution*, 80(2):86–97, 2015.

[142] Sudha Rajamani, Alexander Vlassov, Seico Benner, Amy Coombs, Felix Olasagasti, and David Deamer. Lipid-assisted Synthesis of RNA-like Polymers from Mononucleotides. *Origins of Life and Evolution of Biospheres*, 38(1):57–74, 2008.

[143] Alexander N. Semenov and Michael Rubinstein. Thermoreversible Gelation in Solutions of Associative Polymers. 1. Statics. *Macromolecules*, 31(4):1373–1385, 1998.

[144] Alison Emslie Lewis. Review of metal sulphide precipitation. *Hydrometallurgy*, 104(2):222–234, 2010.

[145] Vishal Agarwal and Baron Peters. Solute precipitate nucleation: A review of theory and simulation advances. *Advances in Chemical Physics: Volume 155*, pages 97–160, 2014.

[146] Giacomo Carrea, Stefano Colonna, David R. Kelly, Antonio Lazcano, Gianluca Ottolina, and Stanley M. Roberts. Polyamino acids as synthetic enzymes: mechanism, applications and relevance to prebiotic catalysis. *Trends in Biotechnology*, 23(10):507–513, 2005.

[147] Luca Schulz, Zhijun Guo, Jan Zarzycki, Wieland Steinchen, Jan M. Schuller, Thomas Heimerl, Simone Prinz, Oliver Mueller-Cajar, Tobias J. Erb, and Georg K. A. Hochberg. Evolution of increased complexity and specificity at the dawn of form I Rubiscos. *Science*, 378(6616):155–160, 2022.

[148] Omar Adame-Arana, Christoph A. Weber, Vasily Zaburdaev, Jacques Prost, and Frank Jülicher. Liquid phase separation controlled by pH. *Biophysical Journal*, 119(8):1590–1605, 2020.

[149] Philip Loche, Douwe J. Bonthuis, and Roland R. Netz. Molecular dynamics simulations of the evaporation of hydrated ions from aqueous solution. *Communications Chemistry*, 5(1):1–8, 2022.

[150] Kazuo Hisatake, Satoko Tanaka, and Youko Aizawa. Evaporation rate of water in a vessel. *Journal of Applied Physics*, 73(11):7395–7401, 1993.

[151] Matthias Heinen and Jadran Vrabec. Evaporation sampled by stationary molecular dynamics simulation. *The Journal of Chemical Physics*, 151(4):044704, 2019.

[152] T. Y. Xiong and M. C. Yuen. Evaporation of a liquid droplet on a hot plate. *International Journal of Heat and Mass Transfer*, 34(7):1881–1894, 1991.

[153] D C Clary. Fast Chemical Reactions: Theory Challenges Experiment. *Annual Review of Physical Chemistry*, 41(1):61–90, 1990.

[154] Chong Meng Kok and Alfred Rudin. Prediction of Flory–Huggins interaction parameters from intrinsic viscosities. *Journal of Applied Polymer Science*, 27(2):353–362, 1982.

[155] Travis H. Russell, Brian J. Edwards, and Bamin Khomami. Characterization of the Flory-Huggins interaction parameter of polymer thermodynamics. *EPL*, 108(6):66003, 2015.

[156] L. A. Underkofler, R. R. Barton, and S. S. Rennert. Production of Microbial Enzymes and Their Applications. *Appl Microbiol*, 6(3):212–221, 1958.

[157] A Zaks and A M Klibanov. The effect of water on enzyme action in organic media. *Journal of Biological Chemistry*, 263(17):8017–8021, 1988.

[158] Alex P. S. Brogan, Kamendra P. Sharma, Adam W. Perriman, and Stephen Mann. Enzyme activity in liquid lipase melts as a step towards solvent-free biology at 150°c. *Nature Communications*, 5(1):5058, 2014.

[159] Volkhard Helms and Rebecca C. Wade. Hydration energy landscape of the active site cavity in cytochrome p450cam. *Proteins: Structure, Function, and Bioinformatics*, 32(3):381–396, 1998.

[160] Alexander M. Klibanov. Why are enzymes less active in organic solvents than in water? *Trends in Biotechnology*, 15(3):97–101, 1997.

[161] Armen Y. Mulkidjanian, Andrew Yu. Bychkov, Daria V. Dibrova, Michael Y. Galperin, and Eugene V. Koonin. Origin of first cells at terrestrial, anoxic geothermal fields. *Proceedings of the National Academy of Sciences*, 109(14), 2012.

[162] Bruce Damer and David Deamer. Coupled Phases and Combinatorial Selection in Fluctuating Hydrothermal Pools: A Scenario to Guide Experimental Approaches to the Origin of Cellular Life. *Life*, 5(1):872–887, 2015.

[163] Thomas D. Campbell, Rio Febrian, Jack T. McCarthy, Holly E. Kleinschmidt, Jay G. Forsythe, and Paul J. Bracher. Prebiotic condensation through wet–dry cycling regulated by deliquescence. *Nature Communications*, 10(1):4508, 2019.

[164] Alexander V. Vlassov, Brian H. Johnston, Laura F. Landweber, and Sergei A. Kazakov. Ligation activity of fragmented ribozymes in frozen solution: implications for the RNA world. *Nucleic Acids Research*, 32(9):2966–2974, 2004.

[165] Kristian Le Vay, Emilie Yeonwha Song, Basusree Ghosh, T.-Y. Dora Tang, and Hannes Mutschler. Enhanced Ribozyme-Catalyzed Recombination and Oligonucleotide Assembly in Peptide-RNA Condensates. *Angewandte Chemie International Edition*, 60(50):26096–26104, 2021.

[166] Alan Ianeselli, Damla Tetiker, Julian Stein, Alexandra Kühnlein, Christof B. Mast, Dieter Braun, and T.-Y. Dora Tang. Non-equilibrium conditions inside rock pores drive fission, maintenance and selection of coacervate protocells. *Nature Chemistry*, 14(1):32–39, 2022.

[167] T. Matreux, K. Le Vay, A. Schmid, P. Aikkila, L. Belohlavek, A. Z. Çalışkanoğlu, E. Salibi, A. Kühnlein, C. Springsklee, and B. Scheu, et al. Heat flows in rock cracks naturally optimize salt compositions for ribozymes. *Nature Chemistry*, 13(11):1038–1045, 2021.

[168] W. A. Zehring, D. A. Wheeler, P. Reddy, R. J. Konopka, C. P. Kyriacou, M. Rosbash, and J. C. Hall. P-element transformation with period locus DNA restores rhythmicity to mutant, arrhythmic Drosophila melanogaster. *Cell*, 39(2 Pt 1):369–376, 1984.

[169] T. A. Bargiello, F. R. Jackson, and M. W. Young. Restoration of circadian behavioural rhythms by gene transfer in Drosophila. *Nature*, 312(5996):752–754, 1985.

[170] Jay C. Dunlap. Molecular Bases for Circadian Clocks. *Cell*, 96(2):271–290, 1999.

[171] Albert Goldbeter. *Biochemical Oscillations and Cellular Rhythms: The Molecular Bases of Periodic and Chaotic Behaviour*. Cambridge University Press, Cambridge, 1996.

[172] Michael B Elowitz and Stanislas Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–338, 2000.

[173] Jesse Stricker, Scott Cookson, Matthew R Bennett, William H Mather, Lev S Tsimring, and Jeff Hasty. A fast, robust and tunable synthetic gene oscillator. *Nature*, 456(7221):516–519, 2008.

[174] William C. Bray. A PERIODIC REACTION IN HOMOGENEOUS SOLUTION AND ITS RELATION TO CATALYSIS. *Journal of the American Chemical Society*, 43(6):1262–1267, 1921.

[175] Anatol M. Zhabotinsky. A history of chemical oscillations and waves. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 1(4):379–386, 1991.

[176] Alexander Pechenkin. B P Belousov and his reaction. *Journal of Biosciences*, 34(3):365–371, 2009.

[177] BP. BELOUSOV. A periodic reaction and its mechanism. *Ref. Radiats. Med.*, 1958.

[178] I. Prigogine and G. Nicolis. On Symmetry-Breaking Instabilities in Dissipative Systems. *The Journal of Chemical Physics*, 46(9):3542–3550, 1967.

[179] I. Prigogine and R. Lefever. Symmetry Breaking Instabilities in Dissipative Systems. II. *The Journal of Chemical Physics*, 48(4):1695–1700, 1968.

[180] J. E. Marsden and M. McCracken. *The Hopf Bifurcation and Its Applications*. Springer Science & Business Media, 2012.

[181] Tomás Caraballo, José A. Langa, and James C. Robinson. A stochastic pitchfork bifurcation in a reaction-diffusion equation. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 457(2013):2041–2061, 2001.

[182] Mark C Cross and Pierre C Hohenberg. Pattern formation outside of equilibrium. *Reviews of Modern Physics*, 65(3):851, 1993.

[183] Alan Mathison Turing. The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 237(641):37–72, orig. 1952, reprint 1997.

[184] Irving R. Epstein and Kenneth Showalter. Nonlinear Chemical Dynamics: Oscillations, Patterns, and Chaos. *The Journal of Physical Chemistry*, 100(31):13132–13147, 1996.

[185] Chengjie Luo and David Zwicker. Influence of physical interactions on spatiotemporal patterns. *Physical Review E*, 108(3):034206, 2023.

[186] Iris B. A. Smokers, Brent S. Visser, Wojciech P. Lipiński, Karina K. Nakashima, and Evan Spruijt. Phase-Separated Droplets Can Direct the Kinetics of Chemical Reactions Including Polymerization, Self-Replication and Oscillating Networks. *ChemSystemsChem*, 7(1):e202400056, 2025.

[187] Feng-Shu Hsieh, Duy P. M. Nguyen, Mathias S. Heltberg, Chia-Chou Wu, Yi-Chen Lee, Mogens H. Jensen, and Sheng-hong Chen. Plausible, robust biological oscillations through allelic buffering. *Cell Systems*, 15(11):1018–1032.e12, 2024.

[188] Mathias S. Heltberg, Alessandra Lucchetti, Feng-Shu Hsieh, Duy Pham Minh Nguyen, Sheng-hong Chen, and Mogens H. Jensen. Enhanced DNA repair through droplet formation and p53 oscillations. *Cell*, 185(23):4394–4408.e10, 2022.

[189] Mathias S. Heltberg, Yuanxu Jiang, Yingying Fan, Zhibo Zhang, Malthe S. Nordentoft, Wei Lin, Long Qian, Qi Ouyang, Mogens H. Jensen, and Ping Wei. Coupled oscillator cooperativity as a control mechanism in chronobiology. *Cell Systems*, 14(5):382–391.e5, 2023.

[190] Job Boekhoven, Aurelie M. Brizard, Krishna N. K. Kowlgi, Ger J. M. Koper, Rienk Eelkema, and Jan H. van Esch. Dissipative Self-Assembly of a Molecular Gelator by Using a Chemical Fuel. *Angewandte Chemie*, 122(28):4935–4938, 2010.

[191] Job Boekhoven, Wouter E. Hendriksen, Ger J. M. Koper, Rienk Eelkema, and Jan H. van Esch. Transient assembly of active materials fueled by a chemical reaction. *Science*, 349(6252):1075–1079, 2015.

[192] Scientific Image and Illustration Software | BioRender, 01.10.2023. https://www.biorender.com/.

[193] Xi Chen, Jens-Uwe Sommer, and Tyler Harmon. Surviving the first famine: fuel-driven polymerization-induced phase separation stabilizes protocells. *ChemRxiv*, 2025.

[194] Michael Cross and Henry Greenside. *Pattern formation and dynamics in nonequilibrium systems*. Cambridge University Press, 2009.

[195] Steven H Strogatz. *Nonlinear dynamics and chaos with student solutions manual: With applications to physics, biology, chemistry, and engineering*. CRC press, 2018.

[196] Ivar S. Haugerud. Github: Emergent behavior in non-dilute reacting mixtures away from equilibrium: Supplementary movies. `https://github.com/ivarhaugerud/Emergent-behavior-in-non-dilute-reacting-mixtures-away-from-equilibrium`.

[197] F. Jülicher and J. Prost. Generic theory of colloidal transport. *The European Physical Journal E*, 29(1):27–36, 2009.

[198] Boualem Boashash. Estimating and interpreting the instantaneous frequency of a signal. i. fundamentals. *Proceedings of the IEEE*, 80(4):520–538, 2002.

[199] Sergey N. Semenov, Albert S. Y. Wong, R. Martijn van der Made, Sjoerd G. J. Postma, Joost Groen, Hendrik W. H. van Roekel, Tom F. A. de Greef, and Wilhelm T. S. Huck. Rational design of functional and tunable oscillating enzymatic networks. *Nature Chemistry*, 7(2):160–165, 2015.

[200] Yun-Yi Luo, Jun-Jun Wu, and Yan-Mei Li. Regulation of liquid–liquid phase separation with focus on post-translational modifications. *Chemical communications*, 57(98):13275–13287, 2021.

[201] Carsten Donau and Job Boekhoven. The chemistry of chemically fueled droplets. *Trends in Chemistry*, 5(1):45–60, 2023.

[202] Simone M Poprawa, Michele Stasi, Brigitte AK Kriebisch, Monika Wenisch, Judit Sastre, and Job Boekhoven. Active droplets through enzyme-free, dynamic phosphorylation. *Nature Communications*, 15(1):4204, 2024.

[203] Judit Sastre, Advait Thatte, Alexander M Bergmann, Michele Stasi, Marta Tena-Solsona, Christoph A Weber, and Job Boekhoven. Size control and oscillations of active droplets in synthetic cells. *Nature Communications*, 16(1):2003, 2025.

[204] Alan Ianeselli, Annalena Salditt, Christof Mast, Barbara Ercolano, Corinna L. Kufner, Bettina Scheu, and Dieter Braun. Physical non-equilibria for prebiotic nucleic acid chemistry. *Nature Reviews Physics*, 5(3):185–195, 2023.

[205] Patrick W. Kudella, Alexei V. Tkachenko, Annalena Salditt, Sergei Maslov, and Dieter Braun. Structured sequences emerge from random pool when replicated by templated ligation. *Proceedings of the National Academy of Sciences*, 118(8):e2018830118, 2021.

[206] Adriana Calaça Serrão, Felix T Dänekamp, Zsófia Meggyesi, and Dieter Braun. Replication elongates short DNA, reduces sequence bias and develops trimer structure. *Nucleic Acids Research*, 52(3):1290–1297, 2024.

[207] Adriana Calaça Serrão, Sreekar Wunnava, Avinash V. Dass, Lennard Ufer, Philipp Schwintek, Christof B. Mast, and Dieter Braun. High-Fidelity RNA Copying via 2-,3-Cyclic Phosphate Ligation. *Journal of the American Chemical Society*, 146(13):8887–8894, 2024.

[208] Taishan Hu, Nilesh Chitnis, Dimitri Monos, and Anh Dinh. Next-generation sequencing technologies: An overview. *Human Immunology*, 82(11):801–811, 2021.

[209] Yusuke Sato, Tetsuro Sakamoto, and Masahiro Takinoue. Sequence-based engineering of dynamic functions of micrometer-sized DNA droplets. *Science Advances*, 6(23):eaba3471, 2020.

[210] Zhongyang Xing, Alessio Caciagli, Tianyang Cao, Iliya Stoev, Mykolas Zupkauskas, Thomas O'Neill, Tobias Wenzel, Robin Lamboll, Dongsheng Liu, and Erika Eiser. Microrheology of DNA hydrogels. *Proceedings of the National Academy of Sciences*, 115(32):8137–8142, 2018.

[211] Michi Nakata, Giuliano Zanchetta, Brandon D. Chapman, Christopher D. Jones, Julie O. Cross, Ronald Pindak, Tommaso Bellini, and Noel A. Clark. End-to-End Stacking and Liquid Crystal Condensation of 6– to 20–Base Pair DNA Duplexes. *Science*, 318(5854):1276–1279, 2007.

[212] G. Zanchetta, M. Nakata, M. Buscaglia, N. A. Clark, and T. Bellini. Liquid crystal ordering of DNA and RNA oligomers with partially overlapping sequences. *Journal of Physics: Condensed Matter*, 20(49):494214, 2008.

[213] William M. Aumiller, Fatma Pir Cakmak, Bradley W. Davis, and Christine D. Keating. RNA-Based Coacervates as a Model for Membraneless Organelles: Formation, Properties, and Interfacial Liposome Assembly. *Langmuir: the ACS journal of surfaces and colloids*, 32(39):10042–10053, 2016.

[214] Diana M. Mitrea and Richard W. Kriwacki. Phase separation in biology; functional organization of a higher order. *Cell Communication and Signaling*, 14(1):1, 2016.

[215] David Zwicker. The intertwined physics of active chemical reactions and phase separation. *Current Opinion in Colloid & Interface Science*, 61:101606, 2022.

[216] Tobias Göppel, Joachim H Rosenberger, Bernhard Altaner, and Ulrich Gerland. Thermodynamic and kinetic sequence selection in enzyme-free polymer self-assembly inside a non-equilibrium rna reactor. *Life*, 12(4):567, 2022.

[217] Fabian Kilchherr, Christian Wachauf, Benjamin Pelz, Matthias Rief, Martin Zacharias, and Hendrik Dietz. Single-molecule dissection of stacking forces in dna. *Science*, 353(6304):aaf5508, 2016.

[218] Siobhan E. Toal, Nina Kubatova, Christian Richter, Verena Linhard, Harald Schwalbe, and Reinhard Schweitzer-Stenner. Randomizing the Unfolded State of Peptides (and Proteins) by Nearest Neighbor Interactions between Unlike Residues. *Chemistry A European Journal*, 21(13):5173–5192, 2015.

[219] Yan Mu and Yi Qin Gao. Effects of hydrophobic and dipole-dipole interactions on the conformational transitions of a model polypeptide. *The Journal of Chemical Physics*, 127(10):105102, 2007.

[220] Bercem Dutagaci, Grzegorz Nawrocki, Joyce Goodluck, Ali Akbar Ashkarran, Charles G Hoogstraten, Lisa J Lapidus, and Michael Feig. Charge-driven condensation of rna and proteins suggests broad role of phase separation in cytoplasmic environments. *Elife*, 10:e64004, 2021.

[221] Ankur Jain and Ronald D Vale. Rna phase transitions in repeat expansion disorders. *Nature*, 546(7657):243–247, 2017.

[222] Christine Roden and Amy S Gladfelter. Rna contributions to the form and function of biomolecular condensates. *Nature Reviews Molecular Cell Biology*, 22(3):183–195, 2021.

[223] Devin Tauber, Gabriel Tauber, and Roy Parker. Mechanisms and regulation of rna condensation in rnp granule formation. *Trends in biochemical sciences*, 45(9):764–778, 2020.

[224] Bernat Corominas-Murtra, Luís F Seoane, and Ricard Solé. Zipf's law, unbounded complexity and open-ended evolution. *Journal of the Royal Society Interface*, 15(149):20180395, 2018.

[225] George Kingsley Zipf. *Selected Studies of the Principle of Relative Frequency in Language*. Harvard University Press, Cambridge, MA and London, England, 1932.

[226] Marcello Buiatti, C Acquisti, G Mersi, Patrizia Bogani, and M Buiatti Jr. The biological meanings of dna correlations. In *Fractals in Biology and Medicine*, pages 235–245. Springer, 2002.

[227] Martin Ackermann and Lin Chao. Dna sequences shaped by selection for stability. *PLoS Genetics*, 2(2):e22, 2006.

[228] Albino Bacolla, Jacquelynn E Larson, Jack R Collins, Jian Li, Aleksandar Milosavljevic, Peter D Stenson, David N Cooper, and Robert D Wells. Abundance and length of simple repeats in vertebrate genomes are determined by their structural properties. *Genome research*, 18(10):1545–1553, 2008.

[229] Adam Ciesiolka, Magdalena Jazurek, Karolina Drazkowska, and Wlodzimierz J Krzyzosiak. Structural characteristics of simple rna repeats associated with disease and their deleterious protein interactions. *Frontiers in cellular neuroscience*, 11:97, 2017.

[230] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948.

[231] Arieh Ben-Naim and Diego Casadei. *Modern thermodynamics*. World Scientific Hackensack, NJ, USA, 2017.

[232] A. Lempel and J. Ziv. On the complexity of finite sequences. *IEEE Transactions on Information Theory*, 22(1):75–81, 1976.

[233] Ming Li, Paul Vitányi, et al. *An introduction to Kolmogorov complexity and its applications*, volume 3. Springer, 2008.

[234] Terry A. Welch. A technique for high-performance data compression. *Computer*, 17(06):8–19, 1984.

[235] Joachim H. Rosenberger, Tobias Göppel, Patrick W. Kudella, Dieter Braun, Ulrich Gerland, and Bernhard Altaner. Self-Assembly of Informational Polymers by Templated Ligation. *Physical Review X*, 11(3):031055, 2021.

[236] Karina Kinuyo Nakashima, Fatma Zohra Mihoubi, Jagandeep Saraya, Kieran Russell, Fidan Rahmatova, James Robinson, Maria Julia Maristany, Jan Huertas, Roger Rubio-Sánchez, Rosana Collepardo-Guevara, et al. Compositional and functional diversity of minimal primitive coacervates in a nucleic acid-peptide world. *ChemRxiv*, 2025.

[237] Itay Budin and Jack W. Szostak. Expanding roles for diverse physical phenomena during the origin of life. *Annual Review of Biophysics*, 39:245–263, 2010.

[238] Daniel T Gillespie. Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, 81(25):2340–2361, 1977.

[239] J Nicholas Housby, Sigrídur H Thorbjarnardóttir, Zophonías O Jónsson, and Edwin M Southern. Optimised ligation of oligonucleotides by thermal ligases: comparison of thermus scotoductus and rhodothermus marinus dna ligases to other thermophilic ligases. *Nucleic acids research*, 28(3):e10–e10, 2000.

[240] Tianbing Xia, John SantaLucia Jr, Mark E Burkard, Ryszard Kierzek, Susan J Schroeder, Xiaoqi Jiao, Christopher Cox, and Douglas H Turner. Thermodynamic parameters for an expanded nearest-neighbor model for formation of rna duplexes with watson- crick base pairs. *Biochemistry*, 37(42):14719–14735, 1998.

[241] Kenneth D James and Andrew D Ellington. Surprising fidelity of template-directed chemical ligation of oligonucleotides. *Chemistry & Biology*, 4(8):595–605, 1997.

[242] Benedikt Obermayer, Hubert Krammer, Dieter Braun, and Ulrich Gerland. Emergence of information transmission in a prebiotic rna reactor. *Physical Review Letters*, 107(1):018101, 2011.

[243] Lijun Zhou, Dian Ding, and Jack W Szostak. The virtual circular genome model for primordial rna replication. *Rna*, 27(1):1–11, 2021.

[244] Elizaveta Guseva, Ronald N Zuckermann, and Ken A Dill. Foldamer hypothesis for the growth and sequence differentiation of prebiotic polymers. *Proceedings of the National Academy of Sciences*, 114(36):E7460–E7468, 2017.

[245] Brittany Rauzan, Elizabeth McMichael, Rachel Cave, Lesley R Sevcik, Kara Ostrosky, Elisabeth Whitman, Rachel Stegemann, Audra L Sinclair, Martin J Serra, and Alice A Deckert. Kinetics and thermodynamics of dna, rna, and hybrid duplex formation. *Biochemistry*, 52(5):765–772, 2013.

[246] VI Levenshtein. Binary codes capable of correcting deletions, insertions, and reversals probl. *Inf. Transm*, 1:8–17, 1965.

[247] Shoichi Toyabe and Dieter Braun. Cooperative ligation breaks sequence symmetry and stabilizes early molecular replication. *Physical Review X*, 9(1):011056, 2019.

[248] Ronny Lorenz, Stephan H Bernhart, Christian Höner zu Siederdissen, Hakim Tafer, Christoph Flamm, Peter F Stadler, and Ivo L Hofacker. Viennarna package 2.0. *Algorithms for molecular biology*, 6(1):26, 2011.

[249] Ivo L Hofacker, Walter Fontana, Peter F Stadler, L Sebastian Bonhoeffer, Manfred Tacker, and Peter Schuster. Fast folding and comparison of rna secondary structures. *Monatshefte für Chemie/Chemical Monthly*, 125(2):167–188, 1994.

[250] The RNAcentral Consortium et al. Rnacentral: a comprehensive database of noncoding rna sequences. *Nucleic acids research*, 45(D1):D128–D134, 2017.

[251] Federico Caimi, Juliette Langlais, Francesco Fontana, Sreekar Wunnava, Tommaso Bellini, Dieter Braun, and Tommaso P Fraccia. High-yield prebiotic polymerization of 2, 3-cyclic nucleotides under wet–dry cycling. *ACS Central Science*, 2025.

[252] Roberto Barrio, Stephen Coombes, Mathieu Desroches, Flavio Fenton, Stefan Luther, and Esther Pueyo. Excitable dynamics in neural and cardiac systems, 2020.

[253] Bertil Hille. Ionic channels in excitable membranes. current problems and biophysical approaches. *Biophysical journal*, 22(2):283–294, 1978.

[254] Chengjie Luo, Yicheng Qiang, and David Zwicker. Beyond pairwise: Higher-order physical interactions affect phase separation in multicomponent liquids. *Physical Review Research*, 6(3):033002, 2024.

[255] Michael S Wertheim. Fluids with highly directional attractive forces. i. statistical thermodynamics. *Journal of statistical physics*, 35(1):19–34, 1984.

[256] Michael S Wertheim. Fluids with highly directional attractive forces. ii. thermodynamic perturbation theory and integral equations. *Journal of statistical physics*, 35(1):35–47, 1984.

[257] Walter G Chapman, George Jackson, and Keith E Gubbins. Phase equilibria of associating fluids: chain molecules with multiple bonding sites. *Molecular Physics*, 65(5):1057–1079, 1988.

[258] JM Tavares, CS Dias, NAM Araújo, and MM Telo da Gama. Dynamics of patchy particles in and out of equilibrium. *The Journal of Physical Chemistry B*, 122(13):3514–3518, 2017.

[259] Patrice Paricaud, Szabolcs Varga, and George Jackson. Study of the demixing transition in model athermal mixtures of colloids and flexible self-excluding polymers using the thermodynamic perturbation theory of wertheim. *The Journal of Chemical Physics*, 118(18):8525–8536, 2003.

[260] Werner Kuhn. Über die gestalt fadenförmiger moleküle in lösungen. *Kolloid-Zeitschrift*, 68(1):2–15, 1934.

[261] Joseph Malcolm Schaeffer. *Stochastic simulation of the kinetics of multiple interacting nucleic acid strands*. PhD thesis, Citeseer, 2013.

[262] Joseph Malcolm Schaeffer, Chris Thachuk, and Erik Winfree. Stochastic simulation of the kinetics of multiple interacting nucleic acid strands. In *International Workshop on DNA-Based Computers*, pages 194–211. Springer, 2015.

[263] Aditya Sengar, Thomas E Ouldridge, Oliver Henrich, Lorenzo Rovigatti, and Petr Šulc. A primer on the oxdna model of dna: when to use it, how to simulate it and how to interpret the results. *Frontiers in Molecular Biosciences*, 8:693710, 2021.

[264] Hannah K Wayment-Steele, Wipapat Kladwang, Alexandra I Strom, Jeehyung Lee, Adrien Treuille, Alex Becka, Eterna Participants, and Rhiju Das. Rna secondary structure packages evaluated and improved by high-throughput experiments. *Nature methods*, 19(10):1234–1242, 2022.

[265] Jaswinder Singh, Jack Hanson, Kuldip Paliwal, and Yaoqi Zhou. Rna secondary structure prediction using an ensemble of two-dimensional deep neural networks and transfer learning. *Nature communications*, 10(1):5407, 2019.

[266] Atul K Rangadurai, Lisa Ruetz, Rashik Ahmed, Kristen Lo, Martin Tollinger, Julie D Forman-Kay, Christoph Kreutz, and Lewis E Kay. Phase separation modulates the thermodynamics and kinetics of rna hybridization. *Journal of the American Chemical Society*, 146(29):19686–19689, 2024.

[267] Jerelle A Joseph, Jorge R Espinosa, Ignacio Sanchez-Burgos, Adiran Garaizar, Daan Frenkel, and Rosana Collepardo-Guevara. Thermodynamics and kinetics of phase separation of protein-rna mixtures by a minimal model. *Biophysical Journal*, 120(7):1219–1230, 2021.

[268] Giuseppe Carleo and Matthias Troyer. Solving the quantum many-body problem with artificial neural networks. *Science*, 355(6325):602–606, 2017.

[269] Ao Chen and Markus Heyl. Empowering deep neural quantum states through efficient optimization. *Nature Physics*, 20(9):1476–1481, 2024.

[270] Ao Chen. Large-scale simulation of deep neural quantum states. *Universität Augsburg*, 2025.

[271] Moritz Reh, Markus Schmitt, and Martin Gärttner. Time-dependent variational principle for open quantum systems with artificial neural networks. *Physical Review Letters*, 127(23):230501, 2021.

[272] Markus Schmitt and Moritz Reh. jvmc: Versatile and performant variational monte carlo leveraging automated differentiation and gpu acceleration. *SciPost Physics Codebases*, page 002, 2022.

[273] Giovanni Acampora, Andris Ambainis, Natalia Ares, Leonardo Banchi, Pallavi Bhardwaj, Daniele Binosi, G Andrew D Briggs, Tommaso Calarco, Vedran Dunjko, Jens Eisert, et al. Quantum computing and artificial intelligence: status and perspectives. *arXiv preprint arXiv:2505.23860*, 2025.

[274] Markus Schmitt and Markus Heyl. Quantum many-body dynamics in two dimensions with artificial neural networks. *Physical Review Letters*, 125(10):100503, 2020.

[275] Markus Schmitt and Markus Heyl. Simulating dynamics of correlated matter with neural quantum states. *arXiv preprint arXiv:2506.03124*, 2025.

[276] Tianhao Li, W Benjamin Rogers, and William M Jacobs. Interplay between self-assembly and phase separation in a polymer-complex model. *Physical Review E*, 108(6):064501, 2023.

[277] William M Jacobs, David W Oxtoby, and Daan Frenkel. Phase separation in solutions with specific and nonspecific interactions. *The Journal of Chemical Physics*, 140(20), 2014.

[278] Béla Novák and John J. Tyson. Design principles of biochemical oscillators. *Nature Reviews Molecular Cell Biology*, 9(12):981–991, 2008.

[279] Yicheng Qiang, Chengjie Luo, and David Zwicker. Scaling of phase count in multicomponent liquids. *arXiv preprint arXiv:2405.01138*, 2024.

[280] Yicheng Qiang and David Zwicker. flory: A python package for finding coexisting phases in multicomponent mixtures. *Journal of Open Source Software*, 10(107):7388, 2025.

[281] Daniel T Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *The Journal of Chemical Physics*, 115(4):1716–1733, 2001.

[282] Michael A Gibson and Jehoshua Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. *The Journal of Physical Chemistry A*, 104(9):1876–1889, 2000.

[283] Simon Alberti, Shambaditya Saha, Jeffrey B. Woodruff, Titus M. Franzmann, Jie Wang, and Anthony A. Hyman. A user's guide for phase separation assays with purified proteins. *Journal of Molecular Biology*, 430(23):4806–4820, 2018.

[284] Jingzhi Zhu, Long-Qing Chen, Jie Shen, and Veena Tikare. Coarsening kinetics from a variable-mobility Cahn-Hilliard equation: Application of a semi-implicit Fourier spectral method. *Physical Review E*, 60(4):3564–3572, 1999.

[285] Abdel R Minalla and Luc J Bousse. Characterization of microchip separations. In *Microfluidic Devices and Systems III*, volume 4177, pages 123–130. SPIE, 2000.

[286] Nancy C Stellwagen, Soffia Magnusdottir, Cecilia Gelfi, and Pier Giorgio Righetti. Measuring the translational diffusion coefficients of small dna molecules by capillary electrophoresis. *Biopolymers: Original Research on Biomolecules*, 58(4):390–397, 2001.

[287] Daniel C Tessier, Roland Brousseau, and Thierry Vernet. Ligation of single-stranded oligodeoxyribonucleotides by t4 rna ligase. *Analytical biochemistry*, 158(1):171–178, 1986.