Compartmentalization of Synthetic Biochemical Systems

Maximilian Weitz

München 2014

Technische Universität München

Physik Department

Lehrstuhl für Experimentalphysik – E14

Compartmentalization of Synthetic Biochemical Systems

Maximilian Weitz

Vollständiger Abdruck der von der Fakultät für Physik der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. Ulrich Gerland

Prüfer der Dissertation:

1. Univ.-Prof. Dr. Friedrich C. Simmel
2. Univ.-Prof. Dr. Tim Liedl,
   Ludwig-Maximilians-Universität München

Die Dissertation wurde am 17.04.2014 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 17.06.2014 angenommen.

Contents

List of Figures

List of Tables

Zusammenfassung

Mikroemulsionen bieten großartige Möglichkeiten für Studien in femto- bis pikolitergroßen Volumina.
In dieser Arbeit wurden Mikrotröpfchen benutzt um die Auswirkung von Kompartimentierung auf einen *in vitro* Transkriptionsoszillator zu untersuchen. In einer Population von tausenden, oberflächlich identischen Reaktionen fanden wir große Unterschiede in Amplitude, Frequenz und Dämpfungskoeffizient, was auf statistische Variationen in den Konzentrationen der Reaktanten zurückzuführen war.
In einem zweiten Projekt wurde ein künstliches Kommunikationssystem zwischen Gruppen von *Escherichia coli* in einer Mikroemulsion implementiert. Amphiphile Moleküle, die sich von Reservoiren oder mit „Sender"-Bakterien gefüllten Tröpfchen diffusiv ausbreiteten, induzierten in abstandsabhängiger Weise Genexpression in „Empfänger"-Bakterien enthaltenden Tröpfchen. Mathematische Modellierung deutete auf einen stark reduzierten effektiven Diffusionskoeffizienten in der Emulsion hin. Gentechnisch veränderte Bakterien, welche zwei verschiedene chemische Signale integrieren, wurden des Weiteren benutzt um das Potential dieses Systems für genetisch programmierte Musterbildungsstudien zu demonstrieren.

Abstract

Microemulsions offer great opportunities for studies in femto-picolitre sized volumes. In this thesis, microdroplets were utilised to study effects of compartmentalization on an *in vitro* transcriptional oscillator. We observed large variations in amplitude, frequency and damping coefficient in populations of thousands of superficially identical reactions, which can be attributed to statistical variations in reactant concentrations created during partitioning into droplets.
In a second project, an artificial communication system between groups of *Escherichia coli* encapsulated into microemulsion droplets was implemented. Diffusive spreading of amphiphilic molecules from inducer reservoirs or droplets containing "sender" bacteria induced gene expression in droplets containing "receiver" cells in a distant dependent manner. Computational modelling suggested a strongly reduced effective diffusion coefficient in the emulsion. Engineered bacteria that integrated two different chemical signals were further used to demonstrate the system's potential for studies on genetically programmed pattern formation.

1 Introduction

1.1 Synthetic Biology

Synthetic biology aims at realising systems composed of biological molecules with properties that do not exist in naturally occurring systems. One of its very basic motivations is well explained by Richard Feynman's popular quote "What I cannot create, I do not understand". The term *creation* is interpreted very widely. While one part of the synthetic biology community focuses on implementing new functions in existing living systems (hence *in vivo*) like for instance genetic oscillators in *Escherichia coli* [1, 2], another subgroup tries to understand naturally systems' complexity by building biochemical circuits *in vitro* from scratch. Fully synthetic reaction circuits have been realised outside cells which exhibit bistability [3] or even oscillations [4, 5]. The development of an artificial cell or protocell is a further aim for synthetic biologists to study design principles of living organisms [6].
The following paragraph will give a brief overview about the historical development of synthetic biology. In section 1.1.1 important synthetic biology studies and projects related to the *in vivo* work presented in this thesis in chapter 5 are introduced. In the same sense, in section 1.1.2 publications are summarised which serve as foundation as well as motivation for results of this thesis discussed in chapter 4. Since the field has been expanding rapidly for more than a decade, a complete summary would be beyond the scope of this thesis. However, its purpose is to impart a basic idea of synthetic biologists' intention and their success in approaching their vision.
Synthetic biology semi-officially originated around the turn of the millennium but its roots date back much earlier. The first visions of building artificial biological systems by mixing biomolecules already existed at the beginning of the 20th century even though the understanding of biological systems on a microscopic scale was only rudimental. The French biologist Stéphane Leduc was one of the first scientists who argued that problems in biological processes had to be treated from a physical or chemical point of view in order to describe phenomena like development and growth. Almost a century before scientists themselves called their work "synthetic biology", the term occurred repeatedly in his works [7, 8]. However, he was not the only and first one having the idea that living matter underlies same physical laws as - or even arose from - inanimate matter [9, 10]. Even though the ideas already existed the proof could not be provided because the tools and instruments to study and manipulate microorganisms had not been developed to sufficient extent at this time. A break through was made by Elizabeth Szybalska and Waclaw Szybalski with their work on DNA-mediated genetic transformation which built the foundation for, and in which they even coined the term "gene therapy" [11]. From todays perspective, Waclaw Szybalski not only used the term synthetic biology in the context as it is known today, he made an almost visionary prediction in a book chapter published in 1974 when he

stated “up to now we are working on the descriptive phase of molecular biology. [...] But the real challenge will start when we enter the synthetic biology phase of research in our field. We will then devise new control elements and add these new modules to the existing genomes or build up wholly new genomes. This would be a field with the unlimited expansion potential and hardly any limitations to building "new better control circuits" and [...] finally other "synthetic" organisms [...]" [12]. A few years later in 1978 Arber, Nathans and Smith were awarded with the Nobel Prize in Physiology or Medicine for the discovery of "restriction enzymes and their application to problems of molecular genetics" which opened the gates towards completely new possibilities in microbiology. Even though genomes still had to be decrypted at this time, it was from now on in principle possible to deliberately modify, add or remove genes by cutting or pasting of the particular DNA sections in the genome.
The official introduction of the term "synthetic biology" happened at the annual meeting of the American Chemical Society in San Francisco in 2000 by Eric Kool, an organic chemist, chemical biologist, and biophysicist at Stanford University [13, 14]. Since then, the field and number of publications has been growing rapidly. Especially, the decryption of *E. coli*'s genome [15] entailed a boost of numbers of publications in the field. Factors from which the steadily expanding research area benefits are the constantly decreasing prices for sequencing and DNA synthesis. Furthermore, the great future potential for engineering such as biofuel generation [16, 17, 18], drug synthesis [19, 20, 21], and the use of tailored microorganisms as agents for the detection of heavy metal pollution in non-accessible locations [22, 23] favours the progress.
Despite the increasing diversity in synthetic biology subfields a shared idea is the concept to use characterised, biological standard parts to assemble or modify larger systems in an engineering sense. An annual student competition which has grown rapidly since 2003 is also based on this idea. Just 10 years later, in 2013, student teams from 203 universities from all over the world participated in this contest, called the International Genetically Engineered Machine (iGEM) competition. Participants are mostly undergraduates who work on various synthetic biology projects ranging from fundamental research to biotechnological applications. Registered teams are provided with a starter set containing genetic parts with different functions - called DNA Distribution. The genetic parts are - in analogy to electronic parts such as resistors, resonators, or capacitors - listed and described in a registry (http://parts.igem.org). As for other engineering principles, these parts can be used to be connected into larger building blocks like for example a promoter, a ribosomal binding site, a gene for a green fluorescent protein (GFP), and a terminator are merged to form a gene for GFP whose production can be turned on and off. At the finals the iGEM teams' projects are assessed by experienced scientists according to several criteria. Contribution to the parts registry is one of the most important tasks and finally decides about the teams' success. Hence, a central idea of iGEM is to create and characterise genetic parts in order to be used from scientists all over the world. Even though many participants have only beginner skills in scientific research, part registry entries are constantly updated with new experiences and results. Of course, the construct 'parts registry' is filled with information by scientific beginners but the huge number of contributors and users generates a self-regulating effect as it is the case for the open source website wikipedia.com which is becoming more and more reliable since it start in 2001 due to a form of peer review by experts for particular topics. This

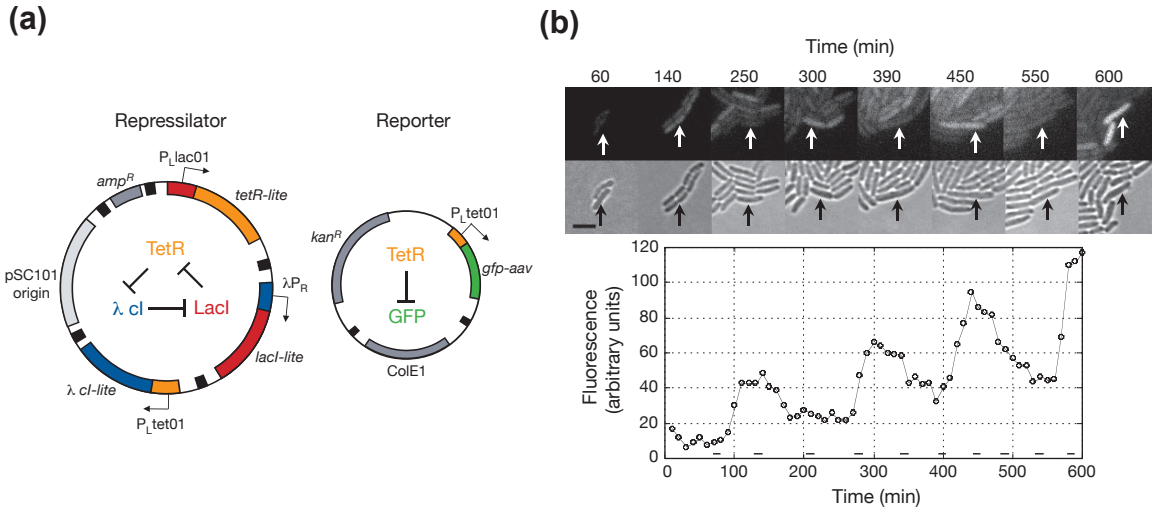

Figure 1.1: The repressilator - A synthetic three node ring oscillator. (a) The repressilator gene circuit. Genes are wired in a cyclic fashion such that each gene's transcription activity is negatively regulated by the protein transcription factor encoded in the upstream gene as indicated schematically in the centre of the plasmid map. The gene oscillations are monitored by expression of GFP whose transcriptional control is in synchrony with one of the three genes. (b) The repressilator gene circuit operated *in vivo*. The microscopy image time series shows repressilator cells in fluorescence (top row) and brightfield (bottom row) mode. The fluorescence time course represents the GFP intensity evolution in the cell indicated in the microscopy images by an arrow. Interestingly, oscillations persist even though cells divide periodically. The scale bar is 4 µm. Reprinted by permission from Macmillan Publishers Ltd: Nature [1] , 2000.

means, even though a few parts might be not or only barely functional, parts which were used several times and positively rated are likely to work.
The great future potential of synthetic biology is also reflected by the financial support funding agencies are willing to provide for this research field. For instance, on July 11th 2013, the British government released an article on its website with headliner "Synthetic biology, one of the most promising areas of modern science, is to receive boost of 60 million pounds". Many more grants and foundations exist who support research on a national and an international scale, for instance the "Human Science Frontier Program" (HFSP), or ERASynBio.

1.1.1 *In vivo* synthetic biology

Synthetic biology aims to build and modify gene circuits which respond to stimuli in the surrounding medium. François Jacob and Jacques Monod contributed significantly to the understanding of gene expression activation and deactivation with their study of the Lac-operon [24] for which they were awarded with the Nobel Prize in Physiology or Medicine in 1965. They also proposed that genetic programs can be generated by careful assembly of genetic elements [25]. However, almost 40 years had passed until first successes were achieved and computational elements could be integrated into living cells [26]. As the term synthetic biology was re-introduced in 2000 also a number of important publications were released around this year. One of the first completely synthetic circuits *in vivo* was realized by Gardner and co-workers who introduced a genetic toggle switch in *E. coli* [27]. The toggle switch was composed of

two different repressible promoters arranged in a circuit where the genes under control of the corresponding promoters mutually inhibited expression of each other. The system could be flipped between two states by chemical or thermal signals. In the same year, Elowitz and Leibler published the first synthetic gene oscillatory circuit in *E. coli*, known as the 'repressilator' [1]. Oscillations were generated in a genetic network consisting of three repressible promoters where each inhibited the expression of another's gene in a ring-like fashion as shown in Fig. 1.1 (a). Interestingly, the oscillation periods were longer than a cell division cycle hence the circuit's dynamics were inherited from generation to generation (Fig. 1.1 (b)). They also found that the dynamical behaviour was 'noisy' due to stochastic fluctuations of components comprising the circuit. Those stochastic fluctuations play a central role when dealing with microorganisms like *E. coli*. Due to its volume in the femtoliter range molecules are present in only small numbers. In this case, biochemical reactions can not be treated deterministically and each reaction event has to be regarded from a probabilistic point of view [28]. Even though Gillespie's algorithm provided a method to simulate the probabilistic nature of those events quite accurately, predictions of the behaviour of larger network structures is computationally very expensive since every single event is taken into account. To further study the origin of noise, bacterial cells were engineered to distinguish between the different sources by deliberately programmed genetic circuits. Michael Elowitz and co-workers found that both intrinsic noise (fluctuations due to small molecule numbers) as well as extrinsic noise (fluctuations of environmental parameters) contributed significantly to the overall dynamics of genetic circuits [29]. It became apparent that the noisy nature of gene expression had to be taken into account when gene expression circuits were supposed to perform on the single cell level as desired. Mechanisms had to be introduced which create robustness against fluctuations in living systems [30, 31]. More aspects of stochastic effects in biological systems are discussed in section. 2.1.6. An *in vitro* approach of a biochemical oscillator in cell-sized microemulsion droplets will be presented in chapter 4.4.2 where stochastic effects will be assessed.

Stochastic effects have less dramatic influences when groups of bacteria are observed since fluctuations are assumed to equal out by the population's mean behaviour [33]. In a seminal work, Weiss and Knight created an artificial communication system between bacteria by isolating the quorum sensing system of *Aliivibrio fischeri*. They dissected the auto-regulatory system into a sender and a receiver part and introduced both into separate bacterial cells hence engineered a intercellular communication mechanism [34]. Bacterial communication by small signalling molecules could now be used to lift synthetic biology from the level of generating intra-cellular genetic programs to a multicellular or even spatial level of genetic programming. These systems can be used to investigate coordinated behaviour in bacterial cell communities. In this sense, Basu and co-workers studied the diffusion mediated communication of sender cells synthesising small signalling molecules and receiver cells containing a genetic pulse-generator circuit [35]. Later they demonstrated in a related study a bacterial pattern formation system with engineered bacteria that responded with fluorescence output only when a diffusively spreading inducer was present in a certain concentration range in a band detector like manner as shown in Fig. 1.2 [32]. Due to well-characterised gene parts, the experimental results were well reproduced in a theoretical reaction-diffusion model. Like single cells in multicellular organisms

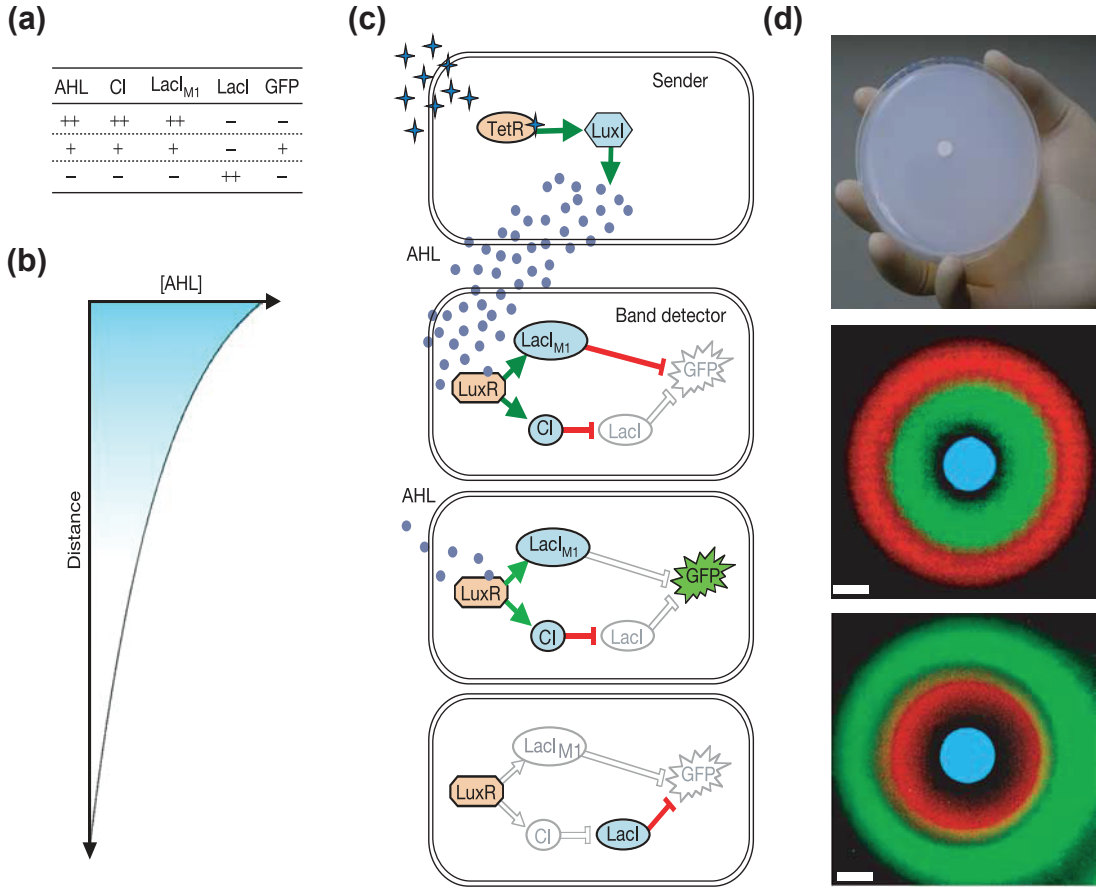

Figure 1.2: Programming bacteria for pattern formation with a genetic band detector for a narrow concentration window of the diffusive signalling molecule AHL. (a) AHL concentration dependent presence of the regulatory proteins CI, LacI$_{M1}$, and LacI, and the reporter protein GFP. ++ represents high, + medium, and - low concentrations. (b) Illustration of the concentration profile for diffusion of AHL from a point source. (c) Scheme of the different pattern formation components and their gene circuit functions. From top to bottom: 1. Sender bacteria express the AHL synthase LuxI which produces the signalling molecule. 2. High levels of AHL cause strong induction of the expression of repressor proteins LacI$_{M1}$ and CI. LacI$_{M1}$ inhibits GFP expression. 3. Medium AHL concentrations activate CI expression strong enough to inhibit expression of LacI which would down-regulate GFP production. At the same time LacI$_{M1}$ is produced only at an intermediate level which is not sufficient to repress GFP expression. Consequently, bacteria start to fluoresce. 4. For low AHL concentration, both LacI$_{M1}$ and CI are only weakly expressed. Therefore, LacI expression cannot be suppressed by CI which causes transcription inhibition of the GFP reporter gene. (d) Top: experimental arrangement of sender bacteria in a petry dish. Centre and bottom: Arrangement of different band detector cells around a sender cell centre (blue) forms macroscopic fluorescent bulls eye patterns. The scale bar is 5 mm. Reprinted by permission from Macmillan Publishers Ltd: Nature [32], 2005.

develop according to their position and function, coordinated behaviour in bacteria combined with sophisticated genetic programs could facilitate three-dimensional tissue engineering, biosensing, and biomaterial fabrication. Also regarding coordinated behaviour in a bacterial cell population, a synthetic gene circuit was engineered which triggered cell death either when the population's cell density crossed a threshold value [36], or when the microbes were subject to external stress which resulted in suicide and consequently "altruistic" release of a public good, thus benefited the survivors [37]. For more sophisticated tasks, additional communication channels are required in order to establish bidirectional communication. A microbial consortium was implemented in *E. coli* cells utilising the LasI/LasR as well as the RhlI/RhlR quorum sensing system from *Pseudomonas aeruginosa* which did only show biofilm formation when two strains were present of which each contained either the one or the other combination of the QS systems [38]. A communication system across species between mammalian cells, bacteria, yeast, and plants was shown by Weber *et al* [39]. A quorum sensing system coupled to a synthetic oscillatory gene expression circuit in *E. coli* was used as signal to synchronise the dynamics of a cell population in microfluidics [40]. In one period of the oscillatory cycle a small, amphiphilic signalling molecule was produced which could freely cross the cell membrane. Due to diffusive spreading of autoinducer molecules in the microfluidic channels, oscillations of the cells were coupled and eventually synchronised despite their spatial separation. Apart from coordination of behaviours, quorum sensing signalling was used to achieve computational tasks with spatially separated bacteria [41]. Here, logic gates were encoded in separate *E. coli* strains. These colonies were arranged apart from each other and wired by diffusive quorum sensing signalling molecules which resulted in the successful construction of 16 two-input logic gates. In chapter 5 the LuxI/LuxR system naturally found in *Aliivibrio fischeri* will be used to study communication and computation between populations of *E. coli* cells compartmentalised into microemulsion droplets [42].

1.1.2 *In vitro* synthetic biology

Already in early synthetic biology studies it appeared that control over single cells is much more challenging than initially expected. A large fraction of scientists therefore concentrates on the accurate description of gene expression at the single-cell level [43, 44, 45]. However, in living systems scientists are faced with an additional interesting problem that hasn't been illuminated thoroughly, namely the influence of the cell membrane on a circuits performance, also known as circuit-chassis interaction [46].
An obvious workaround is to study genetic networks in a cell-free context in batch or in artificial compartments. *In vitro* transcription-translation systems has been used for quite a long time in order to synthesise big amounts of proteins with high yield. A very efficient system - called "protein synthesis using recombinant elements" (PURE) - was realised by reconstitution of about 100 purified components from various organisms [47]. Albert Libchaber and co-workers used a wheat germ extract transcription-translation system to realise synthetic regulatory gene circuits such as cascade and auto-regulatory elements *in vitro* [48]. In high-yield gene expression systems molecular components are conserved and their concentrations increase with time. However,

when genes were connected to circuits it turned out that careful fine-tuning of the degradation machinery is crucial and rapid RNA turnover has to be guaranteed. As a first step towards an artificial cell, a commercially available *E. coli* cell extract was used to carry out gene expression in unilamellar phospholipid vesicles. To prevent nutrients from depletion, the membrane channel protein α-hemolysin was expressed in the vesicle bioreactors. α-hemolysin has a channel diameter large enough to let nutrients pass but small enough that it acts as a barrier for larger molecules like double stranded DNA. When it comes in contact with lipid bilayers, it incorporates and forms a membrane channel. By expressing α-hemolysin inside vesicle bioreactors, the gene expression performance of the *E. coli* cell extract was extended from 5 hours to 4 days (see Fig. 1.3 (a)) [49]. However, the optimisation towards maximum yield and the fact that those expression systems typically consist of components from different organisms makes it hard to compare studies performed in commercially available *in vitro* gene expression kits with those *in vivo*. For this reason, Vincent Noireaux and co-workers developed a gene expression kit which is basically a crude extract from *E. coli* with its endogenous transcription-translation-degradation machinery [50, 51]. Based on this, a coarse-grained model was formulated for simple GFP expression which captured the dynamics well on a qualitative level [52]. Since the crude *E. coli* extract was only based on transcription by the endogenous RNA polymerase rather than bacteriophage RNA polymerase, the build up of more complex circuits was possible. Starting wit holoenzyme E_{70}, a bunch of transcription-translation networks were realised, such as multiple stage cascades, an AND gate and negative feedback loops in bulk and in vesicle bioreactors [53]. In a follow up project, an impressive step further towards an artificial cell was achieved when genome replication as well as expression and self-assembly of bacteriophages was demonstrated *in vitro* [54]. In Fig. 1.3 (b) a schematic of the reaction steps, the results of an experiment with the purpose to proof the functionality of the phages as well as TEM images of the virus capsids are shown. Rather recently, the time to develop a new transcription-translation circuit was considerably reduced due to the application of linear DNA fragments rather than circular. This was enabled by the use of the protein gamS which prevents linear DNA from being degraded by the endogenous degradation machinery [55].

A complementary approach to the implementation of synthetic circuits in living organisms is the generation of biochemical circuits from scratch in a reaction tube or other controlled environments, hence the name "*in vitro*" synthetic biology. Unlike in living systems where still even in the simplest unknown molecular interactions or molecular species exist, *in vitro* circuits only consist of a relatively small number of well-characterised molecules. This approach seems promising in terms of modelling and prediction of the system's behaviour which may facilitate a deeper understanding of the complexity of living systems or even the origin of life.

Even though, several genetic programs in cell-free reactions have been studied in bulk and in lipid vesicles, no dynamically complex circuits like oscillators or multistable systems have been created so far in transcription-translation systems and it still remains challenging to model reactions faithfully. The most accurate description of reaction dynamics in a cell-free approach only focused on the expression of a fluorescent protein. Even though a coarse-grained model was used 10 free parameters were necessary to describe reactions of only 4 enzymes [52]. Apparently, to study the complexity of biological systems further simplification to an even more fundamental

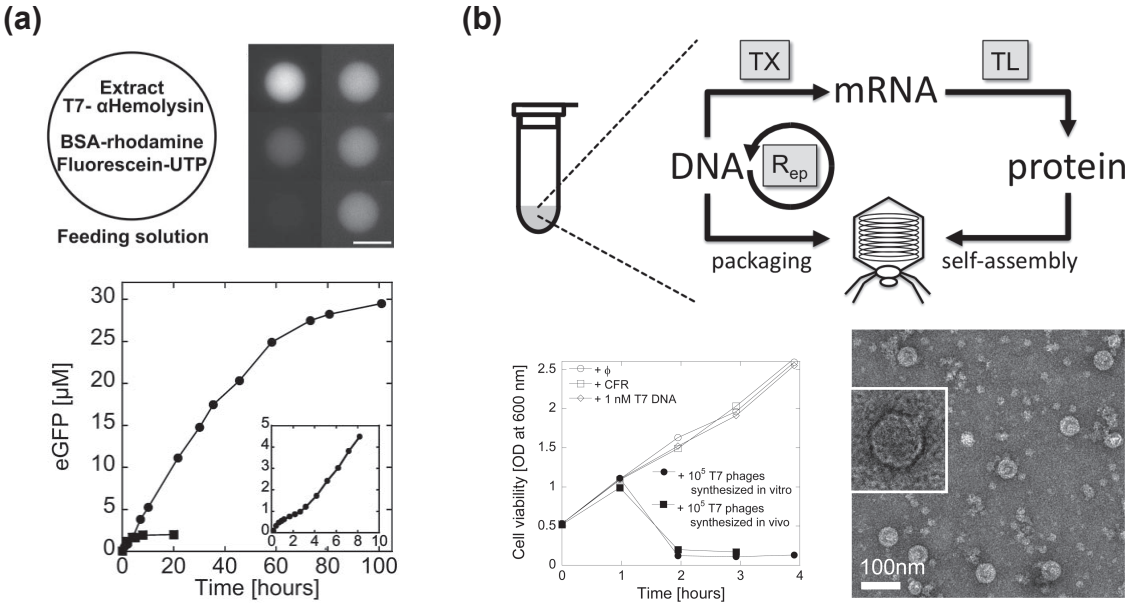

Figure 1.3: Steps towards an artificial cell using a cell-free transcription-translation system. (a) Expression of GFP in vesicle bioreactors. Microscopy images show GFP fluorescence inside vesicles at different times without (left) and with (right) α-hemolysin. The time development of eGFP concentration is shown in the bottom time-course for vesicles without (squares) and with (circles) membrane channels. When α-hemolysin is expressed GFP synthesis lasts for days while without membrane channels gene expression stops after a few hours due to depletion of nutrients as well as accumulation of waste products. Scale bar is 20 µm. Reprinted from [49]. (b) Expression of bacteriophages in an *in vitro* reaction including DNA replication. Top: Illustration of the system. Bottom left: *In vitro* expressed bacteriophages are probed for functionality in *E. coli* viability studies. Addition of pre-produced bacteriophages stops cell growth after 1 h and causes lysis for both the *in vivo* (filled squares) as well as the *in vitro* (filled circles) expressed batch. Bottom right: TEM micrograph of T7 phages synthesised in a cell-free reaction. Adapted with permission from [54]. Copyright 2012 American Chemical Society.

level is required. To this regard, a research area - termed "molecular programming" - started to emerge quite recently whose goal it is to create synthetic reaction systems with designed dynamic behaviour [56]. In one approach, reaction circuits are realised enzyme-free only based on DNA strand displacement reactions [57]. An alternative concept utilises DNA strands and a few enzymes but omit the translation stage completely. Yannick Rondelez and co-workers used a system which they coined the "DNA toolbox" based on isothermal amplification of DNA where information is encrypted and transmitted only in the form of DNA molecules, a mechanism formerly known as nicking amplification [58]. Circuits were driven by three enzymes - a DNA polymerase, a nuclease, and a nickase. DNA signals are produced when a short primer binds to a chemically protected template DNA single strand. The primer is then elongated by the DNA polymerase. The newly synthesised section contains a recognition sequence for a nickase which cuts the elongated strand at the position where the DNA polymerase started to synthesise. The DNA polymerase recognises the short primer again, elongates it and consequently displaces the previously produced single strand, which in the following either can act as primer for a downstream template, or is degraded by the DNase that specifically digests single stranded DNA molecules. Key and simultaneously limiting aspect of this networking strategy is that the maximum length of the template strand is determined by the melting temperature of the hybridised primer-template duplex. Thus, the operation temperature has to be at 46.5 °C which is rather high for biological molecules. However, by proper tuning of experimental parameters an *in vitro* oscillatory reaction circuit was realised comprised by a positive and a delayed negative feedback loop [5]. Slight modification of the main building block enabled the generation of a reaction circuit with which predator-prey dynamics, also found in ecological communities, could be produced. In analogy to population dynamics the system's behaviour was well reproduced using the Lotka-Volterra model [59]. This molecular predator-prey system was later used to study circuit dynamics in emulsion droplets [60] and to follow the development of the population size of "prey" and "predators" on a spatial level [61]. An earlier molecular programming approach using enzymes was introduced by Jongmin Kim and Erik Winfree. Their - more biologically realistic - transcriptional circuit consisted of synthetic, short, linear DNA strands and two enzymes - an RNA polymerase (RNAP) from bacteriophage T7 for RNA production, and an *E. coli* ribonuclease for RNA degradation. Transcription of synthetic DNA genes - coined genelets - can be switched by simply removing or adding a part of the promoter region through strand displacement reactions. The reduced number of components results in a relatively small amount of interactions in contrast to transcription-translation circuits. Thus, computational modelling and a good quantitative understanding of the behaviour of such systems can be achieved. By wiring different genelets with RNA signals - whose concentration level is determined by the transcriptional state of the genelets - repeaters and inverters [62], a circuit which exhibited bistable behaviour [3], and several oscillating transcription networks could be realised [4]. In a follow-up project, a two-node negative feedback oscillator was used as a time keeping device to coordinate the operation of DNA nanodevices and the synthesis of RNA aptamers. By introducing an insulating component, considerable load could be driven while the oscillator's dynamics were only weakly influenced [63]. The genelet concept as well as the two-node oscillator will be discussed in detail in chapter 4.1 and 4.2. In the framework of this thesis, multiple

molecular tasks were coordinated by a transcriptional oscillator presented in chapter 4.3. The same molecular clock reaction circuit was used to study the effects of compartmentalization of a bulk reaction into femto- to picoliter sized emulsion droplets which revealed an impressively high degree of noisiness when systems are downscaled to cell-sizes in chapter 4.4 [64].

1.2 Biological oscillators

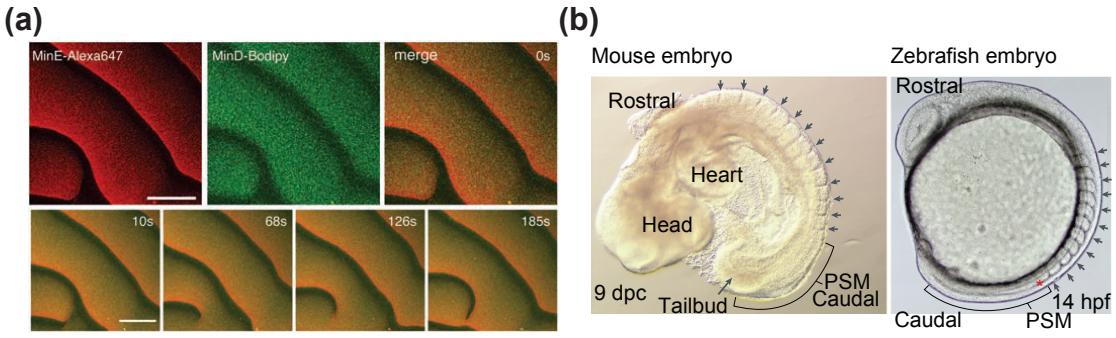

Figure 1.4: Biological oscillators. Many periodic processes in naturally occurring systems are regulated by oscillators. (a) Cell division in *E. coli*. *In vitro* reconstitution of Min proteins exhibit surface waves when placed on a supported lipid membrane. *In vivo*, concentrations of Min proteins oscillate between cell poles and determine the position of cell division. Top row (left to right): fluorescence micrographs of fluorescently labeled MinE, MinD, and merged channels. Bottom row: time-series of merged fluorescence snapshots. The scale bar is 50 µm. Reprinted by permission from "The American Association for the Advancement of Science": Science [65], 2008. (b) Somitogenesis in mouse 9 days post-coitum (left), and in zebrafish 14 hours post-fertilisation (right). The rate of somite formation depends on the growth of the tailbud. Reprinted by permission from Macmillan Publishers Ltd: Nature [66], Science, 2001.

Periodic processes in living organism are typically controlled by oscillators. Prominent processes are circadian rhythms, cell division cycles or morphogenetic processes. Circadian rhythms are responsible for the tuning of organisms' metabolism according to the time of the day, thus have a characteristic period of $\approx 24\,\mathrm{h}$. Circadian rhythms are endogenously regulated and entrainable. They have been found in plants, animals, fungi, and cyanobacteria. In the cyanobacterium *Synechococcus elongatus*, the proteins KaiA, KaiB, and KaiC are responsible for the day-night-rythm. Thereby, the phosphorilation state of KaiC oscillates entrained by a light/dark information through a photosynthetic pathway with a $\approx 24\,\mathrm{h}$ period [67]. Interestingly, the Kai molecules' oscillation is independent of the bacteria's metabolism as a reconstitution in an *in vitro* study showed where oscillations were observed in a sample containing only the three Kai-proteins and ATP [68]. In *Escherichia coli* the mechanism which determines the position at which a cell divides has been associated with oscillations between the cell poles of the proteins MinC, MinD, and MinE [65]. This reaction-diffusion process has been also studied *in vitro* with only two proteins MinD, and MinE, and ATP which generated surface waves on a supported lipid membrane (Fig. 1.4 (a)). In many vertebrates an oscillatory mechanism has been proposed to be responsible for reiterated structure of the somite [66]. The so-called somite-segmentation clock produces

somites in a periodic fashion, for instance with periods of $90 - 120$ min in mouse and 30 min in zebrafish (1.4 (b)). The position where a somite is formed is dependent on the length of the embryos tailbud.
This was just a short excursion into oscillators found in biological systems. However, the main purpose of this paragraph is to give an impression how frequently (chemical) oscillators can be found in nature and which important tasks they control.
In the context with this thesis, studies on a synthetic biochemical oscillator circuit were performed. This oscillator circuit will be used to control molecular downstream processes and to study effects of down-scaling of systems sizes on their dynamics (chapter 4).

1.3 Content

The main topic of this thesis was the study of compartmentalization of synthetic biochemical systems. Encapsulated systems were either *in vitro* reaction circuits or living bacterial cells. Biological, biochemical, and biophysical basics which are used to describe the later presented projects are introduced in chapter 2. Experimental methods which were used to investigate the biochemical systems in bulk and in compartments are presented in chapter 3. Details for the particular topics as well as materials used are added at the end of the corresponding chapters. Results of experimental studies of a biochemical oscillator, and communicating and computing bacteria will be presented and discussed in chapter 4, and 5, respectively.

2 Biological, biochemical and biophysical basics

2.1 Biochemical reactions

2.1.1 Thermodynamics

The free energy F is a thermodynamical system's energy that can be used to perform work. The free energy in dependence of the particle number N, volume V and temperature T for one type of molecule is given by

$$F = F(N, V, T) = U(N, V, S) - TS = -k_B T \ln Z \tag{2.1}$$

with the internal energy U, the system's entropy S, the Boltzmann constant k_B and the partition function Z. The partition function of a canonical ensemble is

$$Z = \sum_i \mathrm{e}^{-\beta E_i} \tag{2.2}$$

where E_i is the total energy E of the system's ith microstate and $\beta = \frac{1}{k_B T}$. With pressure p, an infinitesimal change in free energy is then given by the total differential

$$dF = dU - SdT - TdS = -pdV - SdT + \mu dN \tag{2.3}$$

with the chemical potential μ which represents the change in free energy F for changing particle number N at constant temperature and volume:

$$\mu = \left(\frac{\partial F}{\partial N}\right)_{T,V}. \tag{2.4}$$

The Gibbs free energy (also known as free enthalpy) G is defined as

$$G = G(N, p, T) = F + pV = H - TS \tag{2.5}$$

with the enthalpy of the system H. With equation 2.3 the total differential of 2.5 can be written as

$$dG = -SdT - Vdp + \mu dN. \tag{2.6}$$

The Gibbs free energy is a very useful potential in the context of chemical reactions since those usually take place in test tubes where temperature and pressure are constant. Here, the chemical potential is the change in Gibbs free energy for a change in the molecule number at constant pressure p and temperature T, thus

$$\mu = \left(\frac{\partial G}{\partial N}\right)_{p,T}. \tag{2.7}$$

For systems with more than one type of molecule species, μdN has to be substituted for $\sum \mu_i dN_i$ in equation 2.3 and 2.6 for the chemical potential of the ith particle type.
The chemical potential is the link to chemical reactions which will accompany us throughout this thesis. In the following this relation will be illustrated. Equation 2.2 represents the partition function for one molecule in a heat reservoir like for instance a protein surrounded by water. The partition function for two indistinguishable molecules is

$$Z_2 = \frac{1}{2} \sum_i \sum_j e^{-\beta(E_i+E_j)} = \frac{1}{2} Z_1^2. \tag{2.8}$$

Consequently, for systems with N indistinguishable molecules the partition function can be written as

$$Z_N = \frac{1}{N!} Z_1^N \tag{2.9}$$

with the free energy

$$F = -k_B T \ln Z_N = k_B T\, N! - N k_B T \ln Z_1. \tag{2.10}$$

Using Stirling's approximation $N! \approx N \ln N - N$, equation 2.10 may be approximated as

$$F \approx k_B T (N \ln N - N - N \ln Z_1). \tag{2.11}$$

According to 2.4, the chemical potential then is

$$\mu = \left(\frac{\partial F}{\partial N}\right)_{T,V} = k_B T \left[\ln N - \ln Z_1\right] = k_B T \left[\ln\left(\frac{N}{V}\right) - \ln\left(\frac{Z_1}{V}\right)\right]. \tag{2.12}$$

The second term in 2.12 is usually referred to as standard chemical potential

$$\mu^0 = -k_B T \ln\left(\frac{Z_1}{V}\right).$$

With standard chemical potential and concentration $c = N/V$, equation 2.12 becomes

$$\mu = \mu^0 + k_B T \ln c. \tag{2.13}$$

With respect to a reference concentration c_0 we finally get

$$\mu = \mu^0 + k_B T \ln\left(\frac{c}{c_0}\right). \tag{2.14}$$

Values for μ^0 for many substances can be found in the literature. In honour of the American engineer, chemist and mathematical physicist Josiah Willard Gibbs "Gibbs", abbreviated G (= J/mol) is sometimes used as SI conform unit, but may not be confused with the abbreviation G in the CGS system which stands for the magnetic flux density. The standard conditions are typically:

- Temperature $T = 298\,\mathrm{K}$
- Pressure $p = 101.3\,\mathrm{kPa}$

- Reference concentration $c_0 = 1\,{}^{\text{mole}}/_{\text{L}}$

It has to be noted that the latter derivation only holds for gaseous or dissolved substances in an idealised state without intermolecular interactions. According to Euler's theorem

$$\begin{aligned} f(\lambda x_1, \ldots, \lambda x_N) &= \lambda^n f(x_1, \ldots, x_N) \\ \Rightarrow n\, f(x_1, \ldots, x_N) &= \sum_i x_i \frac{\partial f}{\partial x_i} \end{aligned} \tag{2.15}$$

where λ is a scaling factor, n is the degree of $f(x_1, ..., x_N)$ and i is a running index. Since the particle number N is an extensive variable ($n = 1$) 2.15 also holds true for the Gibbs free energy

$$\begin{aligned} G(\lambda N_1, \ldots, \lambda N_N, \lambda p, \lambda T) &= \lambda^n G(N_1, \ldots, N_N, p, T) \\ \Rightarrow G(N_1, \ldots, N_N, p, T)_{p,T} &= \sum_{i=1}^{N} \left(\frac{\partial G}{\partial N_i} \right)_{p,T} N_i = \sum_i^N \mu_i N_i \end{aligned} \tag{2.16}$$

yielding an expression for changes in G for infinitesimal changes in N_i:

$$(dG)_{p,T} = \sum_i^N \mu_i dN_i \tag{2.17}$$

This enables us now to find a direct relation between the chemical potential and chemical reactions. Considering a reversible reaction of N involved substances X_i, which can be written as

$$-\nu_1[\mathrm{X_1}] - \nu_2[\mathrm{X_2}] \ldots - \nu_j[\mathrm{X_j}] \rightleftharpoons \nu_{j+1}[\mathrm{X_{j+1}}] + \ldots + \nu_N[\mathrm{X_N}] \tag{2.18}$$

with the stoichiometric coefficient ν_i, where

$$\nu_i = \begin{cases} < 0 \text{ for reactants} \\ > 0 \text{ for products} \end{cases} \tag{2.19}$$

and index i indicating the molecular species. We introduce the "extent of a reaction" := ξ to describe the change in G for an infinitesimal change in the particle numbers

$$(dG)_{p,T} = \sum_i \mu_i dN_i = \sum_i \mu_i \nu_i d\xi. \tag{2.20}$$

The infinitesimal change of Gibbs energy for a reaction progress is defined by the expression

$$\Delta_r G = \left(\frac{dG}{d\xi} \right)_{p,T} = \sum_{i=1}^{N} \mu_i \nu_i \tag{2.21}$$

where $-\frac{dG}{d\xi}$ is called the affinity of the reaction. With 2.14, equation 2.21 turns into

$$\begin{aligned} \Delta_r G &= \sum_{i=1}^{N} \left[\mu_i^0 + k_B T \ln \left(\frac{c_i}{c_0} \right) \right] \nu_i \\ &= \Delta_r G^0 + \sum_{i=1}^{N} \nu_i k_B T \ln \left(\frac{c_i}{c_0} \right) \end{aligned} \tag{2.22}$$

which is the change in Gibbs free energy per single reaction turnover. $\Delta_r G^0$ is known as standard free enthalpy change. Exchanging particle concentrations by molar concentrations, 2.22 is

$$\Delta_r G = \Delta_r G^0 + RT \sum_{i=1}^{N} \nu_i \ln\left(\frac{c_i}{c_0}\right) \tag{2.23}$$

with the ideal gas constant R. A reaction is in chemical equilibrium when $\Delta_r G = 0$. In this case, we find a relation between the standard free enthalpy change and the concentration of the ith species for stoichiometric coefficients ν_i with 2.23

$$-\frac{\Delta_r G^0}{RT} = \sum_{i=1}^{N} \nu_i \ln\left(\frac{c_i}{c_0}\right) = \ln \prod_{i=1}^{N} \left(\frac{c_i}{c_0}\right)^{\nu_i}. \tag{2.24}$$

Finally, this leads us to the definition of the equilibrium constant K_{eq}

$$K_{eq} = \prod_{i=1}^{N} \left(\frac{c_i}{c_0}\right)^{\nu_i} = \exp\left(-\frac{\Delta_r G^0}{RT}\right) \tag{2.25}$$

which is known as the "van't Hoff equation". As an example let us consider a reversible reaction in which molecular species A is converted into B and *vice versa*

$$A \rightleftharpoons B \tag{2.26}$$

the Gibbs free energy per single reaction turnover is according to 2.23

$$\Delta_r G = \sum_{i=1}^{N} \nu_i \mu_i = -\mu_A + \mu_B \tag{2.27}$$

where

$$\Delta_r G \begin{cases} < 0 & \\ = 0 & \text{when} \\ > 0 & \end{cases} \begin{array}{l} \mu_B < \mu_A \\ \mu_B = \mu_A \\ \mu_B > \mu_A. \end{array} \tag{2.28}$$

The equilibrium constant of reaction 2.27 is thus

$$K_{eq} = \exp\left(-\frac{\mu_B^0 - \mu_A^0}{RT}\right) = \frac{c_B}{c_A} = \frac{[\mathrm{B}]}{[\mathrm{A}]}. \tag{2.29}$$

2.1.2 Reaction kinetics

In reaction kinetics (bio)chemical reactions are described by rate equations. In this section, first, a general description of a chemical reaction network is given. The following sections will introduce reactions in which up to three reactants are involved called 0th, 1st and 2nd order reactions. Finally, enzymatic reactions, will be treated. Before we go into detail, a few definitions have to be made. In the following, well-stirred systems are regarded in which molecule numbers are assumed to be large enough that a deterministic description of reactions is justified. A method, which

describes reactions in systems with only small numbers of molecules will be introduced in section 2.1.6. A reaction flux v is the change in concentration of involved substances with time. The unit for concentrations is mole per litre $^{\text{mole}}/_{\text{L}}$, abbreviated molar M; the unit for reaction rates is molar per second $^{\text{M}}/_{\text{s}}$. Reaction constants k are proportionality factors which determine how fast reactions occur. Note, while the unit for reaction rates, or flux, respectively, in the deterministic case is always given in concentration per time, the unit of rate constants depends on the order of reactions.

Chemical reaction network

In general, for i species X_i with concentrations $[\mathrm{X}_i]$ the time evolution of the species concentrations is given by a system of i non-linear coupled ordinary first order differential equations, called reaction equations:

$$\begin{aligned}
\frac{d[\mathrm{X}_1]}{dt} &= f_1([\mathrm{X}_1], [\mathrm{X}_2], \ldots, [\mathrm{X}_\mathrm{i}]) \\
\frac{d[\mathrm{X}_2]}{dt} &= f_2([\mathrm{X}_1], [\mathrm{X}_2], \ldots, [\mathrm{X}_\mathrm{i}]) \\
&\ldots \\
\frac{d[\mathrm{X}_\mathrm{i}]}{dt} &= f_i([\mathrm{X}_1], [\mathrm{X}_2], \ldots, [\mathrm{X}_\mathrm{i}]).
\end{aligned} \tag{2.30}$$

The following paragraphs will introduce reactions with up to three reactants. To aid readability, species identified by indices are exchanged by letters A, B and C. [A], [B], and [C] are the time dependent concentrations of species A, B, and C.

Zero-order kinetics

In reactions with zero-order kinetics the reaction velocity v is independent of the concentration of involved molecular species and is proportional to the product of rate constant k and time t thus

$$v = -\frac{d[\mathrm{A}]}{dt} = k. \tag{2.31}$$

For zero-order reactions the rate constant has the unit $^{\text{mole}}/_{\text{L}}$. After integration of equation 2.31, we derive the time dependent concentration of substance A:

$$[\mathrm{A}] = -kt + [\mathrm{A}]_0 \tag{2.32}$$

where $[\mathrm{A}]_0$ denotes the initial concentration of substance A at time $t = 0$. The half-life $t_{1/2}$ is defined as the time after which half of the initial product has reacted. Setting $[\mathrm{A}] = 1/2[\mathrm{A}]_0$ in equation 2.32 yields the half-life of a reaction with zero-order kinetics

$$t_{1/2} = \frac{[\mathrm{A}]_0}{2k} \tag{2.33}$$

Zero-order kinetics are a special case in reaction kinetics and appear in the limits of more general first or second order kinetics. An example for zero-order kinetics will be given in 2.66.

First-order reactions

First-order reactions are reactions of the form:

$$A \underset{k_-}{\overset{k_+}{\rightleftharpoons}} B. \tag{2.34}$$

This is a reversible conversion of a substance A into a substance B and *vice versa*. k_+ and k_- are the rate constants in forward and backward direction respectively. Forward and backward rate constants in first-order reactions have units of $^1/_{\mathrm{s}}$. The change in concentrations [A] and [B] with time t is

$$\frac{d[\mathrm{A}]}{dt} = -k_+[\mathrm{A}] + k_-[\mathrm{B}] \tag{2.35}$$

and

$$\frac{d[\mathrm{B}]}{dt} = k_+[\mathrm{A}] - k_-[\mathrm{B}]. \tag{2.36}$$

With the initial concentrations of species A and B at time $t = 0$, $[\mathrm{A}]_0$ and $[\mathrm{B}]_0$, respectively, solutions for equations 2.35 and 2.36 are

$$\begin{aligned}
[\mathrm{A}] &= \frac{k_+[\mathrm{A}]_0 - k_-[\mathrm{B}]_0}{k_+ + k_-} e^{-(k_+ + k_-)t} + \frac{k_-([\mathrm{A}]_0 + [\mathrm{B}]_0)}{k_+ + k_-} \\
[\mathrm{B}] &= \frac{k_-[\mathrm{B}]_0 - k_+[\mathrm{A}]_0}{k_+ + k_-} e^{-(k_+ + k_-)t} + \frac{k_-([\mathrm{A}]_0 + [\mathrm{B}]_0)}{k_+ + k_-}.
\end{aligned} \tag{2.37}$$

In the special case when k_- is small, thus $k_- \approx 0$, equation 2.37 simplifies to

$$[\mathrm{A}] = [\mathrm{A}]_0 e^{-k_+ t} \tag{2.38}$$

with the half life

$$t_{1/2} = \frac{\ln 2}{k_+}. \tag{2.39}$$

A prominent member of this class of reactions is the concentration independent radioactive decay with decay rate k_+.

Second-order reactions

Second-order reactions are reactions of the form

$$A + B \underset{k_-}{\overset{k_+}{\rightleftharpoons}} C. \tag{2.40}$$

This is the reversible association of molecule species A and B to form product substance C. k_+ and k_- denote the rate constants in forward (association) and backward (dissociation) direction, respectively. Association rate constants have units of $^1/_{\mathrm{M \cdot s}}$, dissociation rate constants have units of $^1/_{\mathrm{s}}$. The resulting reaction equations are:

$$\frac{d[\mathrm{A}]}{dt} = -k_+[\mathrm{A}][\mathrm{B}] + k_-[\mathrm{C}] = \frac{d[\mathrm{B}]}{dt} \tag{2.41}$$

$$\frac{d[\mathrm{C}]}{dt} = k_+[\mathrm{A}][\mathrm{B}] - k_-[\mathrm{C}] \tag{2.42}$$

Reaction 2.40 has reached steady state when

$$\frac{d[\mathrm{C}]}{dt} \stackrel{!}{=} 0. \tag{2.43}$$

Then, a characteristic constant for the stability of a compound can be extracted, the dissociation constant K_d (in units of $^{\mathrm{mole}}/_{\mathrm{L}}$), which is

$$K_d = \frac{k_-}{k_+} = \frac{[\mathrm{A}]_\infty[\mathrm{B}]_\infty}{[\mathrm{C}]_\infty} \tag{2.44}$$

with the steady state concentrations of the involved substances $[\mathrm{A}]_\infty$, $[\mathrm{B}]_\infty$, and $[\mathrm{C}]_\infty$, respectively.
The initial reaction 2.40 appears simple, however, the analytical solution turns out to be a bit more difficult. Due to mass conservation we get boundary conditions for solving 2.42 with the initial concentrations of the involved molecular species $[\mathrm{A}]_0$, $[\mathrm{B}]_0$ and $[\mathrm{C}]_0$:

$$\begin{aligned} [\mathrm{A}]_0 + [\mathrm{C}]_0 &= [\mathrm{A}] + [\mathrm{C}] \\ [\mathrm{B}]_0 + [\mathrm{C}]_0 &= [\mathrm{B}] + [\mathrm{C}]. \end{aligned} \tag{2.45}$$

Together with 2.45 and after applying the parametrisation

$$\begin{aligned} \alpha &:= k_+ \\ \beta &:= -k_+(2[\mathrm{A}]_0 + [\mathrm{B}]_0 + [\mathrm{C}]_0) - k_- \\ \gamma &:= k_+([\mathrm{A}]_0 + [\mathrm{C}]_0)([\mathrm{B}]_0 + [\mathrm{C}]_0) \end{aligned} \tag{2.46}$$

we derive the differential equation for the time evolution of [C]:

$$\frac{d[\mathrm{C}]}{dt} = \alpha[\mathrm{C}]^2 + \beta[\mathrm{C}] + \gamma. \tag{2.47}$$

Further time integration reveals the expression for the time dependent concentration of C:

$$[\mathrm{C}] = \frac{c\exp\left(\sqrt{\beta^2 - 4\alpha\gamma}\right)(\beta + \sqrt{\beta^2 - 4\alpha\gamma}) - \beta + \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha(1 - c\exp\left(t\sqrt{\beta^2 - 4\alpha\gamma}\right))} \tag{2.48}$$

For negligible back reaction rate constants k_- equation 2.41 turns into

$$\frac{d[\mathrm{A}]}{dt} = \frac{d[\mathrm{B}]}{dt} = -k_+[\mathrm{A}][\mathrm{B}]. \tag{2.49}$$

If reactants are initially present in equal stoichiometry, thus $[\mathrm{A}]_0 = [\mathrm{B}]_0$, 2.49 further simplifies equation to

$$\frac{d[\mathrm{A}]}{dt} = \frac{\mathrm{d}[\mathrm{B}]}{dt} = -k_+[\mathrm{A}]^2 \tag{2.50}$$

with the solution

$$[\mathrm{A}] = \frac{[\mathrm{A}]_0}{1 + k_+ t [\mathrm{A}]_0}. \tag{2.51}$$

The half life of an irreversible association reaction with initial reactants in equal stoichiometry turns out to be

$$t_{1/2} = \frac{1}{[\mathrm{A}]_0 k_+}. \tag{2.52}$$

In bimolecular reactions as mentioned previously the binding rate k_+ is determined by the collision of two reactants. An upper limit in well-stirred solutions is set by diffusion to $k_{+,limit} \approx 10^8 - 10^{10}\ {}^1/_{\mathrm{M \cdot s}}$.

Enzyme kinetics

Enzymes are catalysts for biochemical reactions. They speed up the turn over of a substrate (the reactant) into a product by orders of magnitude by lowering the energy barrier between initial and final state. Enzymes are not consumed during this process, which means that their concentration stays constant. In the following paragraph, we will only consider single-substrate reactions.
One of the most basic enzymatic reaction models consists of two main steps: a reversible binding reaction of enzyme E and substrate S into a complex $E : S$ (cf. 2.40), and an irreversible turn over of $E : S$ into free enzyme E and reaction product P (cf. 2.38), hence

$$E + S \underset{k_-}{\overset{k_+}{\rightleftharpoons}} E : S \xrightarrow{k_{cat}} E + P \tag{2.53}$$

where k_+ is the rate constant for association of E and S, k_- is the rate constant for the dissociation of the complex $E : S$ into its constituents, and k_{cat} denotes the so-called turn over number which is a measure for the amount of resulting product P per time, hence its unit ${}^1/_{\mathrm{s}}$. Applying the law of mass action leads to four non-linear reaction equations:

$$\begin{aligned}
\frac{d[\mathrm{E}]}{dt} &= -k_+[\mathrm{E}][\mathrm{S}] + k_-[\mathrm{E:S}] + k_{cat}[\mathrm{E:S}] \\
\frac{d[\mathrm{S}]}{dt} &= -k_+[\mathrm{E}][\mathrm{S}] + k_-[\mathrm{E:S}] \\
\frac{d[\mathrm{E:S}]}{dt} &= k_+[\mathrm{E}][\mathrm{S}] - k_-[\mathrm{E:S}] - k_{cat}[\mathrm{E:S}] \\
\frac{d[\mathrm{P}]}{dt} &= k_{cat}[\mathrm{E:S}].
\end{aligned} \tag{2.54}$$

After some time, the reaction will enter a steady state

$$\frac{d[\mathrm{E:S}]}{dt} \overset{!}{=} 0. \tag{2.55}$$

At steady state (2.55) the concentration of the enzyme-substrate-complex $[\mathrm{E:S}]_{SS}$ is

$$[\mathrm{E:S}]_{SS} = \frac{k_+[\mathrm{E}][\mathrm{S}]}{k_- + k_{cat}} = \frac{[\mathrm{E}][\mathrm{S}]}{K_m}. \tag{2.56}$$

$$K_m = \frac{k_- + k_{cat}}{k_+} \tag{2.57}$$

K_m is widely known as Michaelis-Menten constant named in honor of Michaelis and Menten who proposed an enzyme reaction like shown in 2.53 in 1913 [69]. It is a measure for the affinity between an enzyme and its substrate. The dimension of K_m is concentration, typically $^{\text{mole}}/_{\text{L}}$, the same as for the dissociation constant K_d.
Boundary conditions for 2.54 are the conservation of mass, thus the total enzyme and the total substrate concentrations $[\text{E}]_0$ and $[\text{S}]_0$, respectively:

$$\begin{aligned} [\text{E}]_0 &= [\text{E}] + [\text{E} : \text{S}] \\ [\text{S}]_0 &= [\text{S}] + [\text{E} : \text{S}]. \end{aligned} \tag{2.58}$$

With 2.56 and the first equation of 2.58 we find

$$[\text{E} : \text{S}]_{SS} = \frac{[\text{E}]_0[\text{S}]}{K_m} - \frac{[\text{E} : \text{S}]_{SS}[\text{S}]}{K_m} \tag{2.59}$$

$$[\text{E} : \text{S}]_{SS} = \frac{[\text{E}]_0[\text{S}]}{K_m + [\text{S}]}. \tag{2.60}$$

Here, it is assumed that $[\text{S}] \gg [\text{E} : \text{S}]$. This assumption is key in this derivation. The assumption is reasonable as long as $K_M \gg [\text{E}]_0$. The production rate of P at steady state is

$$\frac{d[\text{P}]}{dt} = k_{cat}[\text{E} : \text{S}]_{SS}. \tag{2.61}$$

Introducing 2.60 into 2.61 leads to

$$\frac{d[\text{P}]}{dt} = \frac{k_{cat}[\text{E}]_0[\text{S}]}{K_m + [\text{S}]} \tag{2.62}$$

and

$$\frac{v}{v_{max}} = \frac{[\text{S}]}{[\text{S}] + K_m} \tag{2.63}$$

where the production rate is the reaction velocity $v([\text{S}])$ and the maximum reaction velocity is $v_{max} = k_{cat}[\text{E}]_0$ achieved at substrate saturation. The meaning of K_m becomes immediately clear when we set $[\text{S}] = K_m$ in 2.63. K_m represents the substrate concentration [S] at which the reaction appears at half of its maximum velocity v_{max}. 2.63 is one of the best known approaches to enzyme kinetics, the Michaelis-Menten equation. However, the derivation used here differs from the original by Michaelis and Menten and was provided by George Briggs and J.B.S. Haldane in 1925 [70]. Michaelis and Menten assumed that binding and dissociation of E and S happens much faster than the conversion step into P which is also called the rapid equilibrium approximation. Briggs and Haldane made no assumption to this end. The Michaelis-Menten equation is especially useful in limiting cases as discussed in the following.
First, when the dissociation rate of the complex becomes much smaller than the turn

over number, thus $k_- \gg k_{cat}$, then the complex' dissociation constant K_d is a good approximation for K_m:

$$K_m = \frac{k_- + k_{cat}}{k_+} \approx \frac{k_-}{k_+} = K_d. \tag{2.64}$$

Second, in the opposite case, when $k_- \ll k_{cat}$, then the dimensions of K_m turn into a rate rather than a concentration and the reaction has to be treated as reported by Van Slyke et al. in [71].
Third, when the substrate's concentration is much smaller than K_m, then the reaction behaves like a first-order reaction and we find

$$\frac{v}{v_{max}} = \frac{[\mathrm{S}]}{[\mathrm{S}] + K_m} \approx \frac{[\mathrm{S}]}{K_m}. \tag{2.65}$$

Finally, when $[\mathrm{S}] \gg K_m$ the enzymes are substrate saturated and the reaction velocity will become independent of [S], thus resulting in zero-order kinetics since

$$\frac{v}{v_{max}} = 1. \tag{2.66}$$

2.1.3 Gene expression

In this section, gene expression in prokaryotes is introduced. Gene expression is the synthesis of a protein which is encoded in a gene. Genes consist of deoxyribonucleic acid molecules (DNA) which serve as storage for the information needed for a functional organism. Genes in prokaryotes can be located on a chromosome (one per cell) or on plasmids which are both circular DNA molecules. Gene expression in eukaryotes is far more complicated than in prokaryotes, thus the question for an universal definition of the word "gene" has been controversial [72]. In the context of this thesis we will use the expression gene for an unit on a DNA molecule which can be transcribed into RNA by an RNA polymerase (abbreviated RNAP). In Fig. 2.1 (a) a gene is illustrated in a rather simplistic view. Furthermore, we will concentrate our discussion on prokaryotes if not mentioned otherwise.
The central dogma of molecular biology describes the information flux in living organisms. It distinguishes between three main stages: replication, transcription and translation. When a cell reproduces, the genome is replicated and its offspring will receive a molecular copy. The information stored in a gene is transcribed into another type of biopolymer, ribonucleic acids (RNA), by special enzymes, RNA polymerases (RNAP). RNA transcribed from genes which serves as messenger to transmit information is called messenger RNA, abbreviated mRNA, and is the most prominent candidate in the family of ribonucleic acids. mRNA molecules consist of multiple units of three nucleotides, the so-called codons. Codons are the units, which are used as template for the synthesis of proteins by ribosomes, a process called translation. They either serve as regulatory sites for initiation or termination of translation or encode for an aminoacid. Chains of aminoacids, assembled according to their mRNA sequence, are proteins. The combination of transcription and translation is called gene expression and is shown schematically in Fig. 2.1 (b). The whole genome contains a huge number of genes. For example, in one of the simplest organisms *Escherichia coli* (*E. coli*) the genome contains more than 4000 genes [15]. First it was - from today's perspective - naively expected that only the knowledge about the genomic sequence

will enable us to control gene expression as desired by simple cut and paste of certain genes into an existing genome. Quite soon it became clear that having control about such systems is far more complex. A detailed understanding of the gene regulatory mechanisms and the connections between superficially independent genes to network structures is crucial to modify existing or integrate completely new functions into living organisms.

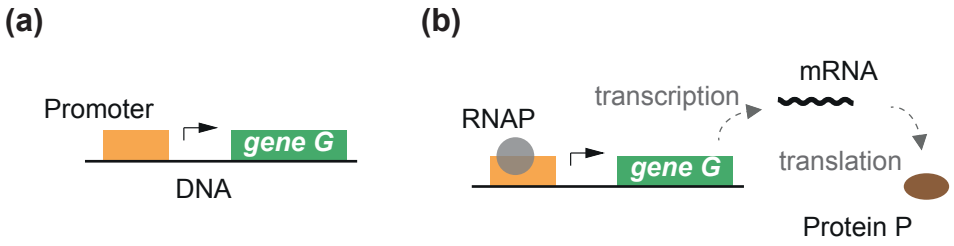

Figure 2.1: Gene expression. (a) Schematic representation of a DNA section which contains a promoter and a gene. A promoter (orange) is a specific DNA sequence, which serves as recognition site for an RNA polymerase (RNAP), an enzyme which catalyzes the synthesis of RNA. The type of promoter regulates the rate in which a gene G is read out. (b) Schematic illustration of gene expression. An RNAP binds the promoter and transcribes the downstream gene G into mRNA which is further translated into a protein P.

2.1.4 Gene regulation

Cells are impressive self organising systems, which can perform tasks such as to replicate on their own, to sense their environment, to adapt to external influences, to respond to and to resist stress, to communicate, and many more. When *E. coli* replicates, the genome is copied and the cell divides into two (almost) identical daughter cells. In the presence of external stimuli, such as changes in the environmental conditions, bacteria can for instance respond with formation of biofilms, secretion of toxins, or conjugation. If nutrients like glucose drop below a critical low concentration in their environment the metabolism can undergo a transition and start to harvest energy from alternative sources like lactose. By evolution, bacteria develop resistances against antibiotics and assure the survival of their population. In 1979, Nealson and Hastings found that cells can measure their population density by a special communication mode coined quorum sensing [73]. If a population density exceeds a critical value, special programs are initiated which are for instance only advantageous in presence of a large number of other individuals. All examples given here are based on changes in gene regulation. In the following, it is discussed how and on which level genes are regulated. Changes in gene expression afford genetic switching mechanisms. Gene regulation can happen on very different levels on the path from coding gene to task fulfilling protein. Evolution tuned regulation according to the time scales on which changes in the environment occur where an organism has to respond and whether this response has to be permanent or transient. The occurring fluctuations to which bacteria are faced, are usually buffered quite efficiently by gene regulation circuits that have been adjusted extremely accurately during thousands of years of evolution. An example for gene regulation on much shorter time scales is the response of metabolism to nutrient shortage. The Lac-operon, for instance, accounts

for a gene regulation mechanism which switches between metabolic pathways according to the availability of either glucose or lactose in the surrounding medium of *E. coli*. Francois Jacob and Jacques Monod were honoured by the Nobel Price in the category Physiology or Medicine in 1965 for the Operon model of gene expression on the example of the Lac-operon [24].
In this work, gene regulation mechanisms are used to build up artificial networks or modify existing which work on the transcriptional level only. Fig. 2.2 shows examples for transcriptional gene regulation mediated by small input molecules: transcription induction upon either activation of a transcription factor by an inducer molecule (Fig. 2.2 (a)), or by deactivation of a repressor due to binding of an input molecule (Fig. 2.2 (d)), and transcription inhibition upon either deactivation of a transcription activator by an input molecule (Fig. 2.2 (b)), or activation of a repressor protein by binding of a ligand (Fig. 2.2 (c)). It has to be stressed that in real gene regulation systems the switching is usually more complex. Activators like LuxR and repressors like LacI whose functionality is modulated by ligands usually form multimers (either in their active or in their native state). In the following, mathematical concepts are introduced which are used to model and predict dynamics within gene circuitry.

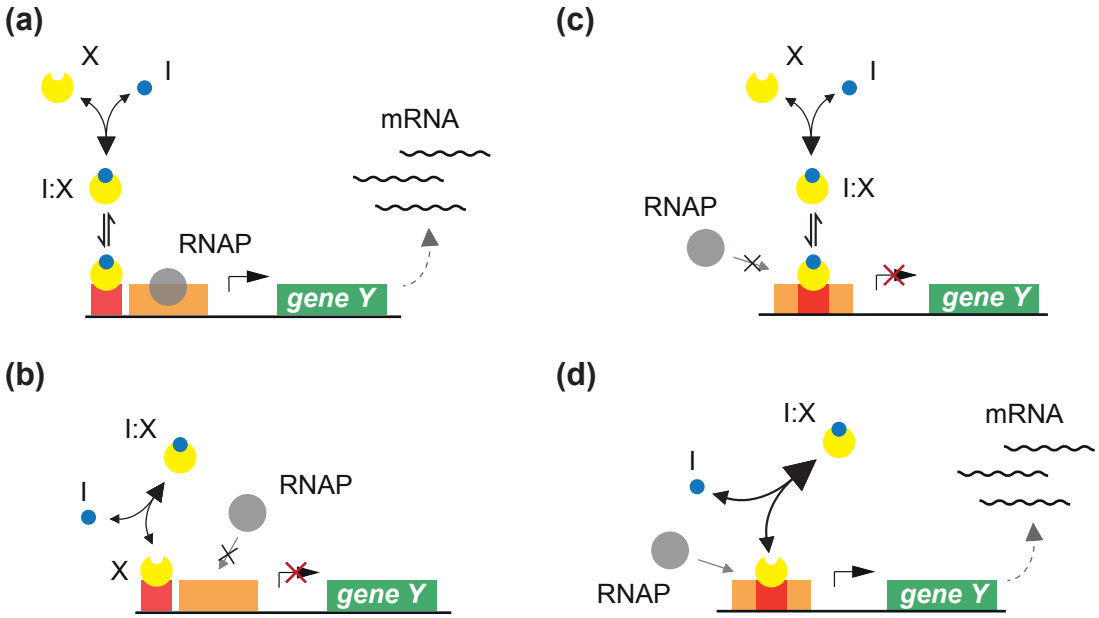

Figure 2.2: Gene regulation by input molecules. (a), (b) Schematic representation of transcription control by an activator X (yellow). (a) An input molecule I (blue) acts as inducer. I and X bind to form the complex $I : X$ which acts as transcription activator. When $I : X$ is bound to its recognition site in front of the promoter the affinity for an RNAP to bind is strongly increased and turns transcription of $geneY$ into $mRNA$ ON. (b) An input molecule acts as inhibitor. When I binds X the complex $I : X$ is formed which cannot act as transcription activator anymore. As a consequence the affinity for RNAP to bind the promoter decreases strongly and transcription of $geneY$ is turned OFF. (c), (d) Schematic representation of transcription control by an repressor X (yellow). (c) An input molecule I binds to repressor X and builds the active complex $I : X$ which can sequester a promoter. Consequently, the promoter is blocked for RNAP and transcription of $geneY$ is inhibited. (d) An input molecule I removes repressor X from its binding site at a promoter due to formation of the inactive complex $I : X$. The promoter in turn is accessible for an RNAP and $geneY$ can be transcribed into RNA, thus transcription is turned ON.

To illustrate how gene regulation can be described mathematically, the gene regulation function for a repressor which releases a promoter upon binding to an input molecule as shown in Fig. 2.2 (d) will be derived. We will start with the second-order binding reaction of a repressor X to a promoter P to form the complex $X : P$:

$$X + P \underset{k_-}{\overset{k_+}{\rightleftharpoons}} X : P \tag{2.67}$$

where k_+ describes the rate of collision of the two molecules in units of $1/\mathrm{M{\cdot}s}$. The backward reaction is described by the dissociation rate k_- in units of $1/\mathrm{s}$. The promoter can either be free P or occupied by the repressor in the complex $X : P$ thus the total amount of promoter $[\mathrm{P}]_0$ follows the conservation law

$$[\mathrm{P}]_0 = [\mathrm{P}] + [\mathrm{X} : \mathrm{P}]. \tag{2.68}$$

The complex formation rate is:

$$\frac{d[\mathrm{X} : \mathrm{P}]}{dt} = k_+[\mathrm{X}][\mathrm{P}] - k_-[\mathrm{X} : \mathrm{P}] \tag{2.69}$$

The steady-state concentration of the complex $X : P$ when $\frac{d[\mathrm{X:P}]}{dt} = 0$ is

$$[\mathrm{X} : \mathrm{P}]_{SS} = \frac{k_+}{k_-}[\mathrm{X}][\mathrm{P}]. \tag{2.70}$$

With equation 2.68 and the dissociation constant $K_d = k_-/k_+$ with units of concentrations we derive the fraction of free promoter

$$\frac{[\mathrm{P}]}{[\mathrm{P}]_0} = \frac{1}{\frac{[\mathrm{X}]}{K_d} + 1}. \tag{2.71}$$

Half of the total promoter is free, when $[\mathrm{X}] = K_d$. The fraction of free promoter can also be seen as the probability that the promoter is available for transcription by an RNAP. The maximal rate with which a gene can be transcribed into RNA is determined by the strength of a promoter. The promoter strength depends on its sequence. A measure for promoter strength is the maximal transcription rate α which yields the RNA transcription rate when multiplied with the probability of free promoter given in equation 2.71:

$$\frac{d[\mathrm{RNA}]}{dt} = \alpha \frac{[\mathrm{P}]}{[\mathrm{P}]_0} = \frac{\alpha}{\frac{\mathrm{X}}{K_d} + 1} \tag{2.72}$$

in units of $1/\mathrm{s}$.

As mentioned in the beginning and shown in Fig. 2.2, transcription factors can be often turned active or inactive by small ligand molecules. Upon ligand binding, activators or repressors change their conformation and their ability to activate or to repress transcription. In the following, the case is treated were a ligand binds to a repressor and causes its dissociation from the promoter. This makes the promoter accessible for transcription by an RNAP. The ligand is therefore also called inducer. Genetic switches are often compared to electronic switches and their operation is

sometimes described using boolean logic. From this context, ligands or inducers are often also called input.

In analogy to equation 2.67, inducer I and repressor X associate to form the complex $I:X$:

$$I + X \underset{k_-}{\overset{k_+}{\rightleftharpoons}} I:X \tag{2.73}$$

with forward rate k_+ for collision of inducer I and repressor X and dissociation rate k_- for the breaking of complex $I:X$. The total amount of repressor $[\mathrm{X}]_0$ is

$$[\mathrm{X}]_0 = [\mathrm{X}] + [\mathrm{I:X}]. \tag{2.74}$$

We derive the complex formation with rate equation 2.73

$$\frac{d[\mathrm{I:X}]}{dt} = k_+[\mathrm{I}][\mathrm{X}] - k_-[\mathrm{I:X}] \tag{2.75}$$

Under steady state conditions where $\frac{d[\mathrm{I:X}]}{dt} = 0$ and conversion of equation 2.74 we find an expression for the concentration of complex $I:X$:

$$[\mathrm{I:X}] = \frac{[\mathrm{I}][\mathrm{X}]_0}{K_{d,I} + [\mathrm{I}]} \tag{2.76}$$

with the dissociation constant $K_{d,I}$ of the complex $I:X$. Equation 2.76 shows three interesting features of the complex formation process. First, for high concentration of inducer I it saturates. Second, for $[\mathrm{I}] << K_{d,I}$ it is linear in [I]. And third, the dissociation constant $K_{d,I}$ is the concentration of inducer I for which half of maximum [I : X] is formed.

In this case, the promoter is available for transcription when the repressor is sequestered by a ligand. The concentration of free repressor [X] in equation 2.72 can now be expressed in dependence of free inducer concentration using the mass conservation in 2.74 as

$$[\mathrm{X}] = \frac{[\mathrm{X}]_0}{1 + \frac{I}{K_{d,I}}}. \tag{2.77}$$

With equation 2.72 and equation 2.77 we find an expression for the RNA transcription rate in dependence of free inducer I

$$\frac{d[\mathrm{RNA}]}{dt} = \frac{\alpha}{1 + \left(\frac{1}{1+\frac{[\mathrm{I}]}{K_{d,I}}}\right).\frac{[\mathrm{X}]_0}{K_d}} \tag{2.78}$$

However, in gene regulation, it is often necessary that the transition between transcriptional active and inactive state is sharp when a certain threshold concentration is exceeded rather than continuous. This is realised in nature by the fact that transcription factors are often composed of multiple monomer subunits from which each can bind inducers and often only reach full functionality when all subunits are occupied by an inducer. If we assume a repressor protein X that can bind n ligands I then the total repressor concentration $[\mathrm{X}]_0$ is given by

$$[\mathrm{X}]_0 = [\mathrm{X}] + [\mathrm{nI:X}] \tag{2.79}$$

where $nI : X$ is the complex consisting of n ligands I bound to X. The complex formation rate is

$$\frac{d[\mathrm{nI : X}]}{dt} = k_+[\mathrm{I}]^n[\mathrm{X}] - k_-[\mathrm{nI : X}]. \tag{2.80}$$

When 2.80 reaches steady state, the concentration of complex $nI : X$ is

$$[\mathrm{nI : X}]_{\mathrm{SS}} = \frac{k_+}{k_-}[\mathrm{I}]^n[\mathrm{X}] = \frac{1}{K_{d,I}}[\mathrm{I}]^n[\mathrm{X}]. \tag{2.81}$$

The conservation of the total amount of repressor X in steady-state, thus 2.79 in 2.81 gives the probability for the repressor to be bound to n ligand molecules I or more general the fraction of occupied sites by a ligand:

$$\frac{[\mathrm{nI : X}]}{[\mathrm{X}]_0} = \frac{[\mathrm{I}]^n}{K_{d,I} + [\mathrm{I}]^n} \tag{2.82}$$

With the dissociation constant of the multimeric repressor-ligand complex defined as $K_H^n = k_-/k_+ = K_{d,I}$, equation 2.82 can be written as

$$\frac{[\mathrm{nI : X}]}{[\mathrm{X}]_0} = \frac{[\mathrm{I}]^n}{K_H^n + [\mathrm{I}]^n}. \tag{2.83}$$

Equation 2.83 (or 2.82) is known as Hill equation which is frequently used to express and quantify cooperativity in transitions between states. In the context with gene regulation K_H is called the threshold value, n is called the Hill exponent. Even though the Hill-equation is widely used to characterise transfer functions or gene regulation functions, however, it has to be noted that in most cases the Hill-exponent can not be interpreted easily. For instance, the derivation assumes a synchronous binding of ligands which is implausible because it affords the simultaneous availability of all ligands. Models for a more accurate description of the ligand binding exist that consider sequential binding or also states in which a certain number of bound ligands is more or less favoured.
For completion, we convert 2.78 into the equation which describes the transcription rate for repressors which are deactivated by the binding of multiple ligands:

$$\frac{d[\mathrm{RNA}]}{dt} = \frac{\alpha}{1 + \left(\frac{1}{1+\left(\frac{[\mathrm{I}]}{[\mathrm{K}]_H}\right)^n}\right)\frac{[\mathrm{X}]_0}{K_d}}. \tag{2.84}$$

When gene induction or repression is studied *in vivo*, the repressor/ inducer molecules are often expressed constitutively hence the concentration is unknown. Furthermore, equation 2.84 describes only the dependence of the RNA transcription rate on the inducer molecule's concentration. When cells are observed, the observable which is used for quantification is often a fluorescence output of a protein situated behind the promoter of interest. Instead of expanding equation 2.84, such processes are typically simplified by a gene regulation or transfer function that describes the rate of fluorescent protein expression as a function of inducer concentration, omitting the detailed mRNA transcription and protein translation step. Using this simplified approach it is usually only distinguished between the influence of the small inducer molecule (ligand)

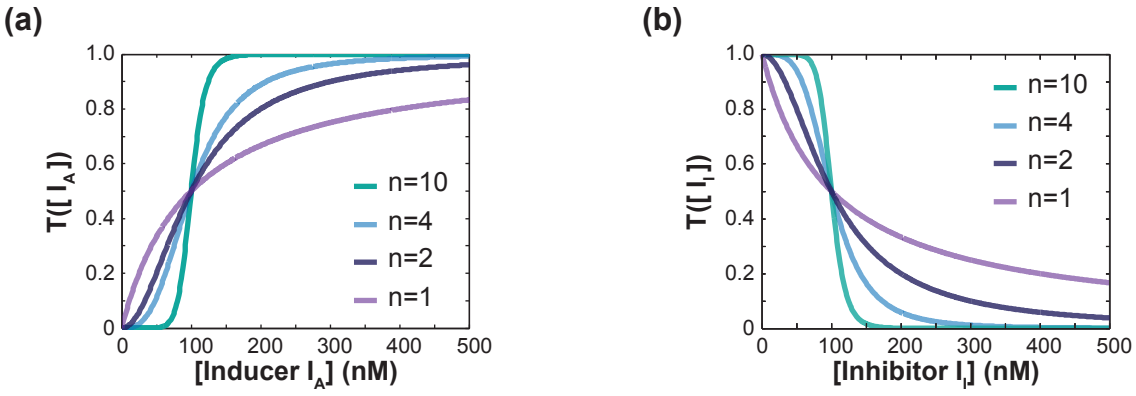

Figure 2.3: Transfer function T for expression induction and inhibition by an effector molecule. (a) Plots of transfer functions for induction by an inducer I_A for Hill exponents $n = 1$, $n = 2$, $n = 4$, and $n = 10$ (equation 2.85). (b) Plots of transfer functions for repression by an inhibitor I_I for Hill exponents $n = 1$, $n = 2$, $n = 4$, and $n = 10$ (equation 2.86). In these cases $\alpha = 1$ and $\alpha_0 = 0$. The threshold value is for both (a) and (b) $K = 100\,\mathrm{nM}$. For $n = 1$, the transfer functions have the same shape as a plot of the Michaelis-Menten equation. The higher n, the more step-like appears the switching behaviour.

that is added rather than whether the transcription regulating protein is an activator or a repressor. A transfer function T in dependence of an expression activating input molecule $[\mathrm{I_A}]$ can be quantified with a transfer function of the form

$$T([\mathrm{I_A}]) = \frac{\alpha}{1 + \left(\frac{K}{[\mathrm{I_A}]}\right)^n} + \alpha_0 \tag{2.85}$$

where α is the maximum protein expression rate, α_0 is the basal or leakage transcription rate in absence of inducer, K is the threshold value describing the concentration of $[\mathrm{I_A}]$ for which have of the maximum expression rate is reached, and n is the Hill exponent which is a measure for cooperativity. Fig. 2.3 (a) shows example transfer functions for different n. It has to be noted that in this case n may not strictly be interpreted according to its introduction in 2.79 but rather as a value which enables comparison of different transcription switching processes without knowing the mechanism in detail. For a system where only one ligand binds to the repressor, $n = 1$, equation 2.83 turns into equation 2.76. As illustrated in Fig. 2.2, inducer molecules can also act inhibiting on gene expression processes. In this case the transfer function T turns into

$$T([\mathrm{I_I}]) = \frac{\alpha}{1 + \left(\frac{[\mathrm{I_I}]}{K}\right)^n} + \alpha_0 \tag{2.86}$$

where $[\mathrm{I_I}]$ is the concentration of the inhibiting effector molecule. Example transfer functions for different n are shown in Fig. 2.3 (b).

For synthetic biologists, switchable genes as described above can be used in principle as parts which further can be connected to generate more sophisticated functions than just activation or inhibition of gene expression. For instance, a genetic AND-gate can be realised by connecting switchable subunits such that a genetic circuits only responds with an output when both inducer molecules are present. As initially introduced, without the normalisation to protein production rates through α and α_0, the Hill function describes the probability for a gene to be activated. Assuming that

in the case of our AND-gate the transfer functions are independent for both inputs, the probability for the output gene to be activated is given by the product of the probability of two transfer functions for two independent inducers (equation 2.85) [74, 75]. With normalisation to the basal expression level α_0 and the output gene expression level at full induction α we derive

$$T([\mathrm{I_{A,1}}],[\mathrm{I_{A,2}}]) = \frac{\alpha}{\left(1+\left(\frac{K_1}{[\mathrm{I_{A,1}}]}\right)^{n_1}\right)\left(1+\left(\frac{K_2}{[\mathrm{I_{A,2}}]}\right)^{n_2}\right)} + \alpha_0 \tag{2.87}$$

where K_1 and K_2 are the threshold values and n_1 and n_2 are the Hill exponents for inducers $I_{A,1}$ and $I_{A,2}$, respectively. In this thesis, transfer functions as introduced were used to describe the gene induction of a Lac-promoter by IPTG and of a Lux-promoter by AHL. Both systems were further combined to create a genetic AND-gate. The genetic circuits and their characterisation in *E. coli* will be presented in chapter 5.2.

From genes to circuits

As we have seen in section 2.1.4 transcription factors can act as transcription activators or repressors. Tasks performed by cells are usually based on the interplay of multiple connected genes which form a network-like architecture. Dependent on how genes are connected with each other network motifs arise, for instance such as positive or negative feedback circuits. Those gene networks are often illustrated as graphs with nodes and edges (Fig. 2.4). Nodes typically represent genes that encode proteins. Proteins with regulatory functions (transcription factors) serve as signals which connect two genes. These signals are drawn as connecting lines (called edges) between two nodes where the shape of the end towards the influenced gene indicates whether the signal regulates the downstream gene in an inhibitory or activatory manner. The first is symbolised by a blunt end, the latter is drawn as arrowhead. By comparing those networks with network structures generated randomly [76] certain motifs were identified that occurred significantly more frequently than others - the so-called network motifs [77]. The frequent occurrence of particular network motifs indicates that those were selected during evolution since they have provided special advantages for the equipped organisms. As we have seen in the previous section, gene regulation mechanisms are often non-linear. When regulatory elements based on such switching principles are connected, dynamical systems arise and phenomena like bistability or oscillations occur which will be the central topic in the next section.

2.1.5 Dynamical systems

Dynamical systems

In dynamical systems the time evolution of variables is described either by differential equations if time is continuous or by difference equations if time is discrete. We will here focus on systems where time is continuous and systems are described by differential equations. In this section, basic principles of dynamical systems are introduced and parallels to examples in biology will be drawn. Non-linearities like for instance due to cooperative switching behaviour of gene expression make dynamical systems

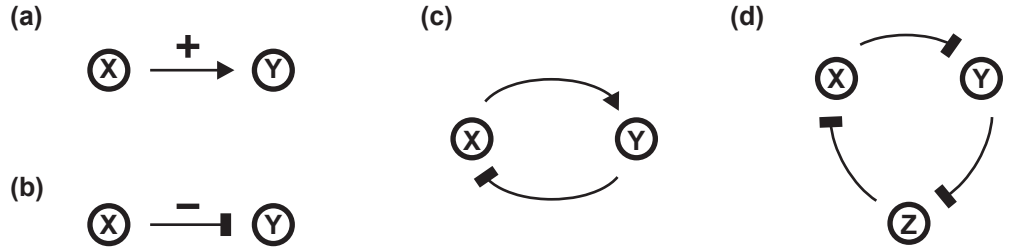

Figure 2.4: Gene regulation networks in a representation adapted from graph theory. Nodes represent genes, edges represent interactions due to signals. (a) Positive regulation: gene X activates gene Y. Activation is indicated by an arrow head. (b) Negative regulation: gene X represses gene Y. Inhibition is indicated by a blunt end. (c) A negative feedback circuit. Gene X activates Y while Y inhibits X. Under certain conditions, this circuit can exhibit oscillatory behaviour. (d) A three-node ring-circuit that can show oscillations. X, Y, and Z inhibit each other in a circular manner.

interesting. Regulatory elements as introduced in section 2.1.4 can be connected to larger network structures. In Fig. 2.4 circuits are shown in a representation adapted from graph theory. In the context with gene networks, nodes represent genes whose activity can be modulated and edges represent molecular signals connecting the genes which can be either activatory (indicated by an arrow head) or inhibitory (indicated by a blunt end). An activating and an inhibiting connection between a gene X and a gene Y are shown in Fig. 2.4 (a) and (b), respectively. For example, let us assume Y's activity is characterised by the transfer function for an activating molecule x given in 2.85, the activator is represented by the arrow headed edge (Fig. 2.4 (a)). If x is produced by gene X, the time dependent concentration of output molecule y is

$$\frac{d[\mathrm{y}]}{dt} = \frac{\alpha_y}{1 + \left(\frac{K_Y}{[\mathrm{x}](t)}\right)^{n_Y}}, \tag{2.88}$$

where α_Y is the production rate of y at full induction of Y, K_Y and n_Y are the threshold concentration and the Hill exponent describing the gene activation of Y by activator molecule x. $[\mathrm{x}](t)$ is the time dependent concentration of x, hence $\dot{[\mathrm{y}]} = f([\mathrm{x}])$. For the following considerations, where a method is introduced to analyse a dynamical system's behaviour, square brackets are omitted since relations are not limited to chemical reactions. Assuming the one-dimensional case where $x(t)$ is a real valued time dependent function and $f(x)$ is a smooth real valued function of x, then

$$\dot{x} = f(x). \tag{2.89}$$

This is of course not necessarily a non-linear system. Non-linearities arise when the function on the right hand side contains terms with time dependences to powers different than one. A dynamical system is characterised by its fixed points x_i. There are two categories of fixed points,

$$x_i \text{ is } \begin{cases} \text{stable fixed point}, & \text{when } f(x_0) = 0 \wedge f'(x_0) < 0 \\ \text{unstable fixed point}, & \text{when } f(x_0) = 0 \wedge f'(x_0) > 0. \end{cases} \tag{2.90}$$

In two-dimensions, a dynamical system is described by

$$\begin{aligned} \dot{x} &= f(x, y) \\ \dot{y} &= g(x, y). \end{aligned} \tag{2.91}$$

The curves for which

$$\begin{aligned} \dot{x} &= 0 = f(x,y) \\ \dot{y} &= 0 = g(x,y) \end{aligned} \tag{2.92}$$

are called "nullclines". The fixed points of this system are given by the intersections of the nullclines, thus

$$\dot{x} = 0 \wedge \dot{y} = 0. \tag{2.93}$$

Be (x_0, y_0) fixed point of 2.91. Fixed points are points in a system's phase space where its state is stationary. The type of fixed point can be determined by studying the system's response to small perturbations. For small deflections from stable fixed points, the system may transiently oscillate around it but eventually return to (x_0, y_0) like a marble in a bowl returns to the bottom. Stable fixed points are therefore often called attractors. A perturbation of a system at an unstable fixed point however will cause the system to leave its position in phase space like a deflection of a ball sitting on top of a hill will cause it to roll down. Unstable fixed points are therefore also called repellers. Linear stability analysis is used to investigate a system's behaviour close to fixed points. Therefore, the differential equations describing a dynamical system are linearised around its fixed points for small deflections from there. To linearise the two-dimensional system in 2.91 we use Taylor's expansion at the fixed point (x_0, y_0)

$$f(x,y) = f(x_0,y_0) + \frac{\partial f}{\partial x}\Big|_{\binom{x=x_0}{y=y_0}}(x-x_0) + \frac{\partial f}{\partial y}\Big|_{\binom{x=x_0}{y=y_0}}(y-y_0) + O(\Delta x^2, \Delta y^2, \Delta x \Delta y)$$

$$g(x,y) = \frac{\partial g}{\partial x}\Big|_{\binom{x=x_0}{y=y_0}} \cdot \Delta x + \frac{\partial g}{\partial y}\Big|_{\binom{x=x_0}{y=y_0}} \Delta y \tag{2.94}$$

where $x - x_0 = \Delta x$ and $y - y_0 = \Delta y$ and $O(\Delta x^2, \Delta y^2, \Delta x \Delta y)$ are quadratic terms which are negligible since deflections are small. Then, locally 2.94 can be written as

$$\frac{d}{dt}\begin{pmatrix}\Delta x \\ \Delta y\end{pmatrix} = \begin{pmatrix}\frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial y} & \frac{\partial g}{\partial y}\end{pmatrix}_{x_0,y_0} \begin{pmatrix}\Delta x \\ \Delta y\end{pmatrix} \tag{2.95}$$

which is a linear equation in (Δx, Δy), thus the linearisation about (Δx, Δy). For further processing of 2.95 it is more convenient to substitute the partial differentials at the fixed points $\left(\frac{\partial f}{\partial x}\right)_{x_0,y_0} = a$, $\left(\frac{\partial f}{\partial y}\right)_{x_0,y_0} = b$, $\left(\frac{\partial g}{\partial x}\right)_{x_0,y_0} = c$, $\left(\frac{\partial g}{\partial y}\right)_{x_0,y_0} = d$, thus

$$\frac{d}{dt}\begin{pmatrix}\Delta x \\ \Delta y\end{pmatrix} = \underbrace{\begin{pmatrix}a & b \\ c & d\end{pmatrix}}_{:=A}\begin{pmatrix}\Delta x \\ \Delta y\end{pmatrix}. \tag{2.96}$$

The eigenvalues are given by the characteristic equation $\det(A - \lambda I) = 0$, with unity matrix I

$$0 = \det\begin{pmatrix}a-\lambda & b \\ c & d-\lambda\end{pmatrix} = (a-\lambda)(d-\lambda) - bc = \lambda^2 - \underbrace{(a+d)}_{=\mathrm{trace}(A)}\lambda + \underbrace{ad-bc}_{=\det(A)}. \tag{2.97}$$

With $\tau = \text{trace}(A)$ and $\Delta = \det(A)$, the eigenvalues are

$$\lambda_{1,2} = \frac{1}{2}(\tau \pm \sqrt{\tau^2 - 4\Delta}) \tag{2.98}$$

which depend only on the trace and the determinant of A. Further are

$$\Delta = \lambda_1 \lambda_2 \tag{2.99}$$

and

$$\tau = \lambda_1 + \lambda_2. \tag{2.100}$$

The diagonalised matrix around the fixed points is then

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}_{x_0,y_0} \longrightarrow \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \tag{2.101}$$

with the transformed coordinates

$$\begin{aligned} \Delta x &\longrightarrow \widetilde{\Delta x} \\ \Delta y &\longrightarrow \widetilde{\Delta y}. \end{aligned} \tag{2.102}$$

After diagonalisation, 2.96 can be written as

$$\frac{d}{dt}\begin{pmatrix} \widetilde{\Delta x} \\ \widetilde{\Delta y} \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\begin{pmatrix} \widetilde{\Delta x} \\ \widetilde{\Delta y} \end{pmatrix} \tag{2.103}$$

with the solutions

$$\begin{aligned} \widetilde{\Delta x}(t) &= \widetilde{\Delta x}(0)\mathrm{e}^{\lambda_1 t} \\ \widetilde{\Delta y}(t) &= \widetilde{\Delta y}(0)\mathrm{e}^{\lambda_2 t}. \end{aligned} \tag{2.104}$$

The eigenvalues can now be used for a simple classification scheme shown in Fig. 2.5. By considering 2.98, 2.99, and 2.100 the domains for different dynamical behaviours can be drawn. τ and Δ represent the vertical and horizontal axes which intersect at

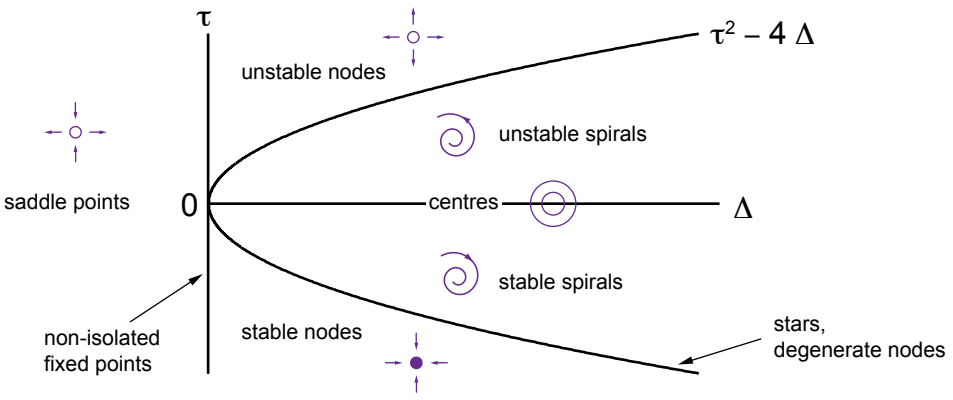

Figure 2.5: Classification of fixed points. The borderlines between the classification domains for fixed points are drawn by τ, Δ and the discriminant of 2.98. The graph was conceptually adapted from [78].

$\tau = \Delta = 0$. The fixed point is a "saddle point" when $\Delta < 0$. For $\Delta > 0$, the parabola draws the border between "nodes" for real eigenvalues with same sign, and "spirals" or "centres" for complex conjugate eigenvalues. Fixed points on the parabola are "star nodes" or "degenerate nodes". There is no isolated fixed point for $\Delta = 0$. Trajectories are attracted by stable fixed points. The stability is global if trajectories converge to it independent of the initial conditions. In linear stability analysis, fixed points are only observed for small perturbations in order to neglect quadratic terms as in 2.94 which facilitates linearisation. In linear systems local fixed points are implicitly globally stable. However, this holds not absolutely true in non-linear systems. Hence, classifications for fixed points in non-linear systems derived by linear stability analysis may only be valid locally.
For complex conjugate eigenvalues, fixed points are centres and trajectories are closed orbits around the fixed points. In this case the system oscillates. A different closed orbit trajectory is a limit cycle. In contrast to centres limit cycles are isolated trajectories where neighbouring trajectories are not closed. They only occur in non-linear systems while centres can also be found for linear systems. A limit cycle can either be stable, unstable or half-stable. It is stable if neighbouring spirals approach the limit cycle. In this case, if the system experiences a perturbation it will eventually return to its closed trajectory. Therefore, limit cycle oscillations play an important role in biological systems, for instance in circadian clocks, by non-equilibrium self-organisation [79]. The Poincaré-Bendixson theorem represents a mathematical method to show that a closed orbit exists in a two-dimensional system. It is valid for a closed sub-set R which is contained in a continuously differentiable vector field. In R exists a trajectory C. Its main statement is, if R does not contain a stable fixed point and C never leaves R, then C is either a limit cycle or a spiral approaching a closed orbit with time. There is also a generalisation for higher than two-dimensional systems by Mallet-Paret and Smith [80].

Design principles for biological oscillators

In the previous section, dynamical systems were introduced from a general view. In biological systems, circuit dynamics are established by time-dependent interactions of molecules.
Novák and Tyson nicely summarise the four basic requirements for non-linear oscillators in biological systems in a review [81]: **Negative feedback** is necessary for reactions to get back to their initial state in order to start a new reaction cycle. **Time-delay** must be sufficiently large since otherwise the back-fed signal causes the reaction to fall in a steady-state. **Non-linearity** of switching mechanisms must be strong in order to generate steep transfer functions. **Time-scales of production and degradation mechanisms** of signals must be balanced.
Time-delay in biological reaction networks arises due to time necessary for transcription and translation reactions as well as by maturation times of proteins. Alternatively also spatial separation were signals have to migrate by diffusion can introduce time-delay. A further way to increase time-delay of negative feedback is a positive feedback loop. Non-linearity is introduced by cooperativity, thus a large Hill coefficient of the feedback function. If the feedback is a series of many sequential reactions it has been shown that also lower Hill-coefficients are sufficient to generate oscillations [82].

2.1.6 Stochastic effects in biochemical systems

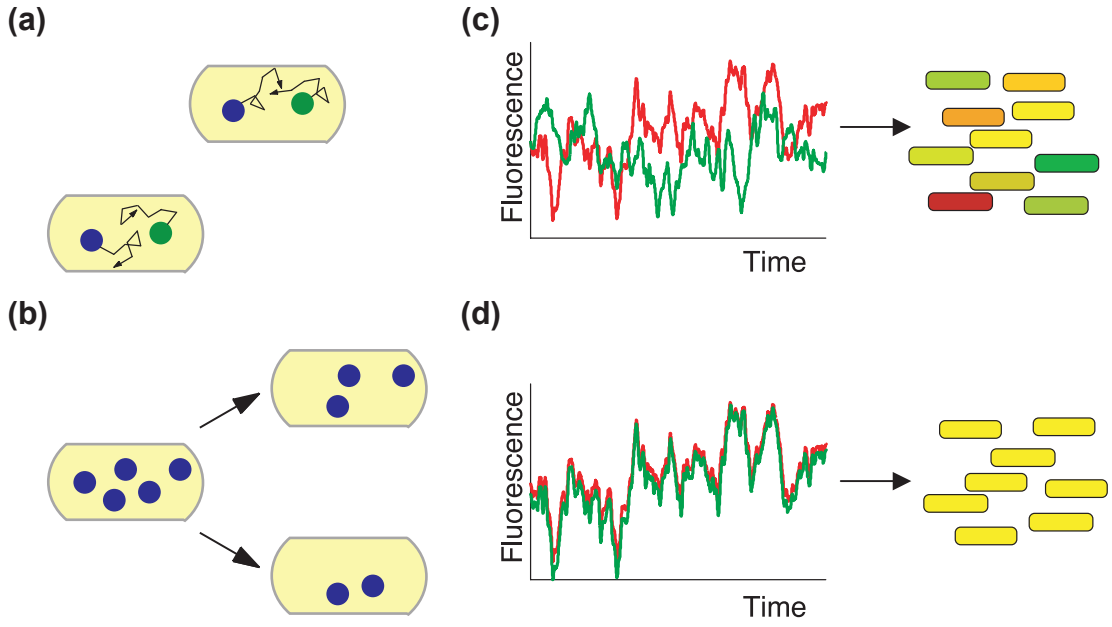

Figure 2.6: Effects in systems with small molecule numbers: Intrinsic and extrinsic noise. (a) Schematic representation of the origin of intrinsic noise. Yellow ovals represent bacteria which contain two different molecules (green and blue circle) like for instance an RNA polymerase and a plasmid. The molecules react as soon as they collide. The molecules diffuse through the cells due to Brownian motion indicated by black arrow headed trajectories. Depending on the random movement, the molecules in the two identical bacteria may meet at different times. As a consequence, the time point when the molecules react cannot be predicted deterministically and is very likely to be different for both systems. (b) Illustration of partitioning error during cell division. Upon cell division two genetically identical bacteria evolve. However, due to uneven partitioning during cell division, the daughter cells may develop differently. (c), (d) A strategy to distinguish intrinsic and extrinsic noise. The red and green time course represent the time development of fluorescent protein levels in a single cell which are controlled by identical regulatory sequences. (c) Intrinsic noise causes both protein levels to develop uncorrelated with each other. Since the fluorescence of both proteins not only develops differently in a single cell, snap shots of cell populations (right) show cell-to-cell-differences in the ratios between protein levels which is expressed by different mixed colours. (d) Extrinsic noise is caused by fluctuations in cellular components between individual cells. In a single cell both fluorescent proteins will be present in the same amount and fluctuate correlated with each other. Cell-to-cell differences occur not in the ratio between both fluorescent proteins but in the total amount causing same mixed colour but different fluorescence intensities (right). (c), (d) reprinted by permission from "The American Association for the Advancement of Science": [83], Science, 2002.

In section 2.1.2, the formal description for chemical reactions was introduced. Molecular interactions in gene regulation were discussed in section 2.1.4. These discussions are valid for large systems in terms of molecule numbers or for the statistical mean of a large number of individual experiments like many cells in a well plate in a plate reader where the behaviour of the population mean is observed. If we regard biochemical systems with identical content, then reactions between molecules depend on the time point of collisions between them (Fig. 2.6 (a)). In systems where large numbers of molecules are present, fluctuations can be typically ignored and reactions may be described deterministically [84]. Studies of systems of the category *in vitro* reaction circuits [57, 85, 56], are typically performed in volumes of tens of microlitres

and concentrations of reacting species are at least nanomolar. Considering a reaction circuit with a component's concentration of $[\mathrm{C}] = 10\,\mathrm{nM}$ in a volume of $V = 10\,\mu\mathrm{L}$, the particle number in this system is

$$n = [\mathrm{C}] \cdot V \cdot N_A = 10^{-9}\,{}^{\mathrm{mol}}/_{\mathrm{L}} \cdot 10^{-5}\,\mathrm{L} \cdot 6.022 \cdot 10^{23}\,{}^{1}/_{\mathrm{mol}} = 6.022 \cdot 10^{9} \tag{2.105}$$

with Avogadro's constant N_A. Consequently, for this example the number of participating particles is high enough to justify a deterministic simulation of reactions. However, biological systems like bacteria are on scales much smaller than these bulk reactions. *E. coli*, for instance, with a volume of about one femtoliter, has typically only one gene for a certain protein. Other reacting species are typically below 100 which manifests in noisy, not deterministically predictable behaviour. Over the past decade, people have got more and more excited by the idea of creating artificial cells which are supposed to be synthetic, cell-scaled systems which contain biochemical reaction circuits and have properties as real cells such as growth, replication, and evolution [86, 87, 88]. To design such tasks with artificial biochemical circuits strategies may be required which provide robustness against stochasticity or exploit randomness to advantage. Even though we are regarding solely problems in biology, randomness comes into play everywhere where ensembles of particles are reduced in size. Miniaturisation is of major interest in many niches of engineering sciences. Examples are lab-on-a-chip devices for medical applications, where minor defects can have strong impact on diagnosis results, or semi-conductors for transistors where doping inhomogeneities may become appreciable through fluctuations in threshold voltages or variations in gate lengths. It is inevitable that if systems are scaled down only small numbers of particles are present and stochastic effects have to be considered. Perhaps the most appreciated effect found in biological systems is the stochasticity of reactions itself. When only small numbers of molecules are present, the fact that the speed of a reaction is determined by the random time point of the collision of both becomes noticeable when two identical reactions are compared as illustrated schematically in Fig. 2.6 (a). In this case the deterministic approach breaks down and reactions have to be treated stochastically. Gillespie's algorithm offers a method to treat reactions on a probabilistic level [28]. This stochastic effect in the context of gene expression is coined "intrinsic noise" [89]. If a bacterium contains two genes for fluorescent proteins controlled by the same regulatory sequence, then intrinsic noise can cause the fluorescence levels of both proteins to fluctuate in an uncorrelated manner as illustrated in Fig. 2.6 (c) [83]. Since these uncorrelated fluctuations happen in individual cells randomly, also the ratio of both protein levels will vary between identical cells as indicated by different mixed colours. "Extrinsic noise" is caused by fluctuations in cellular components as for instance due to environmental fluctuations, stress or partitioning error during cell division shown in Fig. 2.6 (b). This can be also illustrated by means of the two reporter example (Fig. 2.6 (d)). In single bacteria, both genes will be expressed in a correlated manner and are therefore present in a constant ratio. When individual cells are compared (right), the same ratio of the two fluorescent proteins will cause the same mixed colour throughout the colony. However, because the number of components responsible for gene expression may vary, like RNAPs or ribosomes, variations in the overall intensity are likely to occur. The two-colour reporter system is convenient to exemplify the different influences on isogenic bacterial populations. However, there are doubts if these systems always lead to correct conclusions

[90]. Rather recently, it has been shown that partitioning effects during cell division might be underestimated and can also lead to diversity in development of superficially identical systems (Fig. 2.6 (b)) as it would be expected from intrinsic noise. To be more precise, Huh and Paulsson discussed scenarios in which uneven partitioning during cell division reveals the same noise pattern as it would be expected from classical intrinsic noise [91]. In the following, the influence of noise will be demonstrated with

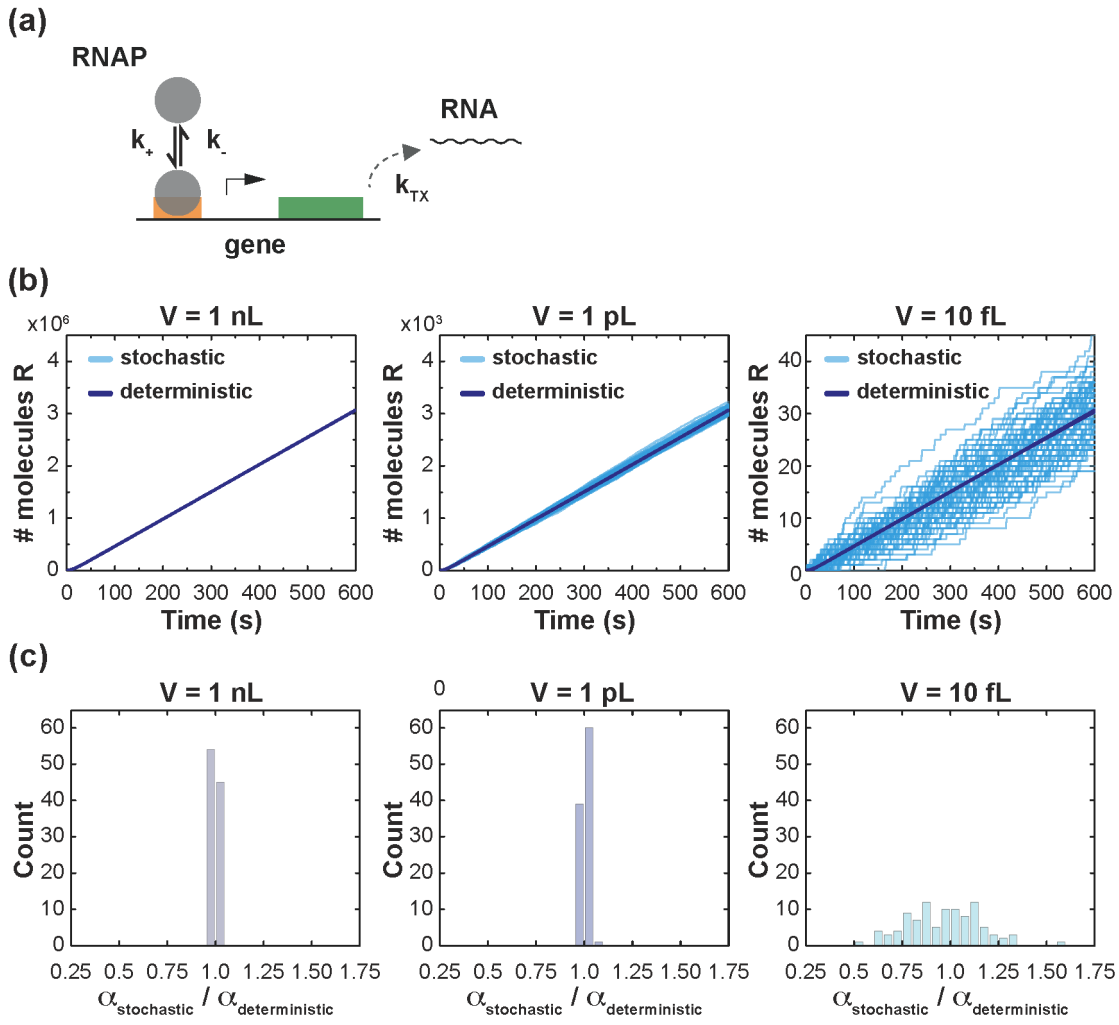

Figure 2.7: Stochastic reaction dynamics: transcription in small volumes. (a) Schematic of the reaction described in 2.106. RNAP binds and unbinds the gene with rates k_+, and k_-, respectively, and eventually transcribes the gene into RNA with rate k_{TX}. (b) Simulation of the reaction 2.106 in volumes of $V = 1\,\text{nL}$, $V = 1\,\text{pL}$, and $V = 10\,\text{fL}$ assuming deterministic (dark blue) or stochastic (light blue) reaction dynamics. For the stochastic case, 100 simulation traces are shown. Interestingly, already for $V = 1\,\text{nL}$ stochastic traces are not completely underneath the deterministically simulated. RNAP and gene concentrations were 10 nM which is converted to particle numbers $6.022 \cdot 10^6$, 6022, and 60, respectively. Simulation was run with COPASI using its built-in Gibson-Bruck algorithm [92]. (c) For comparison purposes, the slope of traces was determined by a linear fit of the form $[\text{R}] = \alpha \cdot t + [\text{R}]_0$ where $[\text{R}]_0 = 0$. The histograms show the distribution of slopes of the stochastically simulated traces normalised by the slope of the deterministic simulation. The linear fit is reasonable since the evolution of RNA is roughly linear in time.

the help of a simple model of a transcription reaction. Let us consider a synthetic *in vitro* transcription circuit which is comprised by only two components, an RNA

polymerase and a gene, transcriptionally permanently active, which encodes for an RNA strand as reaction output (Fig. 2.7 (a)). The system can be modelled in analogy with previous studies [3, 4, 63] and reaction rates adopted from [64] as follows:

$$RNAP + G \underset{k_-}{\overset{k_+}{\rightleftharpoons}} RNAP{:}G \xrightarrow{k_{TX}} RNAP + G + R \tag{2.106}$$

with RNA polymerase *RNAP*, which can associate in a bimolecular reaction with the gene G to form the complex *RNAP:G* with the corresponding rate constant k_+. The complex *RNAP:G* either dissociates into its constituents with the rate constant k_-, or the RNAP transcribes the gene substrate into RNA molecule R with turn-over number k_{TX}. In this process the genes are conserved and the RNA building material, the nucleotides, is assumed to be present in excess on the time scale on which the reaction is regarded and has therefore no influence on the reactions itself. The reaction was simulated in COPASI deterministically with four ordinary differential equations:

$$\begin{aligned}
\frac{\mathrm{d[G]}}{dt} &= -k_+[\mathrm{RNAP}][\mathrm{G}] + (k_- + k_{TX}) \cdot [\mathrm{RNAP : G}] \\
\frac{\mathrm{d[RNAP]}}{dt} &= -k_+[\mathrm{RNAP}][\mathrm{G}] + (k_- + k_{TX}) \cdot [\mathrm{RNAP : G}] \\
\frac{\mathrm{d[RNAP : G]}}{dt} &= k_+[\mathrm{RNAP}][\mathrm{G}] - (k_- + k_{TX}) \cdot [\mathrm{RNAP : G}] \\
\frac{\mathrm{d[R]}}{dt} &= k_{TX} \cdot [\mathrm{RNAP : G}].
\end{aligned} \tag{2.107}$$

However, when this reaction occurs in volumes where only small molecule numbers are present the deterministic formulation of rate equations is not longer sufficient. Correctly, a chemical master equation (CME) has to be solved which treats each reaction according to its probabilistic nature and returns a statistically correct time development of a reaction. The CME for a reaction network is

$$\frac{dP(\vec{x},t)}{dt} = \sum_{k=1}^{r} w_k(\vec{x} - \vec{s_k})P(\vec{x} - \vec{s_k}, t) - w_k(\vec{x})P(\vec{x}, t) \tag{2.108}$$

with the particle numbers of reactants $\vec{x} = (x_1, ..., x_n)^T$, running index k, number of reactions r, the stoichiometry matrix $S = (\vec{s_1}, ..., \vec{s_r})$ and $w_k(\vec{x})\,dt$ the probabilities that the kth reaction takes place in $[t, t + dt]$ which is the propensity. $P(\vec{x}, t)$ is the probability of a chemical reaction network to be in state $\vec{x} = (x_1, ..., x_n)^T$ at time t. Gillespie's algorithm allows the discrete and stochastic simulation of a chemical reaction system as it is described by the CME. The algorithm iteratively simulates numerically one reaction after the other. Initially, the simulation time is set and the reactions, the stochastic rate constants and the initial numbers of reactants are initialised. In the following, two random numbers are generated which are used to determine the waiting time τ for the next reaction to occur as well as which reaction. The probability for a reaction to occur is proportional to the amount of involved reactant. Subsequently, the simulation time is updated $t \longrightarrow t + \tau$ and the numbers of reactants are adjusted according to the reactions which took place. From this point, either a new pair of random numbers is generated and a new reaction cycle is simulated or the simulation is finished because a reactant was completely consumed or

the simulation time was reached.
For stochastic simulations, concentrations are converted into particle numbers as in equation 2.105 and reaction rates are converted into propensities. First order reaction rates are already volume independent and second order reaction rates are converted to propensities by dividing them by the factor $V \cdot N_A$ for corresponding volumes V. Stochastic simulations were performed using COPASI's built-in Gibson & Bruck algorithm which is a derivative of the original algorithm by Gillespie and computationally less expensive [28, 92]. The results of deterministic and stochastic simulations for different volumes are shown in Fig. 2.106 (b). 100 iterative stochastic simulations were performed for each reaction volume, namely $V = 1\,\mathrm{nL}$, $V = 1\,\mathrm{pL}$, and $V = 10\,\mathrm{fL}$ corresponding to particle numbers for both RNAP and gene of $6.022 \cdot 10^6$, 6 022, and 60 for a concentration of $[\mathrm{RNAP}] = [\mathrm{G}] = 10\,\mathrm{nM}$. For $V = 1\,\mathrm{nL}$, all individual stochastic simulation traces are close to the deterministic result. By reducing the system size to $V = 1\,\mathrm{pL}$ stochasticity becomes apparent. With further downscaling to $V = 10\,\mathrm{fL}$ strong variations of the deterministically simulated trace appear. In order to compare the variances for the three volumes, we determined the mean slope for each simulated trace by a simple linear fit $[\mathrm{R}] = \alpha \cdot t + [\mathrm{R}]_0$, where α is the average RNA production rate and $[\mathrm{R}]_0$ is the concentration of RNA when the simulation is started. To start with identical initial conditions we fixed the initial RNA concentration to $[\mathrm{R}]_0 = 0$. For each volume V, the statistical mean $\bar{\alpha}_{stochastic} = \left(\sum_{n=1}^{N} \alpha_{stochastic,n}\right)/N$ for the population of N slopes was determined (which is identical to the deterministic case for large numbers of iterations). The slopes were then normalised to the slope of a fit to the corresponding deterministic trace $\alpha_{deterministic}$, and the frequency was determined for slope bins with a width of 0.05 each, in order to visualise the dispersion. The resulting histograms are shown in Fig. 2.7 (c). The distributions for $V = 1\,\mathrm{nL}$ and $V = 1\,\mathrm{pL}$ seem to be more similar than the traces in Fig. 2.7 (b) would suggest. In fact for $V = 1\,\mathrm{nl}$ all normalised values are within the interval $[0.999, 1.002]$, hence the assignment to different bins is a result of the choice of the bin borders.

Extrinsic noise due to fluctuations in the number of cellular components, for instance, due to uneven partitioning during cell division (Fig. 2.6 (b)) may also lead to diversity. Vessels with volume V contain N_j molecules of species j. If the molecules were filled into vessels independently of each other from a bulk volume with concentration $[\mathrm{C_j}]$, then the probability to find N_j molecules in a vessel is expected to follow a Poisson distribution:

$$p(N, \lambda) = \frac{\lambda^N \mathrm{e}^{-\lambda}}{N!} \tag{2.109}$$

where λ is the expected number of molecules in a volume depending on the bulk concentration $\lambda = [\mathrm{C_j}] \cdot V \cdot N_A$, and N is an integer. Example Poisson distributions are shown in Fig. 2.9 (a) for expectation values of $\lambda = 10$, $\lambda = 40$, and $\lambda = 100$. To illustrate the effect of number fluctuations in components of reaction circuits due to Poisson partitioning we simulated the model 2.106 for different volumes. As for intrinsic noise, 100 simulations were performed for each volume, yet deterministically and with initial concentrations of $RNAP$ and G drawn from a Poisson distribution. This simulation was performed in MATLAB which offers a convenient method to generate Poisson random numbers around a mean value with the function "poissrnd". Since the standard deviation of Poisson distributions is related to its expectation value according to $\sigma = \sqrt{\lambda}$, the degree of diversity depends on the system size and

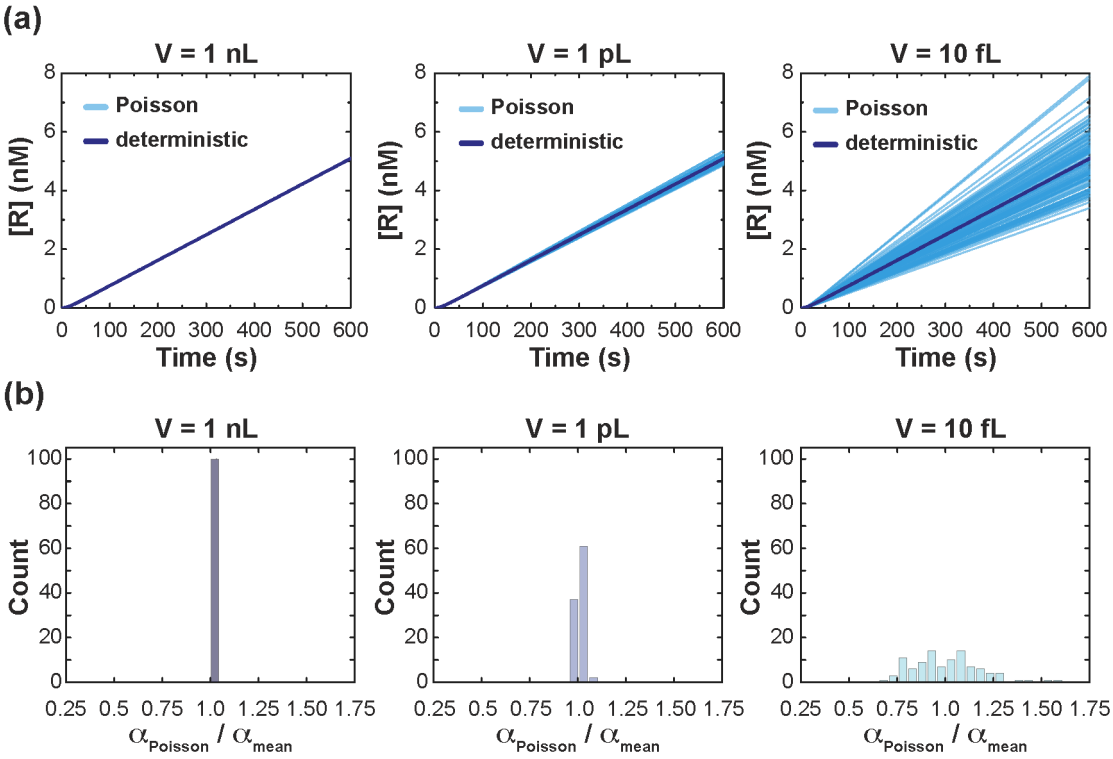

Figure 2.8: Poisson partitioning causes diverse behaviour. Simulation of a simple transcription reaction as shown in Fig. 2.7 (a). Deterministic simulations of reaction equations in 2.107 for initial concentrations of gene G and $RNAP$ drawn by a Poisson probability distribution around $[\mathrm{G}](t = 0) = [\mathrm{RNAP}](t = 0) = 10\,\mathrm{nM}$ for volumes of 1 nL, 1 pL, and 10 fL (light blue). (c) The slopes of concentration time traces were determined by a linear fit $[\mathrm{R}] = \alpha \cdot + [\mathrm{R}]_0$, where $[\mathrm{R}]_0$. The histograms represent the distribution of slopes of traces in (a) normalised by the slope of the mean trace $\alpha([\mathrm{G}](t = 0) = [\mathrm{RNAP}](t = 0) = 10\,\mathrm{nM})$.

decreases with increasing volume. For example, for $\lambda = 100$ the standard deviation is 10% of the mean while for $\lambda = 10000$ the standard deviation drops to 1%. The results of the simulations are shown in Fig. 2.8 (a). The time developments of the concentration of R show as expected increasing variability with decreasing size. For $V = 1\,\mathrm{nL}$, the simulated traces cannot be distinguished from the trace for which initial concentrations were fixed to 10 nM. As the volume decreases the simulated traces for Poisson distributed initial concentrations fan more and more out. To compare the distribution of slopes, we analysed the individual traces in analogy to the stochastic simulations. Histograms of resulting slopes are shown in Fig. 2.8 (b). In probability theory, a measure for the degree of variability of a population from its mean is the coefficient of variation (CV), the standard deviation $\sigma(x)$ divided by the mean $\bar{x}$ for a distribution x,

$$CV = \frac{\sigma(x)}{\bar{x}}. \tag{2.110}$$

With the standard deviation defined as

$$\sigma(x) = \sqrt{Var(x)} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{2.111}$$

with running index i. The CV is useful when data sets are compared with different mean values, thus when the relative dispersion of a distribution is of interest rather than absolute values. The CV enables us to compare the processes in different volumes. The CVs determined by averaging across time points for the number or concentration of R molecules among corresponding time courses are listed in table 2.1 and plotted in Fig. 2.9 (b). For Poisson distributions, the coefficient of variation

Volume V	Stochastic	Poisson
10 fL	0.4152	0.1993
1 pL	0.0455	0.0177
1 nL	0.0014	0.0001

Table 2.1: Coefficient of variation (CV) for transcription in small volumes.

is $CV_{Poisson} = \frac{1}{\sqrt{\lambda}}$. The CV for Poisson distributed particles from a bulk concentration of 10 nM in volumes of 10 fL, 1 pL, and 1 nL are shown in Fig. 2.9 (b). We

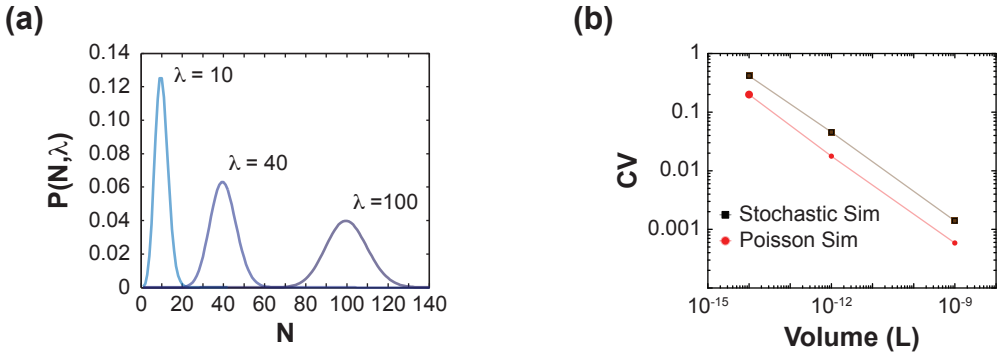

Figure 2.9: Poisson distribution and coefficient of variation. (a) Binomial partitioning leads to probability density functions described by Poisson distributions. Poisson distributions $P(N, \lambda)$ are shown for expectation values of $\lambda = 10$, $\lambda = 40$, and $\lambda = 100$. The standard deviation of Poisson distributions is $\sigma = \sqrt{\lambda}$, thus the coefficient of variation is $CV = 1/\sqrt{\lambda}$. (b) Coefficient of variation in dependence of transcription volumes for stochastic (brown) and deterministic simulation with Poisson distributed initial molecule numbers (red). Lines connecting points are just guides for the eye. The CV for Poisson distributions of molecules with mean concentration 10 nM in same volumes as transcriptions are also shown (blue).

have seen that even though the discussed system was simpler than systems usually observed in context of gene expression noise considerable variations occurred when system sizes approached bacteria-like volumes. Literature on stochastic effects in biological systems usually describes whole expression pathways in growing and dividing cell-like volumes. We here omitted completely any kind of degradation or decay, translation and growth of the compartments. The main purpose of this action is to give an impression how stochastic effects can cause diversity in superficially identical systems. Therefore, it is hard to compare the CV discussed here with discussions in the literature. It has to be added that if transcription and translation are modelled mRNA is the dominant source of noise, thus the level of transcription [93]. Binomial partitioning of circuit components mimics uneven partitioning which could occur during cell division. However, previous studies showed that under certain assumptions

like expression 'bursts' or cellular heterogeneity molecule-number distributions are better described by functions other than a Poisson distribution [94, 91]. A function which was empirically used when classic binomial partitioning was not sufficient is the gamma distribution with probability density function:

$$p(x;\alpha,\beta) = (\beta^{\alpha}\,\Gamma(\alpha))^{-1}\,x^{\alpha-1}\mathrm{e}^{-x/\beta} \tag{2.112}$$

where $\Gamma(\alpha)$ is the gamma function, α is the shape parameter, and β is a scaling parameter. The mean for gamma distributions is $\bar{x} = \alpha\beta$ and the variance $Var(x) = \bar{x}\beta$. Rather recently, we also found diversity in a population of compartmentalized *in vitro* reaction circuits that showed variability beyond Poisson. The results will be presented in section 4.4.2 and were published in [64].

2.2 Diffusion

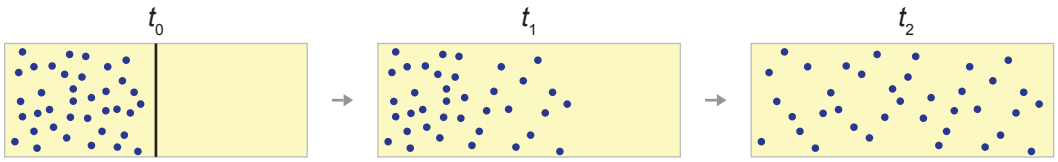

Figure 2.10: Equilibration of a concentration gradient by diffusion. At t_0 particles (blue circles) in a solution (yellow) filled box are confined to one half of the volume by a barrier (black line). After the barrier is removed, a flux of particles develops, caused by the concentration gradient between the left and the right part of the box. At t_1, the particles have started to explore the initially unoccupied volume. After waiting time t_2, there is no concentration gradient and the particles are homogeneously distributed in solution.

Diffusion is a type of mass transfer which causes heterogenous distributions of particles in a solvent to distribute equally in space with time. Diffusion is based on the undirected random movement of particles due to thermal motion. If particles are heterogeneously distributed in space statistically more particles move from region with high to regions with low concentrations which in closed systems eventually causes a homogenous distribution in space as schematically indicated in Fig. 2.10. During the equilibration process, a macroscopic particle flux appears. In 1855 Adolf Fick derived on empirical basis the two fundamental laws of diffusion. Fick's first law is

$$\vec{j} = -D\nabla c(\vec{r}) \tag{2.113}$$

where $\vec{j}$ is the flux of the net number of particles passing through an area per time, D is the diffusion coefficient and $c(\vec{r})$ is the concentration of particles assuming steady state, thus $\dot{\vec{j}} = 0$. Assuming mass conservation concentration can only change in time when in- or out-flow of material happens.

$$\frac{\partial c}{\partial t} + \nabla\vec{j} = 0, \tag{2.114}$$

is the continuity equation. With Fick's first law 2.113 and 2.114, Fick's second law is given by the partial differential equation

$$\frac{\partial c(\vec{r},\vec{t})}{\partial t} = \nabla(D\nabla c(\vec{r},\vec{t})) \tag{2.115}$$

where $D = D(\vec{r})$ for heterogenous media or $D = D_{ij}$, a symmetric tensor for anisotropic diffusion. If D is constant 2.115 reduces to

$$\frac{\partial c(\vec{r},\vec{t})}{\partial t} = D\nabla^2 c(\vec{r},\vec{t}). \tag{2.116}$$

Fick's second law is also known as diffusion equation. The solution of the diffusion equation in three dimensions is generally laborious if an analytic solution exists. It is usually more convenient to solve the diffusion equation numerically, for instance using Matlab's solver for partial differential equations "pdepe". A solution for the problem illustrated in Fig. 2.10 assuming that the box expands to $[-\infty, +\infty]$ is:

$$c(x,t) = \frac{1}{2} c_0 \mathrm{erfc}\left(\frac{x}{2\sqrt{Dt}}\right) \tag{2.117}$$

with initial conditions $c(x, t = 0) = c_0$ for $x \leq 0$ and $c(x, t = 0) = 0$ for $x > 0$, where $c(x,t)$ is the concentration of particles and erfc $= 1 -$ erf is the complementary error function. Solutions of 2.117 for one-dimensional diffusion are plotted in Fig. 2.11 (a) for time points $t = 1, 10, 100, 1000, 10^9$ s. The diffusion coefficient was $D = 1000\,\mu\mathrm{m}^2/\mathrm{s}$. Diffusion coefficients for small molecules in water are in the range $100 - 1000\,\mu\mathrm{m}^2/\mathrm{s}$. Even though the parameters were chosen for aqueous solutions and small solutes, similar examples can be found more likely in material sciences. At metal-dielectrics interfaces, the diffusion of metal atoms into dielectrics can be described by a solution as 2.117 [95]. However, diffusion constants are orders of magnitudes smaller and diffusion is only considered in one direction. Bi-directional diffusion in the discussed example is accounted for by a factor 1/2 in 2.117. Beyond natural sciences, diffusion is widely used to describe phenomena also in other fields such as sociology, finance or economics.

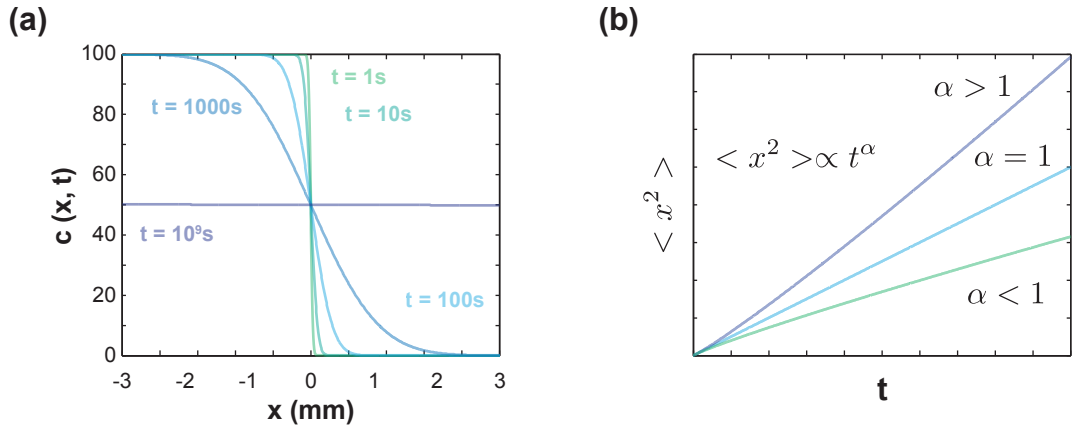

Figure 2.11: Diffusion. (a) An in x-direction infinitely expanding container filled with water is separated at $x = 0$ by a barrier. The left half contains a small molecule in concentration $c = 100$. At $t = 0$, the barrier is removed and the concentration starts diffusion mediated to equilibrate. Concentration profiles are plotted for at $t = 1, 10, 100, 1000, 10^9$ s assuming $c_0 = 100$ and $D = 1000\,\mu\mathrm{m}^2/\mathrm{s}$.

After the empirical introduction of the diffusion equation by Adolf Fick, Albert Einstein derived the same law based on thermodynamic considerations. The Einstein-

Smoluchowski equation relates the diffusion constant D to the friction coefficient ζ:

$$D = \frac{k_B T}{\zeta}, \tag{2.118}$$

where k_B is Boltzmann's constant, and T is the absolute temperature. Stokes' law is a relation for friction and size of a spherical object:

$$\zeta = 6\pi\mu R, \tag{2.119}$$

where μ is the solvent's viscosity, and R is the object's radius. With 2.119, 2.118 gives

$$D = \frac{k_B T}{6\pi\mu R}, \tag{2.120}$$

which is the Stokes-Einstein law for the diffusion coefficient of a spherical object. Even though most objects are not spherical, 2.118 represents an useful estimate for many compounds. As briefly mentioned before, diffusive processes in inhomogenous environments may require a diffusion coefficient $D(\vec{r})$ dependent on the locus. However, the heterogeneity is sometimes periodic like fluid phases separated by membranes as in embryogenesis where morphogens diffuse along roughly periodically arranged cells. Instead of $D(\vec{r})$ it is sometimes sufficient to describe the periodic inhomogenous medium as homogenous medium with a constant effective diffusion coefficient D_{eff}. For instance, one-dimensional diffusive migration of compounds in an environment with periodically placed permeable barriers the effective diffusion coefficient was found to be well described by

$$\frac{1}{D_{eff}} = \frac{1}{D_0} + \frac{1}{\kappa a}, \tag{2.121}$$

where D_0 is the bulk diffusion coefficient between the barriers which are separated by spacings a and permeability κ [96]. Later, the relation 2.121 was further generalised in regard to varying diffusion coefficients in different areas, and for geometries which are better mimics of tissue microstructures, for instance [97, 98]. An intuitive measure in the context of diffusive processes is the diffusion length $x_D = \sqrt{Dt}$ which explains the depth to which particles diffuse in dependence of the time t. Einstein found that the mean displacement of a particle subject to Brownian motion is not proportional to the elapsed time but to its square root. The mean square displacement is

$$\langle x^2 \rangle = 2Dt. \tag{2.122}$$

In some special cases the mean square displacement is not linear in time, thus $\langle x^2 \rangle \propto t^\alpha$ with $\alpha \neq 1$. Such cases are called anomalous diffusion. For instance crowding effects have been found to cause "sub-diffusion", where $\alpha < 1$. Active transport processes like through molecular motors in cells in contrast may result in "super-diffusion", where $\alpha > 1$. The mean square displacement for ordinary diffusion, sub- and super-diffusion are shown in Fig. 2.11 (b) [99, 100].

An important extension of pure diffusion systems are reaction-diffusion systems which combine the diffusively spreading of compounds with local chemical reactions. Local sources or sinks of chemical species may cause systems to remain out of equilibrium

and spatial effects establish time-delays which are required for interesting dynamics as introduced in section 2.1.5. The diffusion equation, expanded by an additional term that accounts for chemical reactions is

$$\partial_t c_i(\vec{r},t) = D_i \nabla^2 c_i(\vec{r},t) + F_i(\vec{c}), \tag{2.123}$$

where $c_i(\vec{r},t)$ is the concentration of species c_i at position $\vec{r}$ and time t, and $F_i(\vec{c})$ describes a chemical reaction of concentrations of involved species $\vec{c}$. Equation 2.123 is a partial differential equation, which describes a reaction-diffusion system. Reaction-diffusion systems has been used widely to model pattern-formation processes. A famous pioneer in this field was Alan Turing who showed theoretically how two morphogens can create spatial patterns due to combined diffusion and chemical reaction [101]. Spatial effects were later associated with pattern-formation in several biological systems [102] like in the embryonic state of *Drosophila* or in the segmental development of the spine in mouse embryos [103]. An example for a synthetic reaction-diffusion system is shown in Fig. 1.2.
In the context of this thesis, a reaction-diffusion system is used in chapter 5 to model diffusion mediated distance dependent communication in spatially separated bacteria.

2.3 Droplet emulsions

The aim of this chapter is to briefly introduce emulsions, discuss requirements for applications of emulsions in science, and present a microfluidic technique to generate monodisperse droplets.
Emulsions are widely known due to their broad application in industry as lotions in cosmetics, and pharmacology, or as milk or mayonnaise in food technology. Emulsions consist of liquid droplets dispersed into a carrier fluid. The two fluids are immiscible, thus the carrier fluid, often termed the continuous phase, seals the tiny droplet microreactors. Surfactants are used for the purpose of stabilising droplets against coalescence. Today, microdroplet emulsions find applications in biological and chemical research. They offer the opportunity to conduct studies in small volumes, to isolate single cells [104, 105, 106] or single molecules [107, 108], and to perform high throughput studies [109, 110]. Using microdroplets, considerable achievements have been made for instance in single-molecule enzymology [111], emulsion PCR [112], and *in vitro* evolution experiments for the selection of ribozymes or functional proteins [113, 114].
A wide range of droplet generation techniques exist. The choice of the emulsification technique determines the size distribution which can be created - ranging from almost monodispersed droplets to distribution widths covering orders of magnitude in diameter. Typical sizes range from hundreds of nanometers up to hundreds of micrometer in diameter [115, 116, 60]. Advantages of microdroplet studies are the low consumption of material due to the small sized microreactors, and of course the simple observation of thousands of reaction compartments in parallel with standard (fluorescence-) microscopy. Conceptually, microdroplet long-time measurements are very similar to studies in a plate-reader with the difference that the 'wells' represented by individual droplets are downscaled from micro- to picoliters. Fig. 2.12a shows the scaling relation between droplet radius and its volume assuming a spheric compartment. An

interesting opportunity for extremely low concentrated compounds is that the concentration can be effectively increased when a bulk solution is compartmentalized into microdroplets. Due to stochastic partitioning generated droplets will contain either one or no molecule, which effectively increases the concentration in droplets occupied by a target molecule or particle. For instance, when one molecule of interest is present in one microliter solution. The concentration is sub-attomolar. Partitioning of this solution into micron diameter droplets leads to an effective concentration jump to the nanomolar range in one droplet (as shown in Fig. 2.12 (b)), thus similar to gene concentrations in cells. If this gene encodes a fluorescent protein, it can be detected due to amplification by gene expression as it was shown for single genes in picoliter droplets [107].

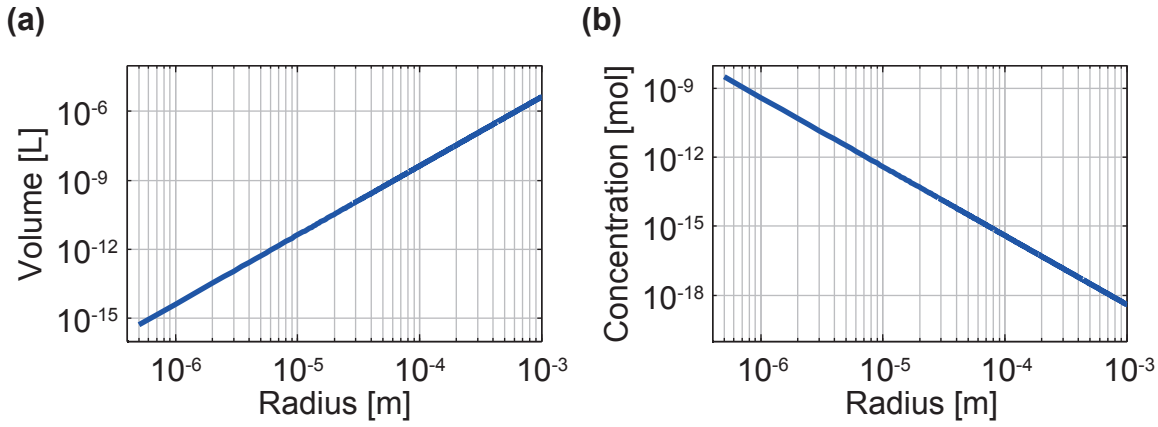

Figure 2.12: Droplet size and concentration in dependence of the radius. (a) Spheric volume versus radius. (b) Concentration for a droplet containing one particle in dependence on droplet radius.

2.3.1 Emulsions for biotechnological applications

For the following considerations the continuous phase is assumed to be oil, the dispersed phase is aqueous if not mentioned else wise. A bulk solution can be considered as a very large droplet with surface tension γ. Depending on the droplet generation technique an emulsion is generated when this solution is deformed through work done by shear stress against the interfacial tension γ. If the shear stress is sufficiently strong larger droplets will be elongated and eventually rupture into smaller ones. During elongation the interface grows while the volume stays constant. The surface energy is proportional to its area, thus elongation increases the energy stored in a droplet's surface. After breaking of a droplet into smaller ones it is energetically favourable for the two droplets to minimise their surface area by returning through coalescence into one droplet. For this purpose, surfactants are utilised. Surfactants (a contraction of *surface-active agent* [117]) are amphiphilic, hence consist of both a hydrophilic and a hydrophobic part. Surfactants stabilise emulsions because they accumulate at interfaces between immiscible substances - like water and oil - in order to minimise their free energy. In the case of water droplets dispersed into oil, a surfactant with a hydrophobic tail group creates a layer between adjacent droplets which acts repulsive due to steric reasons similar as in colloid dispersions [118]. When droplets are used

as bioreactors for biological samples such as cells or biochemical in vitro reaction circuits further requirements in addition to stabilisation properties have to be fulfilled by the surfactant. Biomolecules are often charged and tend to denature when they adhere to surfaces. If reaction volumes are scaled down the surface-to-volume ratio increases. Hence, it is mandatory for a surfactant's chemical structure to minimise interaction between compartment boundary and content. Most commercially available surfactants are equipped with ionic head groups and exhibit extraordinary stabilisation properties due to their strong hydrophilicity. Unfortunately, at the same time biomolecules (especially when charged like nucleic acids) will stick to it due to electrostatic attraction. Today, surfactants with satisfying biocompatibility properties consist of polyethylenoxide or polyethylene glycol headgroups, which show reduced protein adsorption and interactions with cell membranes [119].
The molecular composition of tail groups is adopted to the nature of the continuous phase. However they have to be optimised in terms of chain length. They are supposed to be both, long enough to provide stabilisation when droplets are subject to shear forces or packed densely, and short enough that diffusion of surfactant is not the rate limiting step during droplet formation.
Microdroplets provide a large number of independent experiments which can be observed in parallel. To avoid cross-contamination it is mandatory that no unknown components can pass through droplet-droplet boundaries, typically composed of a thin layer of surfactant and oil. While requirements for surfactants were already discussed, the oil has to be chosen carefully, too. Viscosity and density have strong influences on characteristic length scales in microfluidics. Their relation with droplet production will be discussed in section 2.3.2. To avoid molecular exchange between individual compartments oils are used with a low partitioning coefficient for molecules that are supposed to be isolated. Silicone and organic oils have been used frequently in the past, but today fluorinated oils are preferred. Fluorocarbon oils are both hydrophobic and lipophobic hence the solubility for biological compounds is remarkably low [120]. Polydimethylsiloxane (PDMS), a silicone based elastomer, is frequently used as material for microfluidic chips. Except fluorocarbon oils, many solvents, which would in principle qualify as continuous phase fluid, swell PDMS and are therefore not compatible with the channel material [121, 116].
In this work, exclusively a non-ionic fluorosurfactant from [122] composed of two 6000 g/mol PFPE tail groups and 600 g/mol PEG head groups, called E2K0660, dissolved in fluorinated oil Fluorinert FC-40® has been used. The synthesis path as well as the final chemical structure are depicted in Fig. 2.13.

2.3.2 Generation of microemulsions

Existing droplet generation techniques differ in the way how shear stress is introduced which causes droplets to break. In general, a drop of liquid resting in an immiscible fluid has a spherical shape due to its surface tension. The droplet can be ruptured when it is subject to shear stress through viscous drag applied by the carrier fluid which counteracts the surface tension [123]. The shear stress can be generated through motion of the carrier fluid for instance by steering, vortexing or ultra-sonication. These methods are fast, simple and yield broad size distributions

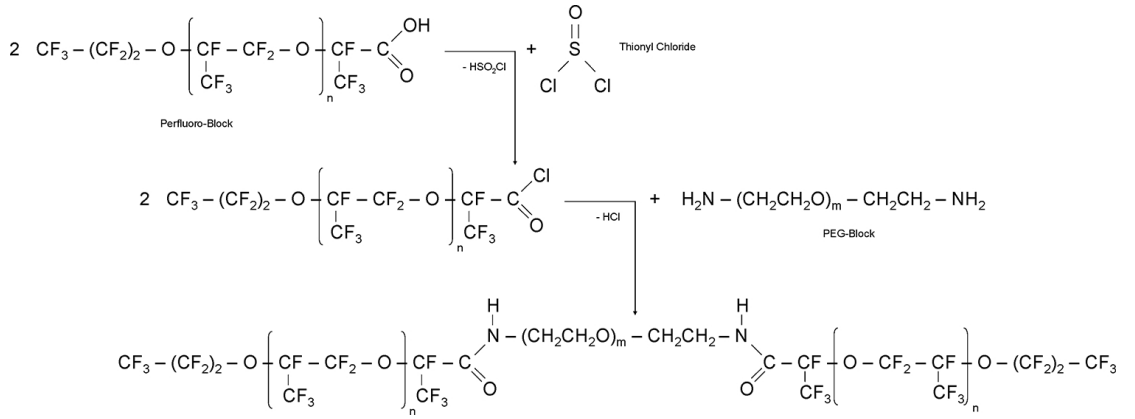

Figure 2.13: Synthesis path and chemical formula of E2K0660 surfactant. Reproduced from [122] with permission of The Royal Society of Chemistry.

of droplets. Furthermore, very small droplets with diameters on the sub-micrometer scale can easily be generated. But, as simple the practical application of those primitive droplet generation techniques is as complicated is the theoretical description of the droplet breaking events due to stochastic processes such as turbulent fluctuations in the fluids' dynamics [116].
In contrast, microfluidics offers a method to generate emulsions consisting of very uniformly sized droplets with a polydispersity index of $I_P = 1-3$. The polydispersity index is defined as

$$I_P = \frac{\sigma(d)}{\bar{d}} \tag{2.124}$$

where d is the size distribution of a population of droplets, $\sigma(d)$ its standard deviation, and $\bar{d}$ its mean [124, 116] which is mathematically identical to the coefficient of variation in probability theory and statistics (2.110). In microfluidics, droplets are generated in a periodic manner at junctions where two immiscible fluids intersect and the break-off happens. The fluid fluxes are controlled independently by syringe pumps or pressure cylinders. There are three different channel geometries that are used: co-flow, cross-flow, and flow-focussing geometry. Co-flow streams are typically realised by concentrically arranged glass capillaries where the inner tube carries the dispersed phase, the outer tube the continuous phase. For this technique, typical droplet diameters range from 80 to a few hundred micrometers with a very low polydispersity index of $1-2$ [116]. Cross-flow and flow-focussing microfluidic chips instead are mostly made from a PDMS mould, but can also be etched in glass or silicon. In Fig. 2.14 a T-junction (cross-flow) [125, 126] and a flow-focussing [127, 124, 128] droplet generator are illustrated. The parameters which determine the droplet production are the viscosities of dispersed (indicated by subscript d) and continuous (indicated by subscript c) phase, μ_d and μ_c, the densities, ρ_d and ρ_c, the interfacial tension, γ, and, depending on the particular setup, either the applied constant volumetric flow, Q_d and Q_c, or pressure, p_d and p_c. How the break-off of the dispersed stream into droplets happens is determined by the channel geometry which forms the local flow-field when flow is laminar. The geometric parameters are the widths of the channels upstream of the junction where dispersed and continuous phase meet, w_d and w_c, the orifice width, w_{or}, or the outlet channel width, w_o, as indicated in Fig. 2.15, and the channel

depth, h, which is constant throughout the chip in case of monolayer PDMS devices. The parameter values for the microfluidic channels used for work presented later are given in table 2.2.
In the following paragraphs, droplet generation is discussed and characteristic parameters in the context of microfluidics are introduced and illustrated by the example of particular channel geometries used for studies reported later in this thesis.
The continuous and the dispersed fluid are injected into separate microchannels. The two fluids meet in a junction where the dispersed phase is forced by the continuous phase stream to form a bulge or jet as illustrated for a T-junction and a flow-focussing device in Fig. 2.15. The growing protruding bulge alters the gap available for the continuous phase flow to pass, and changes the dynamic pressure upstream of the evolving droplet. The local flow-field in the junction depends on the channel geometry and determines the shape of the 'constriction' of the dispersed phase stream. Consequently pressure in the continuous phase increases until a free surface instability, similar to Plateau-Rayleigh instability in water dripping from a faucet, leads to droplet pinch-off eventually [116]. The responsible forces are viscous stress mediated by the continuous phase flow which elongates the bulge, and interfacial tension that leads to a stress jump across the droplet surface. It is thereby essential for controlled drop formation that the channel walls posses proper wetting properties. Droplet breakup only occurs periodically if the channel wall's surface is preferably wetted by the continuous phase. If this is the case then contact between the dispersed phase and channel walls is prevented by a thin film of oil in between. In order to minimise interaction of dispersed phase content and channel surfaces it is therefore important to first fill channels completely with the continuos phase before the aqueous phase is injected.
The fact that droplet pinch-off in microfluidic channels appears in a repeating manner is a result of the characteristic micrometer length scales, and flow rates in microfluidics where flow is typically laminar. In fluid dynamics motion of liquids is described by the Navier-Stokes-equation. Based on this expression, non-dimensional parameters can be derived which are very useful to estimate the transition between characteristic regimes in fluid dynamics. The most popular characteristic parameter is the Reynolds Number Re, which describes the ratio between inertia and dynamic viscosity. It is defined as

$$Re = \frac{\rho v D_H}{\mu} = \frac{2\rho_i Q_i}{\mu_i(w_i + h_i)} \tag{2.125}$$

with the hydraulic diameter, D_H, in a microchannel, and a characteristic velocity, v. Flow is laminar when viscous forces dominate. This is expressed by a small Reynolds number, $Re < 1$. Using the values listed in table 2.2, the Reynolds number for the dispersed phase is $Re_d = 0.8$, for the continuous phase is $Re_c = 0.7$ for the highest flow rates used in our experiments ($Q_{d,max} = 100\,\mu\text{L/h}, Q_{c,max} = 500\,\mu\text{L/h}$), hence flows can be considered laminar. An important characteristic parameter in the context of microfluidics is the capillary number Ca. It describes the ratio between viscous stress and capillary pressure. The capillary number is

$$Ca = \frac{\mu_c v}{\gamma} = \frac{\mu_c G w_d}{2\gamma}. \tag{2.126}$$

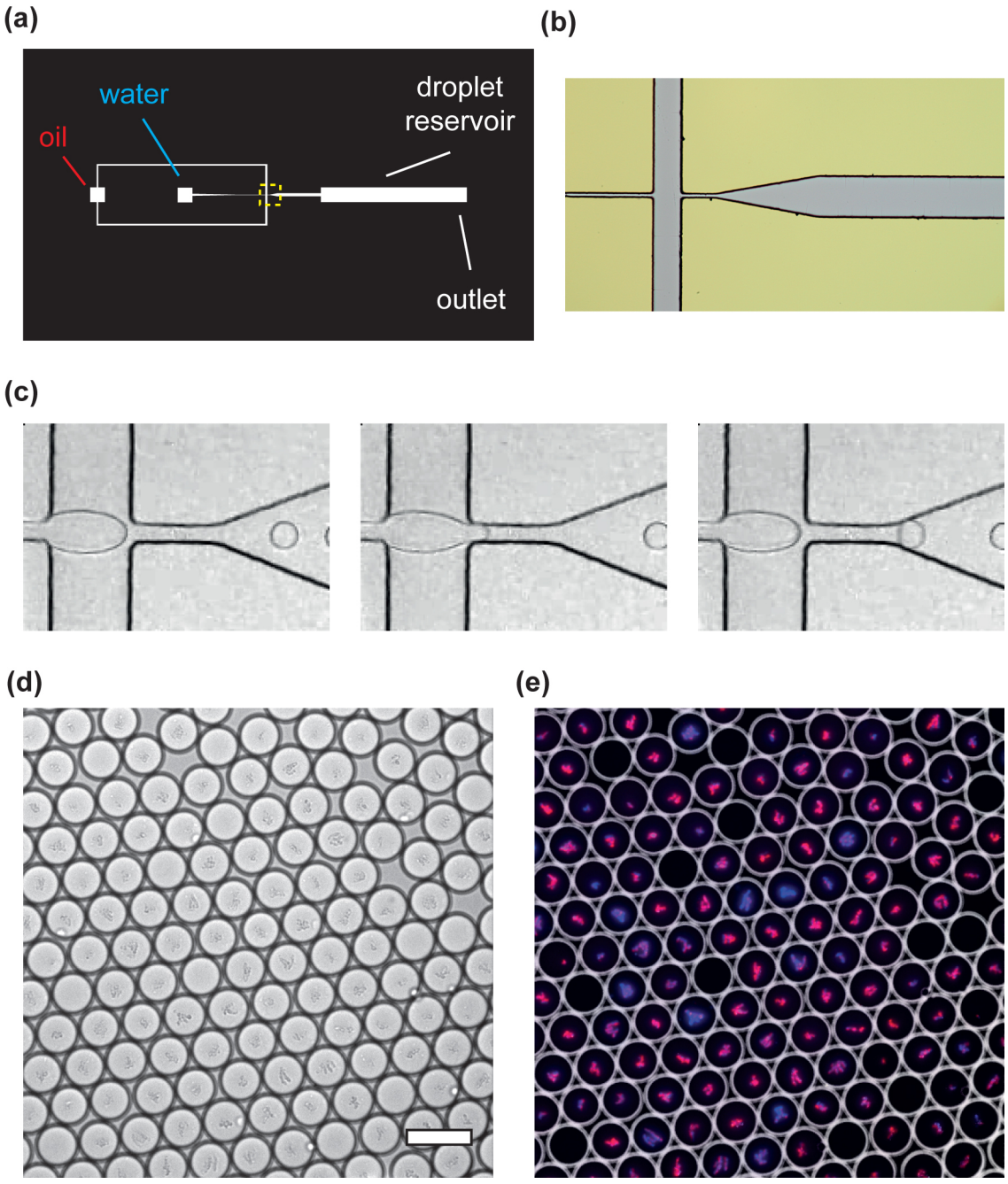

Figure 2.14: Droplet generation in microfluidics. (a) Transparent for soft-lithography with mask for a flow-focussing droplet device. (b) Flow-focus junction of a droplet generation device (indicated in (a) by a yellow box) in SU8 photoresist (grey) on silicon (yellow). (c) Image series of a droplet generation movie. The droplet flux is from left to right. (d) Droplets filled with E. coli bacteria produced in microfluidics stored on an ibidi microslide imaged in bright-field mode and (e) fluorescence mode.

The characteristic velocity, v, is substituted by the product of elongation rate G and the half width of the dispersed phase channel [116]. For flow-focussing channel geometries, a reasonable estimate of the capillary number is given by

$$Ca = \frac{\mu_c w_d \Delta v}{2\gamma \Delta Z} = \frac{\mu_c w_d Q_c}{2\gamma h \Delta Z}\left(\frac{1}{w_{or}} - \frac{1}{2w_c}\right) \tag{2.127}$$

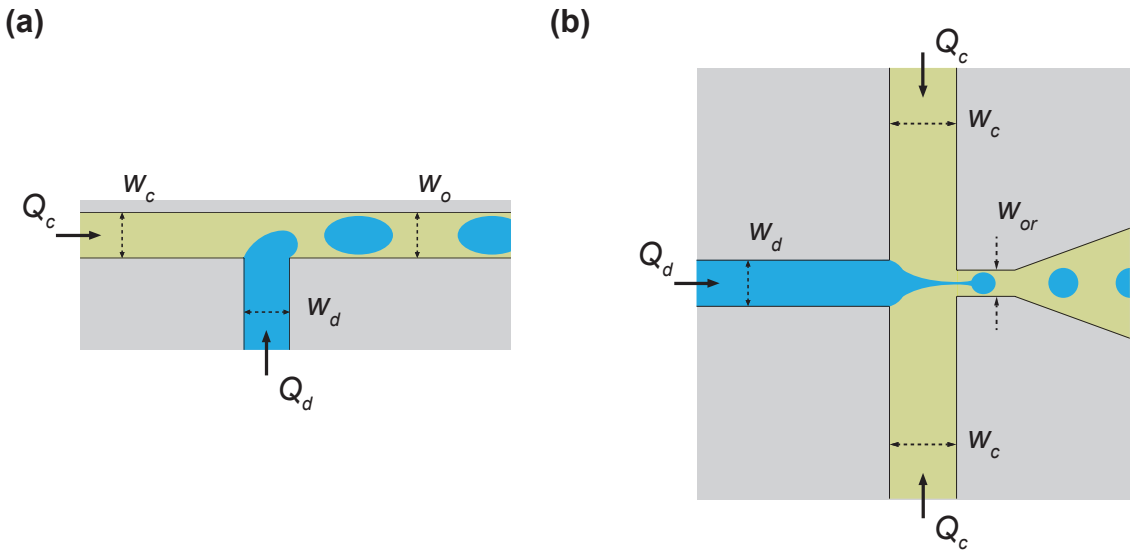

Figure 2.15: Microfluidic chip channel geometries for monodisperse emulsions. (a) Illustration of droplet generation in a T-junction and (b) in a flow-focussing device. Arrows indicate the flow directions of the dispersed (blue) and the continuous (yellow) phase with flow rates Q_d, and Q_c, their inlet channels widths w_d, and w_c, the outlet channel width w_o, and the nozzle/ orifice width w_{or}. Channels are typically planar with height h.

where $G = \Delta v/\Delta Z$ with $\Delta v = v_{c,or} - v_{c,up}$ as the difference of the continuous phase velocity upstream of the junction and in the orifice. ΔZ is the distance from where the dispersed phase channel opens into the junction and the orifice entrance [129]. Using the specific parameters listed in table 2.2 for the flow-focussing device shown in Fig. 2.14 the Capillary number is

$$\begin{aligned} Ca &= \frac{\mu_c w_d Q_c}{2\gamma h \Delta Z}\left(\frac{1}{w_{or}} - \frac{1}{2w_c}\right) = \\ &= \frac{4.1\cdot 10^{-3}\,\mathrm{N\cdot s/m^2}\cdot 2\cdot 10^{-5}\,\mathrm{m}\cdot 1.4\cdot 10^{-10}\,\mathrm{m^3/s}}{2\cdot 3\cdot 10^{-3}\,\mathrm{N/m}\cdot 5\cdot 10^{-5}\,\mathrm{m}\cdot 136\cdot 10^{-6}\,\mathrm{m}}\left(\frac{1}{2\cdot 10^{-5}\,\mathrm{m}} - \frac{1}{2\cdot 136\cdot 10^{-6}\,\mathrm{m}}\right) = \\ &= 1.3\cdot 10^{-2} \end{aligned} \tag{2.128}$$

for $Q_c = 500\,\mu\mathrm{L/h} = 1.4\cdot 10^{-10}\,\mathrm{m^3/s}$, and $\gamma = 3\,\mathrm{mN/m}$ [122]. Literature values for the capillary number in the context of droplet production in microfluidics are typically in the range from 10^{-3} to 10^{1} [116, 130]. Further important characteristics are the volumetric flow rate ratio ϕ,

$$\phi = Q_d/Q_c, \tag{2.129}$$

and the ratio of viscosities λ,

$$\lambda = \mu_d/\mu_c. \tag{2.130}$$

In flow-focussing channel geometries, droplet sizes can be varied by means of the Capillary number, by adjusting the flow-rates. In most studies, the flow rate of the dispersed fluid Q_d is kept constant, while Q_c is varied. When a wide range of flow-rates is screened, different droplet generation mechanism regimes are observed, called squeezing, dripping, and jetting. The transitions between the regimes can be estimated by the capillary number Ca. In contrast to T-junction configurations, no

w_d	$20\,\mu m$	
w_c	$136\,\mu m$	
w_{or}	$20\,\mu m$	
h	$50\,\mu m$	
Q_d	$\leq 10 - 100\,\mu L/h$	
Q_c	$\leq 100 - 500\,\mu L/h$	
Z	$136\,\mu m$	
Density Fluorinert FC-40	$1855\,kg/m^3$	
Dynamic viscosity Fluorinert FC-40	$4.1\,mPa/s$	
Density H_2O	$1000\,kg/m^3$	
Dynamic viscosity H_2O	$1\,mPa/s$	
Interfacial tension FC-40/ H_2O	$3 - 6\,mN/m$	[122, 131]

Table 2.2: Parameters for the operation of flow-focus droplet devices shown in Fig. 2.14.

simple model for flow-focussing channel geometries in terms of tuneable setup parameters has been found so far. Numerical studies attempted to explain the scaling behaviour, but different studies lead to different dependences and a experimental validation is still missing. Just recently, a close similarity between the droplet pinch-off processes in cross-flow and elongation dominated flow was suggested by experimental results. The pressure profiles at three positions on a flow-focussing chip were monitored using Laplace sensors. Each local pressure could be assigned to the droplet breaking progress as the process was recorded simultaneously with a camera at more than 30000 fps. Comparison of the data with previously published work on pressure development during drop formation in a T-junction suggested that both geometries generate droplets by the same mechanism [132, 133]. Further observations at different capillary numbers Ca indicated the existence of the squeezing regime as it was already known for T-junction geometries when $Ca < 0.01$ [134]. For increasing flow rates of the continuous phase, Ca increases, shear forces by the outer fluid become more and more dominant, and droplet generation can be assigned to the dripping regime. In the dripping regime droplets pinch off close to the entrance of the orifice. After a droplet breaks off, the dispersed phase tracks back and starts to grow again. In this regime droplets are highly monodisperse with polydispersity index $< 2\%$ [116], and the size of emerging droplets can be decreased by increasing Q_c, and Ca, respectively. Further increase of the capillary number turns the droplet production mechanism from the dripping regime to the jetting regime. Shape-wise characteristically for the jetting regime is the development of a thin jet which passes the orifice and does not break into droplets at least 3 orifice diameters beyond its exit [135]. Resulting droplets are larger than in the dripping regime and less uniform. The transition from dripping to jetting was shown to occur at around $Ca \approx 1$ [135, 119]. In Fig. 2.16 bright field microscopy images are shown for different Ca. The droplet generation shown at the smallest capillary number $Ca = 6.5 \cdot 10^{-3}$ for $Q_C = 250\,\mu L/h$ tracks back after droplet pinch off and the bulge starts to grow again. This image is also shown in the time series in Fig. 2.14 (c). As the capillary number increases to $Ca = 1.3 \cdot 10^{-2}$ due to higher $Q_C = 500\,\mu L/h$ the droplet pinch off occurs inside the orifice. With further increasing capillary number to to $Ca = 2.6 \cdot 10^{-2}$ due to higher $Q_C = 1000\,\mu L/h$, the

point of droplet pinch off migrates further into the orifice. With the channel geometry of the microfluidic chip used in our setup higher pumping rates are hard to realise since the tubing can't be fixed to the chip anymore due to high pressures. Hence, droplet production in our experiments was usually well in the dripping regime. However, for small Ca the droplet production process reminds to the squeezing regime while it is more likely to be situated somewhere on the boundary between squeezing and dripping.

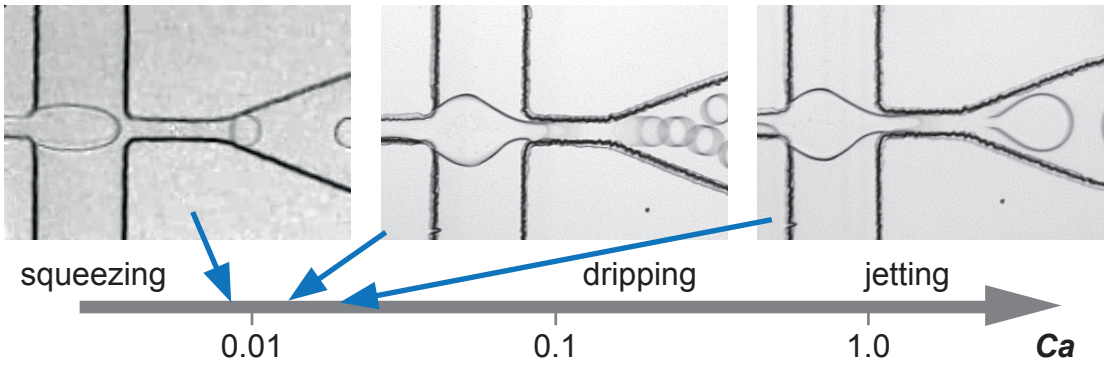

Figure 2.16: Droplet production for different capillary numbers. Capillary numbers are indicated on the logarithmically scaled Ca-axis. The squeezing regime is expected for $Ca < 0.01$. According to the literature, jetting is usually observed for $Ca \geq 1$. Situated between squeezing and jetting is the dripping regime. Bright field microscopy images of the droplet production process are associated with their corresponding Ca values by blue arrows. The tunings of capillary numbers from left to right are $Ca = 6.5 \cdot 10^{-3}$, $Ca = 1.3 \cdot 10^{-2}$, and $Ca = 2.6 \cdot 10^{-2}$, respectively. The image at the left-hand side is identical to that shown in Fig. 2.14 (c).

The previous paragraphs concentrated mainly on the components for, and the generation of emulsions. In studies performed in the framework of this thesis only generation and long-term observation of droplets was applied rather than sophisticated microfluidic setups to manipulate droplets. However, to give a brief impression what the possibilities for microdroplet manipulation are, a few studies should be mentioned. Droplet splitting [136], and fusing [137] was reported mediated solely by the design of channel geometries. Triggered fusion of different droplets can be used to deliberately initiate reactions. Mixing in compartments can be accelerated by winding of channels [138]. In technically more elaborate approaches, droplets were deflected by dielectrophoresis [139] or surface acoustic waves (SAW) [140], which eventually lead to a droplet-equivalent to fluorescence activated cell sorting (FACS), namely fluorescence activated droplet sorting (FADS) [109].
The droplet generation methods applied in the context of this thesis will be described in detail in chapter 3.5.1. Studies in microdroplets will be discussed in chapters 4, and 5.

2.4 Luminescence

Luminescence is a physical phenomenon which occurs when a molecule's electronic state relaxes under emission of a photon after photoexcitation. Section 2.4.1 aims at introducing the basic principles of the two classes of luminescence - fluorescence and phosphorescence. In section 2.4.2, a mechanism which is widely used in the context of fluorescence to determine conformation changes on nanometer length scales is discussed.

2.4.1 Fluorescence and phosphorescence

When molecules, which are exposed to light, start to emit light fluorescence or phosphorescence phenomena occur. Macroscopically these two phenomena can be distinguished by the characteristic time scales on which they happen given by the duration between absorption and emission of light. Microscopically the difference can be nicely illustrated in a so-called Jablonski diagram as shown in Fig. 2.17 (a). In both cases, the molecule's electronic state is excited by absorption of an incoming photon. Absorption of a photon causes an electron to change from a ground state S_0 to an excited singlet-state, S_1 or S_2, depending on the incoming photon's energy, or excitation wavelength, respectively. After excitation, the electron's state relaxes quickly to the lowest excited state through a radiationless transition. From here, it can relax to the ground state S_0 through different paths. The phenomenon is called fluorescence when the electron relaxes to S_0 under emission of a photon. A typical fluorescence lifetime is

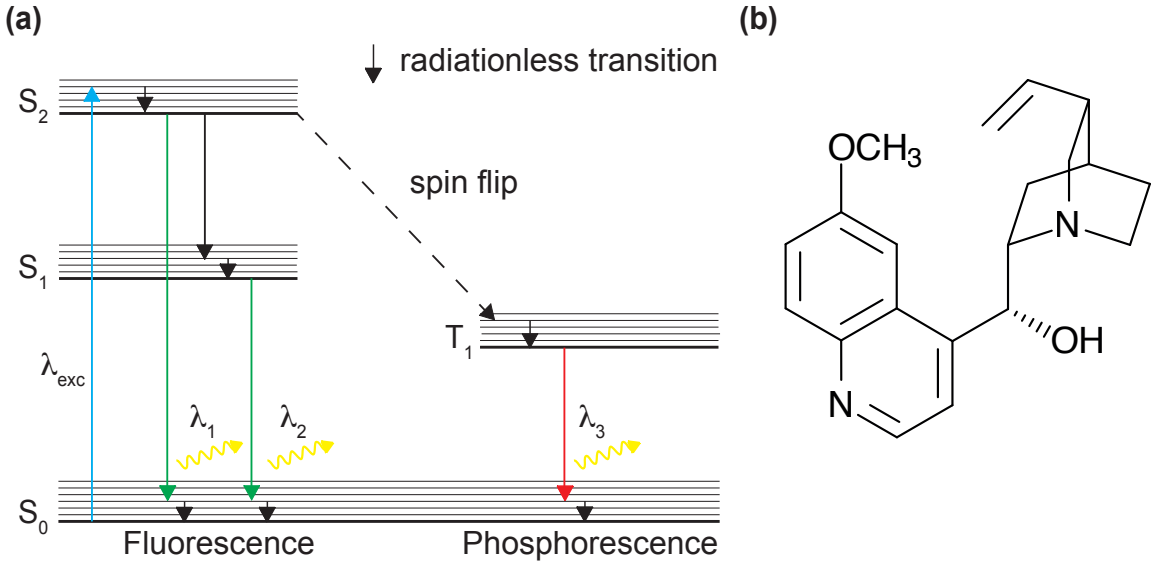

Figure 2.17: (a) Jablonski diagram of transitions between electronic states in a molecule for fluorescence and phosphorescence. A fluorophore's orbit electron is excited by an incoming photon with wavelength λ_{exc} from its ground singlet state S_0 to a singlet state S_2 (blue). The first relaxation is radiationless by heat transfer to the surrounding medium through vibrational substates. The path of relaxation into the ground state determines its luminescence category. The emitted light is associated with fluorescence when the electron in the excited state relaxes rapidly to the orbital electron in the ground state with antiparallel spin orientation (green). In phosphorescence, the electron's spin flips, enters a triplet state (dashed black) and can only relax to the ground state when its spin flips again (red). (b) Molecular structure of quinine adapted from [141].

on the order of 10 ns [142]. In contrast, phosphorescence occurs when the electron's spin flips, thus changes into a triplet-state T_1. Relaxation to a energetically lower singlet state is impossible since the Pauli exclusion principle forbids two fermions to be in the same quantum state at the same time. Thus, relaxation to the ground state dependents on the probability for the spin of the electron in T_1 to flip again, which causes phosphorescence life-times to be typically as long as milliseconds to seconds [142]. The difference between excitation and emission wavelengths is called Stokes-shift. Since the relaxation happens through radiationless intermediate states, the emitted light is less energetically, hence shifted towards red wavelengths. The Stokes-shift enables the observation of only emitters while light with excitation wavelength can be filtered out easily for instance with fluorescence filters or monochromators. Molecules, which have fluorescence properties are called fluorophores or fluorescence dyes. Fluorophores are usually aromatic molecules with several π-bonds which are responsible for the fluorescence properties. Classical fluorophores are in the molecular weight range of 200 to 1000 Dalton. A quite popular example of fluorescence in every day life is quinine in tonic water. The first observation of fluorescing quinine during exposure to sunlight was made by Sir John Frederick William Herschel in 1845 [143]. Its molecular structure is shown in Fig. 2.17 (b). Today, fluorescence is widely used as read-out for instance to sense substances, to measure conformation changes on the nanometer length scale, or to monitor the progress of (bio-)chemical reactions in real-time. However, the usage of conventional fluorescent dyes is limited since chemical modification of a molecule of interest or a characteristic marker is required. Their non-invasive application in living organisms is thus not always possible. For this purpose, fluorescent proteins are today widely used. The first fluorescent protein which was used as marker in living organism was the 26.9 kDa green fluorescent protein (GFP) originally found in the jellyfish *Aequorea victoria* [144]. Today, derivatives of the GFP wild-types exist which cover a wide range of wavelengths of the visible spectrum. Fluorescent proteins can be expressed in an organism simply by utilising its own metabolic machinery. The genetic information is introduced into the organism's genome, for example just behind the gene for a protein whose presence is supposed to be monitored, by standard gene engineering/ transfer techniques. Thus whenever the studied protein is expressed, the fluorescent protein is expressed, too. Consequently, the time evolution of the fluorescence intensity of the fluorescent protein enables direct read-out of the relative concentration of the protein of interest.

2.4.2 Förster resonance energy transfer

Förster Resonance Energy Transfer (abbreviated FRET, also known as Fluorescence Resonance Energy Transfer) is a physical phenomenon occurring in the context of fluorescence techniques which is used to detect and quantify conformation changes on nanometer length scales. In 1948 Theodor Förster observed in a solution containing fluorescein and chlorophyll that energy "travelled" from excited fluorescein to chlorophyll. The efficiency thereby was dependent on the concentration of the molecules, thus the mean distance between them [145]. To be more precise, FRET occurs when the emission spectrum of one dye, called the donor, overlaps with the excitation spectrum of another dye, the acceptor, in a distant dependent manner. When the donor is excited (ideally at its absorption maximum) the absorbed energy is transferred in

a radiationless process to the acceptor which then emits light at the wavelength of its - in respect to the donors absorption maximum - red-shifted emission maximum. The ratio of energy, which is transferred to the acceptor is sensitive to changes on nanometre length scales in distance between both. The rate of energy transfer $k_T(r)$ is

$$k_T(r) = \frac{1}{\tau_D}\left(\frac{R_0}{r}\right)^6 \tag{2.131}$$

with the donor's lifetime in absence of an acceptor τ_D, the distance between donor and acceptor r, and the so-called Förster radius R_0. The Förster radius is determined by the integral of the overlapping donor and acceptor spectra, the donor's quantum yield, the orientation of the dipols, and the refraction index of the surrounding medium. A quite popular measure of changes in distance between two molecules is the FRET efficiency E, which is

$$E = \frac{R_0^6}{R_0^6 + r^6}. \tag{2.132}$$

In 2.132 the Förster radius corresponds to the distance between donor and acceptor for which fluorescence energy is transferred with an efficiency of 50 % (also see Fig. 2.18). Typical values for the Förster radius for good donor-acceptor pairs are in the range of 30 to 60 Å [146]. Since the change of FRET efficiency is inversely proportional to the 6th power of the distance between the molecules, FRET qualifies extraordinarily for studies of conformation changes or only binding events in or of biomolecules.

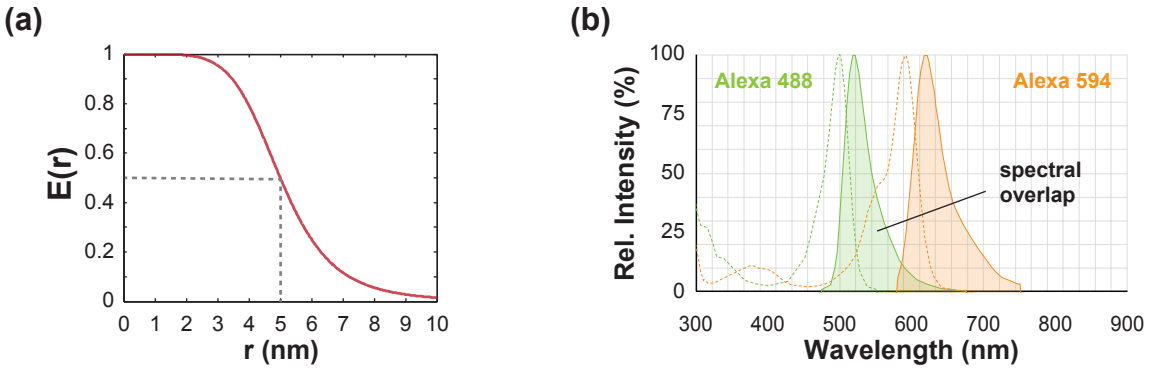

Figure 2.18: Förster resonance energy transfer. (a) Plot of the FRET-efficiency as function of the distance r between donor and acceptor (2.132) for a Förster radius of $R_0 = 5\,\text{nm}$. The Förster radius is the distance between donor and acceptor at which the FRET efficiency is $E(R_0) = 0.5$. Förster radius and the corresponding FRET efficiency E are indicated by dashed lines. (b) Excitation (dashed line) and emission (solid line) spectra of a potential FRET donor and acceptor pair. Alexa 488 and Alexa 594 as FRET pair have a Förster radius of $R_0 = 6\,\text{nm}$. The Förster radius strongly depend on the overlap of donor emission and acceptor excitation spectra. (source: http://www.lifetechnologies.com/de/en/home/references/molecular-probes-the-handbook/tables/r0-values-for-some-alexa-fluor-dyes.html.html). Spectra were generated with life technologies™ online tool "Fluorescence Spectra Viewer".

In the case, that the donor molecule does not emit light, it is called quencher. FRET with two dyes gives much more information of the underlying processes but also requires the measurement of two wavelengths. Quite frequently - and also in the context of this work (chapter 4) - quenchers are used which are sufficient to detect whether a conformation change did happen or not.

3 Methods

In this chapter experimental methods used in the context of this thesis are introduced from a rather general perspective. More detailed information on materials and methods for each result chapter can be found at the end of chapters 4 and 5.
In section 3.1 the determination of concentrations of nucleic acid molecules as well as the measurement of the optical density of bacterial cultures is described. The general principle of fluorescence spectrometers is briefly introduced in section 3.2. Absorption and emission spectra of dyes used as well as a table containing the corresponding excitation and emission wavelengths can be found here. Section 3.3 introduces optical microscopy. Fluorescence spectra and the filter sets used for their imaging in our setup is also contained. Gel electrophoresis was performed for tests of the purity of oligonucleotides and the grade of annealing processes as well as to purify synthesised RNA molecules (section 3.4). Section 3.5 contains technical details on the generation of microemulsions.

3.1 Determination of concentrations and cell densities

Absorption measurements can be used to estimate the amount of a dissolved substance in solution. In general

$$I(d) = I_0 \, \mathrm{e}^{-\alpha d} \tag{3.1}$$

describes the intensity of light $I(d)$, which passed a distance d in a medium with absorption coefficient α which depends of the molecular species, the temperature and the wavelength of the transmitted radiation. The molar extinction coefficient ϵ is

$$\epsilon = \frac{\alpha}{c} \tag{3.2}$$

with the concentration of a dissolved substance c and has the SI unit of $^{\mathrm{m}^2}/_{\mathrm{mol}}$, but more often given in $^{1}/_{\mathrm{M\cdot cm}}$. The Beer-Lambert law is

$$A := c \cdot d \cdot \epsilon = -\ln \frac{I(d)}{I_0} \tag{3.3}$$

which directly relates the extinction coefficient, the penetration length and the concentration with the absorbance A. Hence, upon measuring A the concentration can be in principle easily calculated. The extinction coefficient for nucleic acid molecules depends on their base composition. Responsible for the absorption of nucleic acids are the bases' π-electron systems. Extinction coefficients can be calculated by considering single bases and their interactions with adjacent bases, called the nearest-neighbour approximation [147]. The labelling efficiency can be assessed for DNA strands modified with a label by comparing the absorbance of the DNA strand (typically around

260 nm) and the label at its specific wavelength (absorption maximum). Attention has to be paid when DNA or RNA strands have self-complementary sequences. Paired bases absorb differently than unpaired, an effect which is called hypochromism. For older samples, measuring the absorption spectrum can give information about potential contamination since other molecular species absorb differently. Proteins, for instance, have their absorbance peak around 280 nm. For this thesis, concentration measurements of nucleic acids were performed on an Implen NanoPhotometer.
This instrument can also be used to determine the optical density of bacterial cultures in solution at a wavelength of 600 nm, called OD600. Instead of absorption the underlying physical process used here is light scattering. For measurements of OD600, a sample in a cuvette is transilluminated by light with a wavelength of 600 nm. Standard LB growth medium has relatively strong absorbance. Hence, before cell densities are measured a blank of pure medium is recorded. The instrument used here demands the recording of a blank which is saved and automatically subtracted of cell samples. Plate-readers are powerful instruments in which samples can be measured in 96-well-plates. In the plate-reader used in the context of work reported in chapter 5 both fluorescence and absorption measurements can be performed. This is especially useful when living cells are studied since the absorbance value can be used to normalise fluorescence intensity $I(t)$ by the cell density $OD600(t)$ which gives information about the promoter activity. The time dependent promoter activity is

$$\alpha(t) = \frac{1}{OD600(t)} \left(\frac{dI(t)}{dt} \right), \tag{3.4}$$

where $I(t)$ usually stems from a fluorescent protein which is controlled by a promoter with time dependent activity $\alpha(t)$ [148]. In principle, knowing the optical density enables calculation of the number of cells in the medium. However, this calculation can only be done reliably when calibration curves are determined for the particular instrument and bacteria strain. Conversion values can be found in the literature for *E. coli* ranging over an order of magnitude $10^7 - 10^8$ $^{\text{cells}}/_{\text{mL}}$ at $OD600 = 0.1$ [149].

3.2 Fluorescence spectroscopy

In fluorescence spectroscopy substances with fluorescence properties are studied. The applications are diverse and reach from only detecting molecules to quantitative measurements of conformation changes. As for all scientific applications utilising fluorescence, the Stokes-shift is an important phenomenon since it ensures that the emitted light is due to emission of a fluorescing molecule rather than caused by anything else. Fluorescence spectrometer are universal scientific tools which are used for instance to determine absorption and emission spectra of fluorescing substances or to follow the fluorescence intensity of reporter molecules which give information about chemical reactions or conformational changes. Measurements are typically done for molecule ensembles, thus in bulk. As reaction chambers, cuvettes are used whose volumes are in the range of tens of microlitre up to several millilitre. The basic principle is to excite a sample at a defined wavelength and to observe the sample's emission at a red-shifted wavelength. A scheme of a fluorescence spectrometer setup which is similar to the instrument used in context of this thesis is shown in Fig. 3.1. The excitation

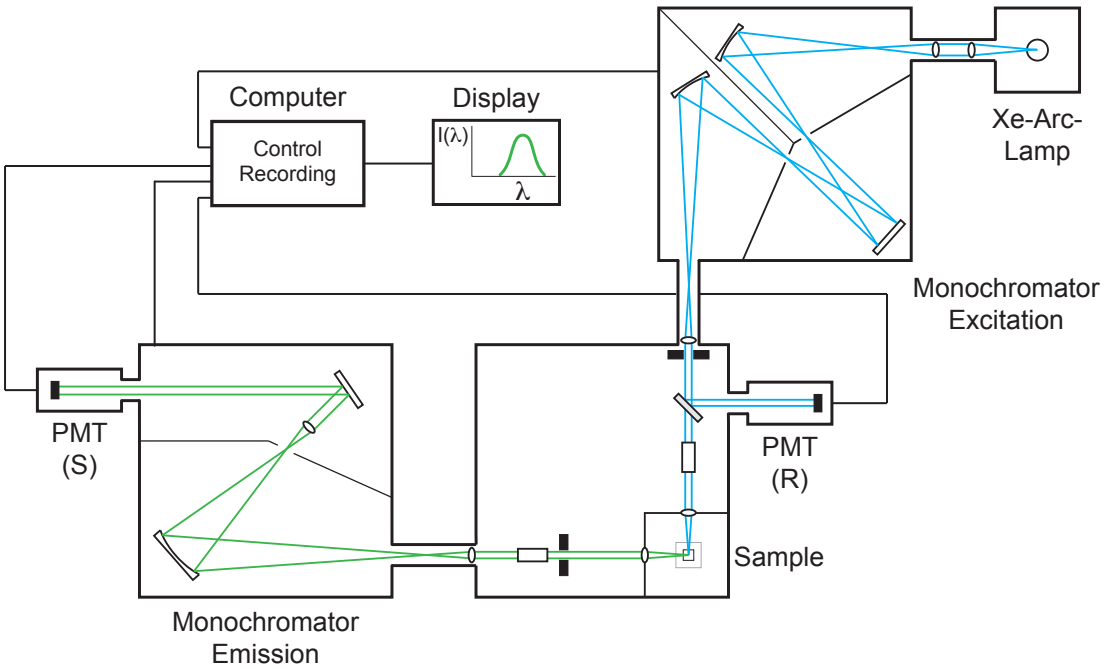

Figure 3.1: Scheme of a fluorescence spectrometer setup. The basic elements of the HORIBA Fluoromax-3 JY Fluorometer used in this work are a Xe-white light source, a excitation monochromator, a reference photomultiplier tube (PMT), a sample holder, an emission monochromator, and the signal detecting PMT. Characteristic for fluorescence spectrometer is the 90° degree angle between excitation and emission light path.

light usually comes from a laser or a white light source, in this case a Xe-Arc-lamp. The wavelength λ_{ex} for excitation is extracted from the light source's spectrum either by a monochromator as shown in Fig. 3.1 or by filter which are transparent only in the range of the specific wavelength. When the light beam leaves the excitation monochromator, it is split in one part which is focussed to the sample and another part which hits a photomultiplier tube (PMT) to be recorded as the reference signal (R). The reference signal is used to eliminate fluctuations of the lamp intensity. The beam which hits the sample excites fluorescent substances as introduced in section 2.4.1 which subsequently emit light. Even though the Stokes-shift strongly reduces background by the excitation light, the emitted light is collected under a 90° angle in respect to the excitation direction to guarantee that no transmitted light is detected. The emitted light leaves the sample and enters the emission monochromator which extracts the signal wavelength which is recorded by a second PMT (S). The here used fluorescence spectrometer is a HORIBA Fluoromax-3 JY Fluorometer controlled by the proprietary software packages DATAMAX or FluorEssence. It offers the consecutive observation of multiple wavelengths which is useful for time course measurements of samples with multiple dyes. For excitation and emission, the spectral width can be adjusted by means of a slit width. For example, for an emission slit width of 5 nm as typically used, the light which is detected by the PMT stems from the wavelength range $\lambda_{det} = [\lambda_{em} - 2.5\,\text{nm}, \lambda_{em} + 2.5\,\text{nm}]$. This is especially useful if the expected fluorescence emission intensity is low for instance due to low concentrated fluorescing species or low quantum yield of the regarded dye. However, increasing slit width is limited by the difference between excitation and emission wavelength which is specific

for every fluorophore. In Fig. 3.2, excitation and emission spectra of dyes are shown. The corresponding excitation and emission wavelengths are indicated for slit widths of 5 nm. The signal which is used as data for further processing is usually corrected for fluctuations in the lamp intensity by the reference signal, thus

$$I(t) = \frac{I_S(t)}{I_R(t)},$$

where $I(t)$ is the fluorescence intensity emitted by the sample $I_S(t)$ divided by the excitation beam intensity recorded by the reference PMT $I_R(t)$. Samples with low concentrations of fluorescing molecules are usually noisy in their emitted intensity. The integration time allows to set the time interval for the recording of emitted photons per data point. Hence, increasing integration time returns a "smoother" signal due to averaging over a longer period which is similar to moving average data smoothing. Since this goes hand in hand with the duration of sample exposure, increasing the integration time will also enhance photobleaching. For fluorescence time-course measurements performed for work presented in the following, the integration time was set to 0.1 s. An effect occurring when dyes are present in a sample is fluorescence cross-talk. In Fig. 3.2 (d) excitation and emission spectra are plotted in the same graph. Cross-talk appears when for instance Texas Red is excited at 598 nm. As can be seen from the excitation spectrum, TAMRA also absorbs at this wavelength (Fig. 3.2 (b)). Hence, the emitted light may contain some contribution of the TAMRA dye even so the Texas Red emission maximum is followed. To correct for cross-talk, a correction factor has to be determined that describes the contribution of one dye to the emitted light intensity of another. The cross-talk for the TAMRA and Texas Red example in our experiments was negligible. However, for the combination of Texas Red and Malachite Green, data had to be corrected for cross-talk (cf. section 4.6.5). The excitation and emission wavelengths for dyes used are listed in table 3.1.

Table 3.1: Excitation and emission wavelengths of dyes in fluorescence spectrometer measurements.

Dye	λ_{ex}	λ_{em}
DFHBI/ Alexa 488	495nm	519nm
Rhodamine Green	504nm	532nm
TAMRA	546nm	570nm
Texas Red	598nm	617nm
Malachite Green	630nm	655nm

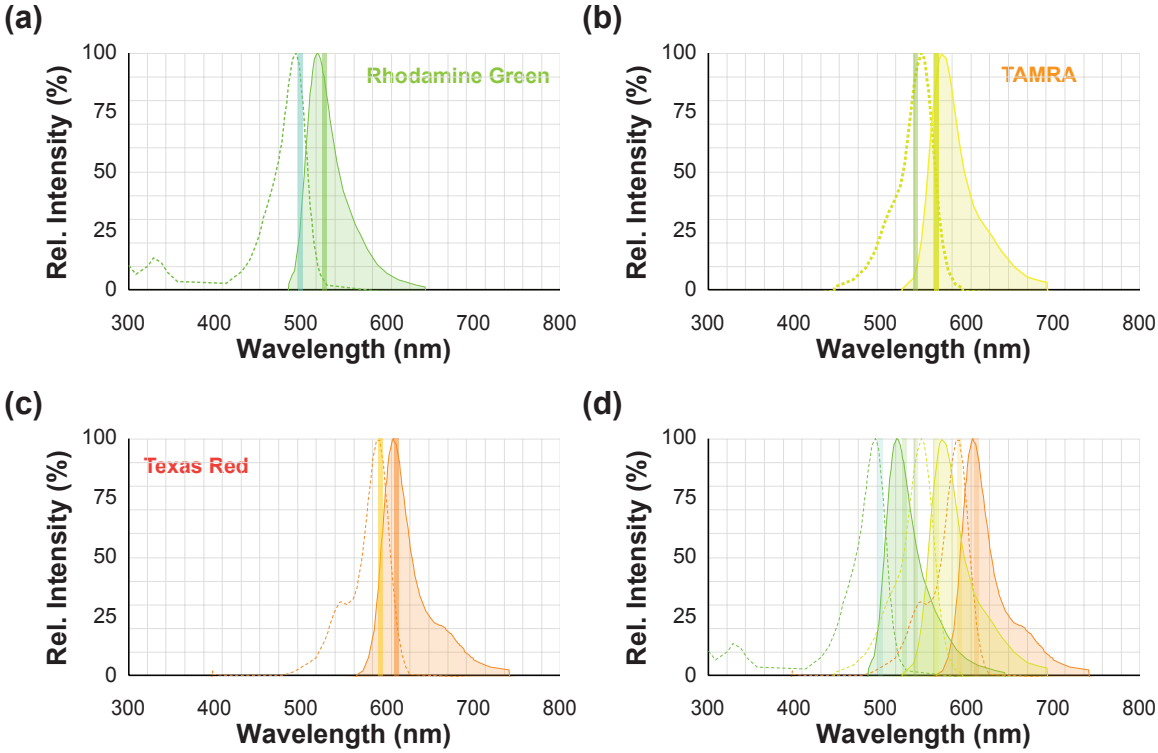

Figure 3.2: Fluorophore excitation (dashed lines) and emission spectra (solid lines) of (a) Alexa 488, (b) TAMRA, (c) Texas Red, and (d) is (a)-(c) together. The used excitation and emission wavelengths for slit widths of 5 nm are indicated as bar and listed in table 3.1. Spectra were generated with life technologies™ online tool "Fluorescence Spectra Viewer".

3.3 Optical microscopy

Optical microscopy is ubiquitous in natural sciences. In general, electro-magnetic radiation in the range of the visible spectrum is used to display small objects by magnification through a lens system. The simplest and probably most prominent optical microscopy technique is bright-field microscopy. Imaging is based on scattering and absorption of white light by a specimen which is either exposed or trans-illuminated. Biological samples are often transparent in the visible spectrum hence have low contrast. Methods to enhance contrast are for instance phase contrast or differential interference contrast (DIC). Phase contrast utilises the phase shift a light beam experiences when it penetrates a specimen. A phase shift ring in the light path can be then used to filter out light which did not pass the specimen. DIC is based on the interference of light which experiences polarisation during transmission through a sample. Here, polarisers are used to filter out light with the same polarisation as the beam before the specimen.
The resolution d is the distance between two features of a structure, like point-shaped objects. Resolution depends on the quality of optical components and the correct adjustment of the microscope setup. However, a fundamental limitation is given by the Abbe diffraction limit. When a wave of light hits an object it is diffracted. Ernst Abbe found in 1873 that due to diffraction, the minimal distance where two point-like objects can still be distinguished as individual objects is

$$d \geq \frac{\lambda}{2\,n \sin\theta} \tag{3.5}$$

where λ is the wavelength of light, n is the refraction index of the medium in which the lens is working and θ is the half angle under which light can leave and enter the lens. An often used estimate for the diffraction limit is $d \approx \lambda/2$ since the refractive index of air is $n_{air} = 1$. However, today lenses are used for high resolution which work in water with $n_{H_2O} = 1.33$ or so-called immersion oil with $n_{oil} \approx 1.5$ (depending on the oil). The product of refractive index and the sine of the half opening angle is called the numerical aperture $NA = n \sin\theta$ and is usually printed on objectives. Nonetheless, low contrast is usually more limiting when dealing with biological samples than the diffraction limit. A microscopy technique that facilitates to reach the diffraction limit is fluorescence microscopy. As in fluorescence spectroscopy, contrast is enhanced by utilising the Stokes-shift of fluorescent dyes and proteins. In Fig. 3.3 (a) a simplified scheme for an epifluorescence microscope setup is shown. Building blocks which are typical for a fluorescence microscope are a light source, an excitation filter, a dichroic mirror, and an emission filter. The light source can either be a white light source (typically mercury arc lamp), a laser or a LED depending on the particular case. If a mercury lamp is used, the characteristic spectrum should be considered when the filter set is chosen (cf. Fig. 3.4 (c) and (f)). Through the excitation filter, the wavelength is chosen for exposure of the sample. The dichroic mirror reflects and transmits light in a particular range of wavelengths in order to prevent excitation light to reach the detector. The fluorescent proteins or dye modifications in the specimen, which are exposed through the objective, are excited and subsequently emit light with a red-shifted wavelength. The emitted light is collected by the objective and passes first through the dichroic mirror and then through the emission filter. The fluorescence

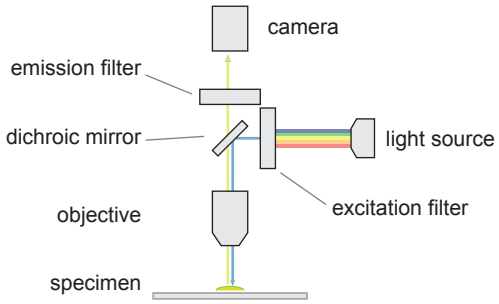

Figure 3.3: Principle of epifluorescence microscopy. A white light source emits light with a broad range of wavelengths. The excitation filter blocks the part of the spectrum which is not desired for exposure of the specimen. The excitation light is reflected to the sample because the dichroic mirror is only transparent for light with higher wavelengths. Fluorophores in the specimen are excited by light incoming through the objective and subsequently emit light with a red-shifted wavelength which passes the dichroic mirror. The emitted light is again filtered by an emission filter and recorded by a camera.

light is eventually recorded by a camera. Because ideally light with all wavelengths but the emitted wavelength is blocked by the emission filter, only emitters are imaged by the camera, thus enables high contrast imaging. Photobleaching when the sample is not observed is prevented by a mechanical shutter that sits directly in front of the light source. The microscope used in the context of this thesis was a fully automated inverted, epifluorescence microscope Olympus IX81, equipped with filter changer, focus drive, x-y-stage, halogen light source for bright-field, and mercury light source for fluorescence microscopy. Spectra of the dyes observed on the microscope as well as the ranges of wavelengths in which the corresponding filters transmit are shown in Fig. 3.4. To provide a constant temperature of 37 °C during observation, an incubator box (Okolab) was used.

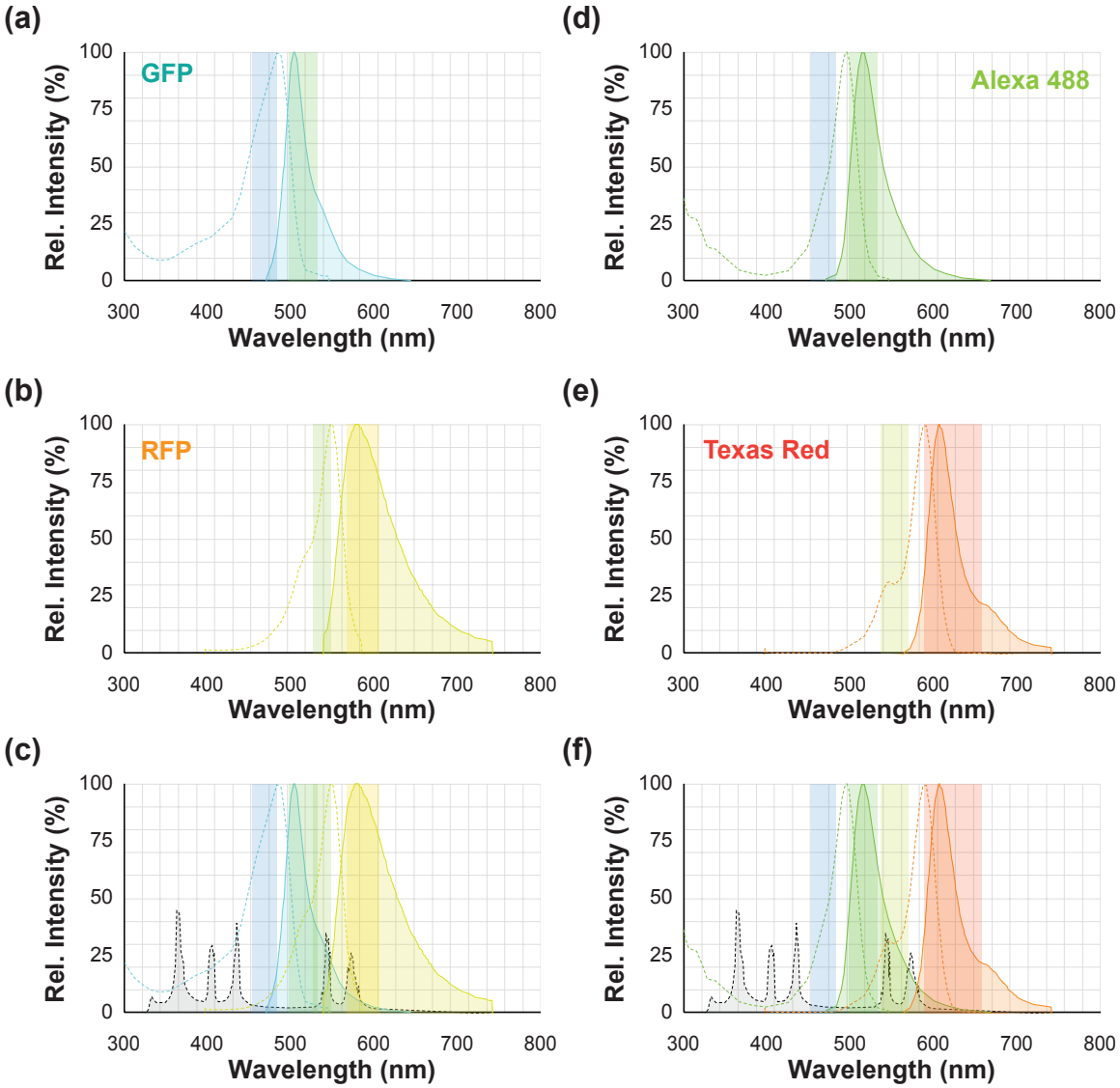

Figure 3.4: Spectra of fluorescent dyes and proteins. Excitation (dashed line) and emission (solid line) spectra of (a) eGFP and (b) RFP. The spectral window which passes through the filter is indicated as transparent bar. (c) eGFP and RFP spectra in one plot with the corresponding filters as shown in (a) and (b). A typical spectrum for a Hg-lamp is indicated as grey area in (c) and (f). (d) and (e) Excitation and emission spectra for Alexa 488 and Texas Red with their corresponding filter sets. Alexa 488 was observed with a filter set for eGFP. (f) Merged Alexa 488 and Texas Red spectra from (d) and (e). Even though, the Texas Red filter in the excitation light path does not exactly overlap with the dye's excitation maximum, the high intensity peak of the Hg-lamp causes efficient excitation at the excitation spectrum's left shoulder. Spectra were generated with life technologies™ online tool "Fluorescence Spectra Viewer".

3.4 Gel electrophoresis and RNA/DNA PAGE purification

Gel electrophoresis is used in biological sciences to separate biomolecules by their size and molecular weight either to analyse samples' content or for purification purposes. Gels made from agarose, are easy to handle and are used for larger proteins (200 kDa) or DNA fragments which are longer than $\approx$ 200 nucleotides. For purification of proteins and nucleic acids, as well as for analysis of shorter fragments polyacrylamide gels are used, commonly termed PAGE gels (abbreviated for "polyacrylamide gel electrophoresis"). Native PAGE gels are used to identify and eventually extract complexes occurring in a sample. The DNA strands migrate through the gel matrix in their native buffer, hence secondary and higher order structures remain stable. Denaturing gels are used when single strands are studied. Denaturing agents such as urea and formamid are used to lower melting temperatures, break hydrogen bonds and prevent complementary sequences from re-hybridising. Due to secondary structures quantitative statements regarding the length of oligonucleotides are usually hard to make with native gels. In denaturing gels, information about the length of an analysed nucleic acid strand can be made since only linear polymers are observed. We here used polyacrylamide gels, typically with 10% density, casted in 8 cm x 8 cm Novex gel cassettes (Art.-Nr.: NC2010).
For analysis gels, combs with 15 wells were used, for purification gels we used combs with three wells. Gels were mounted on a vertical Peqlab gel system which offered the opportunity to temperate using a water circulation temperature control. We applied a voltage of 100 V and ran the gels typically $1 - 1.5$ h. Samples which are loaded on gels are usually mixed with a so-called gel loading buffer (table 3.2). It contains charged reference dyes which migrate in the same direction in an electric field as DNA. In 10% TBE PAGE gels, bromophenol for instance migrates approximately as fast as a 35 nt DNA fragment and can therefore be used as read-out for the migration progress. Table 3.3 lists DNA strand lengths for which Bromophenol Blue and Xylene Cyanol indicate the migration speed depending on the gel density. Analytes are made visible by special stains. Nucleic acids are typically visualised by intercalating dyes, which incorporate between nucleobases. Usually the dyes' quantum yield increases with incorporation, thus improves the contrast between dye and dye-DNA complexes. We stained our gels by 30 min incubation in 1 x TBE buffer which contained 1 x SYBR Gold (Life technologies, Art.-Nr.: S-11494) in a container protected from ambient light. Gels were finally imaged on a gel documentation system. Since the amount of dye depends on the amount of oligonucleotide, the mass of analyte has to be considered before sample preparation. A lower limit for imaging is $\approx$ 50 ng per lane. We typically chose 100 ng in order to visualise minor bands which might be too faint to detect when smaller amounts of material are used.

For purification, it is important to use large amounts of material because samples may not be stained by intercalating dyes. Instead, the band of interest is detected due to absorption of UV-light which is only visible when enough sample is used. We typically loaded 200 µL high concentrated sample (> 20 µM) to the wells. After running, the gel is removed from the cassette and regions with high amounts of DNA

Table 3.2: Gel loading buffer (6x)

Xylene Cyanol FF	0.25%
Bromophenol Blue	0.25%
D (+) Sucrose	40%

Table 3.3: Reference dyes and migration [150]

	Bromophenol Blue	Xylene Cyanol
6%	65 bp	250 bp
8%	45 bp	220 bp
10%	35 bp	120 bp
20%	15 bp	50 bp

are localised. Since UV light destroys biological molecules, it is important that the exposure is only short. When the band was identified, the region in the gel was extracted using a razor blade. The gel pieces were then chopped into 1 mm slices and incubated in 1.5 ml reaction tubes containing 350 µL elution buffer (0.3 M sodium acetate at pH 5.2) at 42 °C overnight. After incubation, the supernatant was extracted from the reaction tubes and mixed with 1 ml freezer cold 100% ethanol. The reaction tubes were then mixed thoroughly and incubated in a freezer for at least 30 min at −80 °C or at least 8 h at −20 °C. After incubation, the samples were spun in a centrifuge with 12000 rpm at 4 °C for 15 min which resulted in precipitation of the RNA or DNA, respectively. Eventually, the supernatant was removed carefully and the pellet was dissolved in nuclease free water for quantification and further processing.

3.5 Droplet generation

In this section, droplet generation is introduced from a technical point of view for the techniques applied in the results chapters. Theoretical fundamentals were discussed in section 2.3. Fabrication of PDMS devices as well as droplet generation in microfluidics is topic of section 3.5.1. In section 3.5.2 an alternative droplet generation technique is introduced, coined the shaken-not-stirred method.

3.5.1 In microfluidics

Water-in-oil droplets were generated in microfluidic flow-focussing devices as shown in Fig. 2.14.

Fabrication of microfluidic chips

Silicon chips with photo-lithographically imprinted channels, in the following called "masters", were arranged in Petri dishes. PDMS elastomer and curing agent (SYLGARD ®184 SILICONE ELASTOMER KIT) were mixed in a ratio 10 : 1 (w/w) by

vigorously stirring in a glass beaker. Gas bubbles which were generated by stirring were removed by evacuation in a desiccator. The bubble-free PDMS mix was casted onto the master in the Petri dish to form an approximately $3-5$ mm layer. The filled petri dishes were again evacuated in a desiccator, and, thereafter baked in an oven pre-heated at 80 °C for at least 1 h until the liquid PDMS was completely cured. The Petri dish was removed carefully and the PDMS-silicon master compound was cut into single master sized pieces utilising a scalpel. The PDMS was then cut along the edges of the silicon master while the chips were oriented with the thick PDMS layer looking downwards. The PDMS layer with the engraved channels was then removed while paying attention that the stress between silicon and PDMS did not cause the master to break. In each PDMS chip holes were punched for the water, and the oil inlet, as well as the outlet using a biopsy punch with 1 mm inner diameter. The PDMS chips were washed together with microscope slides (1.5 mm thickness) in a sonicator alternating in nanopure water, isopropanol, and nanopure water. The cleaned chips and microscope slides were thoroughly dried with a N_2 gas pistol. The channels imprinted in the PDMS chips are closed by bonding them to the microscope slides. In a plasma cleaner, the PDMS chips were arranged such that the channels were on top together with the microscope slides. After activation of the surfaces in the oxygen plasma the two upper sites of PDMS and glass were brought into contact which results in a covalent bond. To further enhance the bonding, glass-PDMS chips were baked again at 80 °C in an oven for at least 2 h. As described in section 2.3.2 the wetting properties are essential for efficient droplet generation. To maintain nicely hydrophobic channel walls, the chips are baked again for 3 h at 200 °C. It was shown that this step yielded a glass surface with comparable hydrophobicity as native PDMS, supposably due to a thin layer of evaporated PDMS [151].

Droplet generation in microfluidic flow-focussing chips

Droplets were generated on an inverted Olympus IX-71 microscope. The droplet generation process was followed in real time on a computer display using an Andor Luca CCD camera.
The oil and water phase fluids were filled into 1 mm syringes and connected to the corresponding inlets by approximately 1 cm long pieces of hypodermic needles and tygon tubing. The connectors were made by carefully cutting hypodermic needles (Sterican, G22, B. Braun) using a side cutter such that the cut face was still circular rather than squeezed by the applied force. Syringes, connectors, and tubing were first assembled and then liquids were taken up. For droplet production without perturbation it is necessary to remove air bubbles completely from syringe and tubing. Syringes are then mounted onto individual syringe pumps (TSE systems) and the connected tubing plugged into the microfluidic chip. The oil used was a mix of Fluorinert®FC-40 (Sigma-Aldrich) with 1.8% $^{w}/_{w}$ EA surfactant (E2K0660, RainDance Technologies, chemical structure shown in Fig. 2.13). It is important to first fill the channels completely with oil. As the oil level starts to incline in the reservoir tubing, the dispersed phase is carefully injected into the chip. During droplet production, the emulsion was stored in a $4-6$ cm piece tygon tubing connected to the chip's outlet by a hypodermic needle connector. When sufficient emulsion was produced, the flow rate of the dispersed phase was turned off in order to clear the channel from the aqueous phase.

For long-term storage, the 'reservoir tube' was removed from the microfluidic chip and the emulsion was poured into a channel on an ibidi microslide (ibidi µ-slide IV 0.4). When droplets were produced with different contents, new PDMS chips, and syringe/connector/tubing were used to prevent cross-contamination. When successively produced droplets were mixed in an ibidi microslide, the slide was gently spun in order to distribute droplets uniformly in a monolayer. Finally, the inlets of the ibidi microslide were sealed using a standard PCR tape to prevent evaporation. In rare cases, a heating plate instead of an incubator box was used to maintain operation temperatures. Then, the inlets were filled with a layer of 10 µL nano pure water covered by 10 µL hexadecane to prevent evaporation due to the vertical temperature gradient.

3.5.2 Shaken-not-stirred

Alternatively, a very primitive method was applied. Droplets discussed in chapter 4 were generated with the 'shaken-not-stirred' method. Therefore, we mixed 10 µL reaction reaction mix with 45 µL FC-40 Fluorinert oil (Sigma-Aldrich, F9755) containing 1.8% $^{w}/_{w}$ E2K0660 (RainDance Technologies, Lexington, MA02421, USA) in Protein LoBind tubes (Eppendorf) using a benchtop Vortex mixer for 60 s at maximum speed. A volume of 45 µL emulsion was transferred with a pipette from the reaction tube to an channel on an ibidi µ-slide IV 0.4 for microscope measurements. The flow chambers were sealed with PCR tape to protect the sample against evaporation. This technique offers the possibility to generate a big range of differently sized compartments in a very short time.

4 A biochemical *in vitro* oscillator

Biological networks can perform highly sophisticated tasks such as signal processing, computation, or the orchestration of molecular processes in time and space. In the past decade a variety of increasingly complex artificial, programmable molecular circuits have been demonstrated *in vitro* as introduced in chapter 1.1.2. On a fundamental level, such systems can be used to prototype and analyse subsystems of more complicated naturally occurring circuits, and thus serve as a training ground for understanding biological complexity. Beyond biology, *in vitro* molecular systems offer considerable design flexibility using a limited number of well characterised components [3, 6, 5]. Thus, they constitute an ideal platform for developing nanosystems that operate in a cell free, well known environment and exploit attractive features of cellular machinery, such as the ability to replicate, self-assemble, and compute at the nanoscale.
In the following, the concept of "genelets" is introduced in section 4.1, which can be in principle used to create arbitrarily complex networks using only well-known and characterised components [152]. In section 4.2, a transcriptional circuit consisting of chemically connected genelets is discussed which will be used in the later sections as molecular clock. How this oscillator circuit was used to coordinate simultaneously the performance of DNA nanodevices as well as the synthesis and annihilation of functional RNA aptamers will be presented in section 4.3. We further utilised the transcriptional oscillator as model systems to study the effect of downscaling to cell-like sizes on reaction networks. In section 4.4 the results of two different compartmentalization methods will be presented and discussed.

4.1 Switchable genes

An approach to build scalable biological reaction networks from scratch *in vitro* is based on the so-called "genelet concept". Genelets are DNA molecules whose transcription activity can be regulated, as introduced for gene regulation in section 2.1.4, by activator or inhibitor molecules. In contrast to genes in naturally occurring systems, genelets only encode for a short RNA strand which is not further used as template for protein synthesis. Since genelet circuits completely omit the DNA replication as well as the translation stage, only a small number of molecules is necessary - such as genelets, activators and inhibitors made from DNA and two enzymes, an RNA polymerase and a ribonuclease - to build up network architectures. The RNA molecules which are transcribed from activated genelets by RNAPs can modulate the transcription activity of another genelet in an activating or inhibiting manner. Hence, upon sequence design, transcription circuits which show interesting properties can be created due to chemical wiring of single subunits driven by only two enzymes. As the components of genelet circuits are relatively simple also the occurring reactions

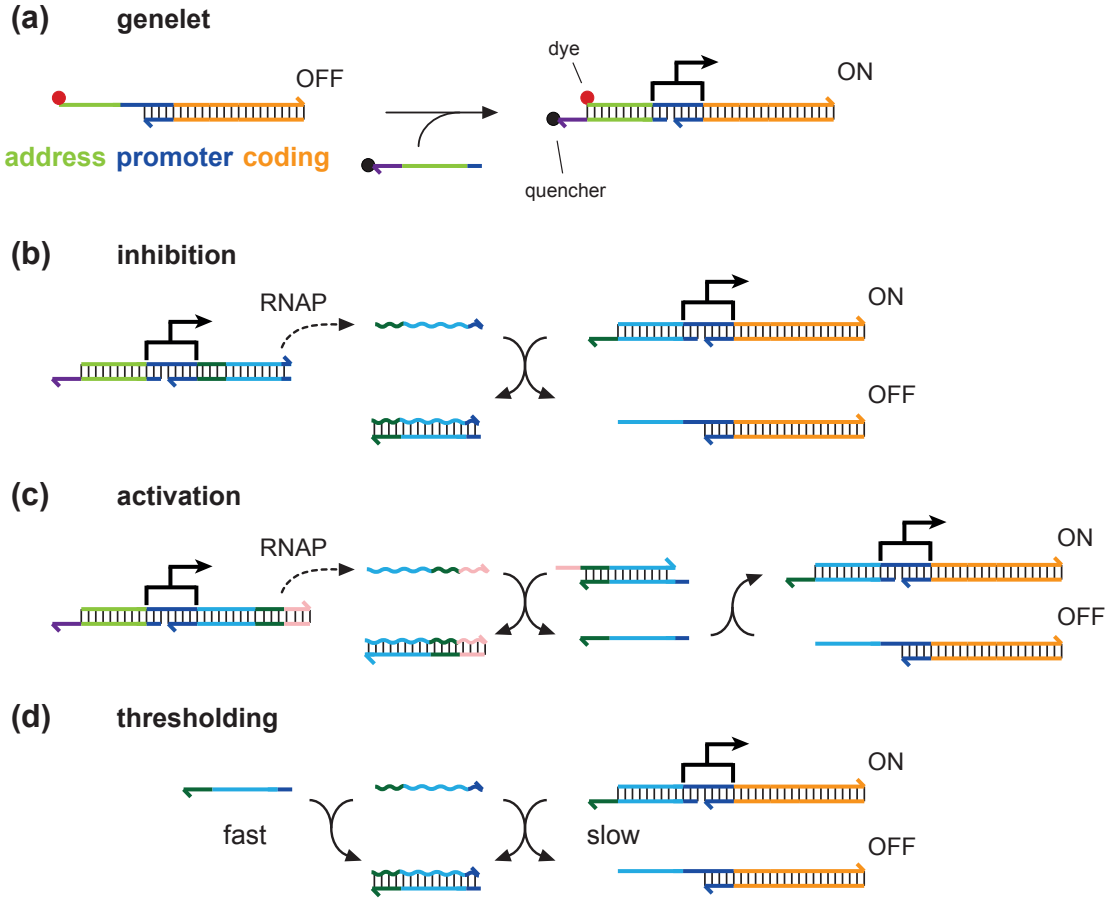

Figure 4.1: Synthetic *in vitro* transcription circuits. DNA strands are represented by straight coloured lines. RNA species are illustrated as wave-like lines. The promoter is indicated by a box with an arrow on top pointing in transcription direction. Enzymatic reactions are indicated by dashed arrow headed lines. Hybridisation and strand displacement reactions are indicated by solid arrow headed lines. (a) A genelet is a short, synthetic gene which contains three domains: a molecular address (green), which serves as recognition site for an input molecule, a promoter (dark blue), and a coding domain (orange). Part of the genelet's promoter and the molecular address are initially single stranded which causes the transcriptional OFF state. Addition of a DNA input strand, which completes the promoter region turns transcription ON. The transcriptional state is read-out by means of fluorescence of a dye-quencher pair (indicated as red and black circle) attached to the $5'-$end of the genelet's non-template strand and the $3'-$end of the activator, respectively. The extended single stranded part on the activator is a toehold (purple) which is required for fast Switching reactions. Genelets can be wired to circuits. (b) By a toehold mediated strand displacement reaction, genelets turn the transcription of downstream genelets OFF when the RNA output of the first genelet is complementary to the second genelet's activator. (c) An additional reaction enables activation. For activation, a DNA activator is initially sequestered by a DNA inhibitor. The RNA output releases the DNA activator upon binding to a toehold region (rose) which in turn completes the promoter region and turns transcription ON. (d) Thresholding enables the creation of non-linear circuits. Excess of DNA activator sequesters synthesised RNA inhibitor and sets a de-activation threshold. The mechanism is based on separated reaction speeds. Thresholding by hybridisation of RNA and DNA is fast while inhibition of the genelet by toehold mediated strand displacement of the activator is comparably slow.

are rather fundamental and have been well studied for years [153, 154, 155, 156]. Hybridisation, strand displacement, RNA polymerisation and degradation reactions can be formulated by means of mathematical models which then can be used to determine the parameters describing the system's dynamics by fits to the experimentally derived data. As a valid model can be used to predict a system's response on changes in parameters, these systems bear the potential to study effects which have not been fully understood so far and are otherwise hard to assess.
The RNA polymerase used in "genelet" reaction circuits is RNAP from bacteriophage T7. T7 RNAP is a 98 kDa single-subunit enzyme that belongs to a family of RNA polymerases from bacteriophages that do not require protein transcription factors [157]. Even though T7 RNAP is relatively specific to its 17 bp consensus promoter sequence, it tolerates considerable modifications in this region [155]. For instance, T7 RNAP also transcribes from DNA templates that contain a nick in the promoter region. If the nick is situated between position -13 and -12 in the promoter sequence on the template strand, transcriptional efficiency is reduced to approximately 50% of that of a fully double-stranded promoter [158]. If part of the promoter is missing, however, transcription is turned OFF.
As shown in Fig. 4.1 (a), a genelet consists of three domains: a molecular address, a promoter, and a coding sequence. Initially, the promoter region is partly single stranded which results in a low RNAP binding affinity and represents the transcriptional OFF state. Upon binding of a DNA strand, which contains the complementary sequence of the molecular address and of the single stranded part of the promoter, transcription is turned ON. The transcriptional state of the genelet is monitored by a dye-quencher pair attached to the $5'-$end of the genelet's non-template strand and the $3'-$end of the short single stranded activating strand, respectively. In the following, a genelet's activating strand will be called activator. The activator is chosen 8 bases longer than the complementary region on the genelet's non-template strand, which leaves a single-stranded toehold region accessible for hybridisation [159, 160]. Genelet circuits are in principle scaleable in terms of individual subunits in a network since both, the molecular address and the coding domain can be chosen arbitrarily. The only part which is pre-determined are the 5 bases which complete the promoter, which do not bind stably in absence of the molecular address domain at the typical operation temperature of 37 °C. By careful sequence design, different genelets can be connected as desired. In Fig. 4.1 (b), an inhibitory link between two different genelets is shown. For this purpose, a genelet's coding sequence is chosen such that its RNA transcript is complementary to the activator of the genelet whose transcription activity is supposed to be inhibited. In the context of genelet circuits, a nucleic acid species that inhibits the transcription of a genelet is called inhibitor. When the inhibitor RNA strand is transcribed from the first genelet, it turns transcription of the downstream genelet OFF due to toehold mediated strand displacement of its activator. The presence of the toehold region is essential for an unidirectional, fast and efficient strand displacement process—a mechanism that has been previously used for the operation of a wide variety of molecular switches based on DNA [161, 162]. Activatory influence of one genelet's transcript on another genelet's transcriptional state cannot be achieved directly by an RNA output molecule since RNA only poorly activates transcription [163]. Instead, an intermediate reaction has to be introduced as illustrated in Fig. 4.1 (c). For activation, an initially present DNA inhibiting strand sequesters the

activator. A toehold on the inhibitor strand serves as binding target for an upstream genelet's RNA output, which forms a dimer with the inhibitor and releases the DNA activator. The activator can now turn transcription of the downstream genelet ON. As introduced in section 2.1.4, in naturally occurring systems non-linear behaviour is achieved by cooperativity, for instance multimerisation of transcription activators, a mechanism which creates step like switching behaviour between transcriptional states. In genelet circuits non-linearity is introduced by thresholding functions as schematically indicated in Fig. 4.1 (d). Free DNA activator in excess sequesters an RNA inhibitor by a fast hybridisation reaction. When free DNA activator is depleted the RNA inhibitor can follow the slower reaction path of toehold mediated transcription inhibition. Hence, activation and deactivation thresholds are set by excess of single stranded activators and inhibitors, a mechanism related to "inhibitor ultrasensitivity", which was used to achieve gene inhibition in synthetic circuits in yeast with Hill exponents of up to 12 [164]. A further requirement to achieve interesting functions is a degradation path for RNA which is necessary to prevent RNA signal accumulation. *E. coli* ribonuclease H (RNase H) seems perfectly suited for this purpose. Instead of randomly degrading all RNA species, the 18.5 kDa endonuclease specifically and processively degrades RNA in RNA:DNA hybrids [156, 165, 166]. However, the particular choice of substrate by RNase H is also a major reason for the limited lifetime of genelet reactions. RNA strands which have lower melting temperatures when forming a hybrid with DNA cannot be degraded and will accumulate with time. As the concentration of these species increases, they start to compete for complementary single stranded circuit components, like toeholds, which changes the kinetics for activation and deactivation, respectively.
Using this toolbox, genelet circuits were designed and demonstrated which showed bistable [3, 62] or oscillatory [4] behaviour. An oscillator circuit will be discussed in detail in the following section 4.2.

4.2 The core oscillator

Based on the genelet concept previously introduced, a negative-feedback oscillator, a negative-feedback oscillator with an additional positive-feedback loop, and a three node ring-oscillator were constructed [4]. In the following, the two-node negative-feedback oscillator, also known as Kim-Winfree-oscillator, will be introduced. For further reading see [4, 63] where the oscillator circuit was already discussed thoroughly. As represented in a simplified scheme in Fig. 4.2 (a), the circuit consists of two genelets, *Sw21* and *Sw12*, which mutually influence the transcriptional state of each other through RNA signals. RNA signal *rA1* produced by *Sw12* activates transcription of *rI2* by *Sw21*. The RNA inhibitor *rI2* in turn de-activates *Sw12*. The reaction dynamics for this circuit were formulated in a simple model in [4] and will be used in section 4.2.1 to understand conceptually why and how this circuit can exhibit oscillations.
Fig. 4.2 (b) shows the oscillator circuit on a more detailed molecular level. The two node negative feedback circuit is established by mutual activatory and inhibitory regulation of the genelets as introduced in Fig. 4.1 (b), and (c). The circuit consists of 7 DNA strands, 2 RNA signals and 2 enzymes. As illustrated in Fig. 4.2 (d),

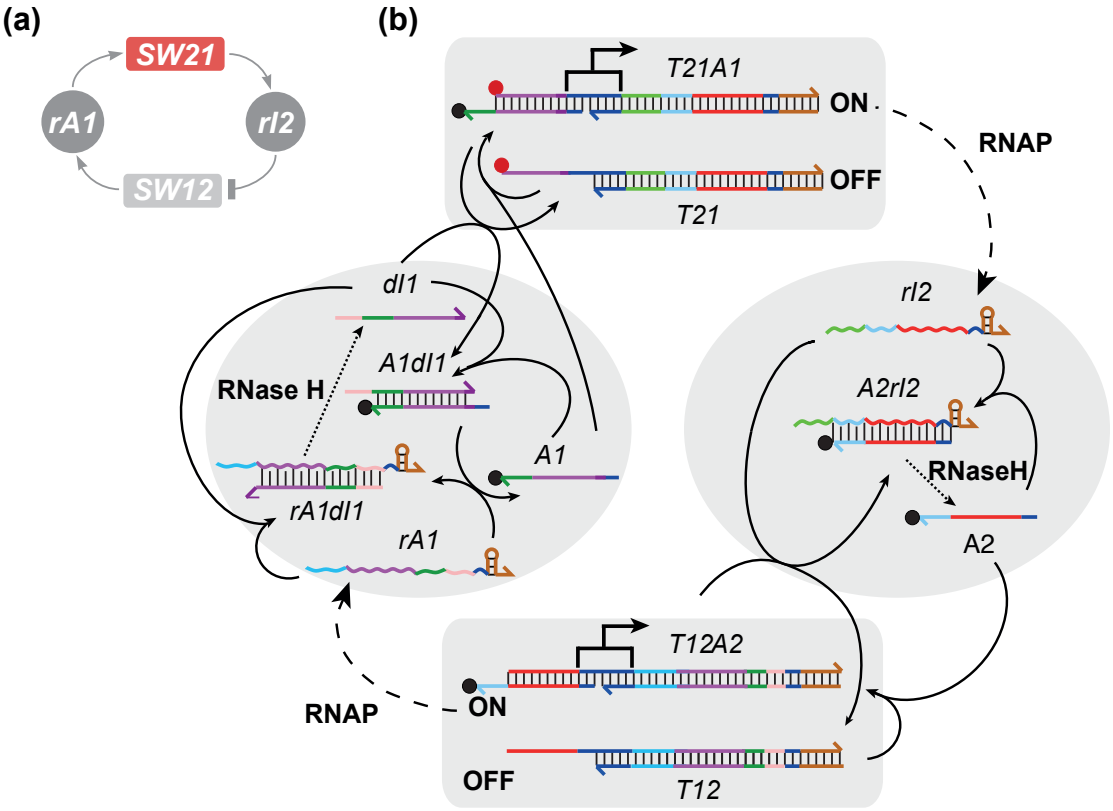

Figure 4.2: A transcriptional negative-feeback oscillator. (a) Abstract representation of the two node negative-feedback oscillator circuit. Genelets are represented by nodes, RNA signals by edges. Activatory interaction is indicated by an arrowhead, inhibitory interaction by a blunt end. (b) More detailed molecular representation of the oscillator circuit. The basic units shown in (a) are indicated by similarly shaped, grey shaded background areas. DNA species are drawn as solid coloured lines. RNA species are illustrated as wave-like lines. Sequence complementarity is indicated by equal colours. Hybridised double strands are connected by faint lines. Promoters are highlighted by a box with an arrow on top directing into transcription direction. Enzymatic reactions are represented by dashed, arrow headed lines. Hybridisation and strand displacement reactions are shown as solid, arrow headed lines. The circuit consists of 7 DNA strands and 2 enzymes. Genelet *Sw21* encodes for an RNA inhibitor which turns genelet *Sw12*'s transcription OFF as it exceeds the threshold set by activator *A2*. *Sw12* encodes for an RNA activator which turns transcription of *Sw21* ON through conversion reaction were DNA activator is released. An activation threshold is set by an excess of DNA inhibitor *dI1*. The systems state is followed by fluorescence of a dye-quencher pair which monitors the transcriptional state of *Sw21*. Figure adapted from [64].

non-linearity is achieved by activation and deactivation thresholds set by *A2* and *dI1*, respectively, and simultaneous degradation of RNA signals which are present in form of hybrids with complementary DNA species.
A typical batch reaction starts in the transcriptional configuration where *Sw21* is OFF since *A1* is sequestered by *dI1*, and *Sw12* is ON. At the very beginning, RNAP transcribes *rA1* from *T12A2*. When [rA1] exceeds [dI1], DNA activator *A1* is released by toehold mediated strand displacement from duplex *A1dI1*. *A1* hybridises with *T21* and turns transcription ON. *Sw21*'s transcript *rI2*, is annihilated by excess *A2*, until free *A2* starts to deplete, and eventually inhibits transcription of *T12* by removing *A2* from *T12A2*. When *Sw12* is deactivated, generation of activator *rA1* gets slower. Thus, degradation of *rA1* bound to *dI1* by RNase H dominates and results in the deactivation of *Sw21*. Finally, *rI2* is not produced at a rate sufficiently high to counteract the steady degradation of RNase H, which resets the system's initial state and thus enables the next reaction cycle to occur for appropriate choice of parameters. The oscillator's dynamics was modelled by an extended model containing 17 ordinary differential equations, 24 molecular species and 24 rate parameters. In addition to the reactions shown in Fig. 4.2 (b), it was further necessary to add reactions to the model which account for incomplete degradation by RNase H as mentioned in section 4.1 and describe slowing down of the frequency with time (cf. for reactions section 6.1).

4.2.1 The simple model

The dynamics of the system can be qualitatively explained by a simple model of 4 ordinary differential equations, which describe the reactions shown in Fig. 4.2 (a). In this 4-dimensional model only the time dependent concentrations of signalling molecules [rA1] and [rI2], and the concentrations of genelets in the ON-state [T12A2] as well as [T21A1] are considered. In the non-dimensional form they are given by

$$\begin{aligned} x &= \frac{[\mathrm{rA1}]}{K_A} & u &= \frac{[\mathrm{T12A2}]}{[\mathrm{T12^{tot}}]} \\ y &= \frac{[\mathrm{rI2}]}{K_I} & v &= \frac{[\mathrm{T21A1}]}{[\mathrm{T21^{tot}}]} \end{aligned} \tag{4.1}$$

where K_I and K_A are the thresholds set by DNA activators *A2* and *dI1*, respectively, and [Tij$^{\mathrm{tot}}$] is the sum of all molecular species containing *Tij*. The set of differential equations which describe the 4-dimensional model is then

$$\begin{aligned} \gamma \frac{dx}{ds} &= \alpha \cdot u - x & \frac{du}{ds} &= \left(\frac{1}{1+y^n} \right) - u \\ \gamma \frac{dy}{ds} &= \beta \cdot v - y & \frac{dv}{ds} &= \left(1 - \frac{1}{1+x^m} \right) - v \end{aligned} \tag{4.2}$$

with Hill exponents n and m for the genelets' response to RNA inputs and the rescaled non-dimensional constants and parameters

$$\begin{aligned} s &= \frac{t}{\tau} & \alpha &= \frac{1}{K_A}\frac{k_p}{k_d}[\mathrm{T12^{tot}}] \\ \gamma &= \frac{1}{k_d \cdot \tau} & \beta &= \frac{1}{K_I}\frac{k_p}{k_d}[\mathrm{T21^{tot}}] \end{aligned} \tag{4.3}$$

where t is time, τ is a relaxation time for hybridisation reactions, k_p is the polymerisation rate which is assumed to be equal for both genelets, and k_d is the decay rate of RNA signals which is assumed to be independent of the RNA species. One immediately notices that the realisation of genelets with identical transcription activities and an universal degradation of RNA species may be hard to achieve experimentally. And if so this would limit the design space considerably which in turn has strong impact on the scalability of genelet circuits and the opportunity to study general problems. However, 4.2 can be used for qualitative understanding of the dynamics underlying circuit. For this purpose, we simulated the simple model in Matlab using its built-in ODE solver "ode23s". Fig. 4.3 shows plotted simulation results for the simple model for transcription rates $k_p = [0.04, 0.03, 0.025]\,{}^1\!/_{\mathrm{s}}$, which can be loosely seen as representative for the RNAP concentration. Hence, variation of k_p qualitatively emulates variation in RNAP concentration. In Fig. 4.3 (a) and (b) the time development of concentrations of the non-dimensional analogs of genelets, v and u, are compared with their corresponding effector RNA signals. Initial parameters were chosen *T21* OFF and *T12* ON. If $x > 1$ the activator exceeds its threshold and v is activated, thus increases. Decreasing transcription rate results in damped oscillations. The inhibitor concentration *rI2* cannot exceed its threshold sufficiently. For $k_p = 0.025\,{}^1\!/_{\mathrm{s}}$ y does not exceed the threshold. Consequently, genelet *T12* or u, respectively, is not sufficiently deactivated and the reaction dampens out quickly. The dependence of sustained oscillations on the total amount of enzyme as well as on the ratio between RNAP and RNase H concentration qualitatively reproduces the behaviour observed in the experiment (and also predicted by the extended model) well. In Fig. 4.3 (c) the time development of u and v is compared representing the fractions of active *T12* and *T21*. As the model simulations reveal, the phase shift between the oscillations of both genelets' activities is smaller than one would expect naively. This is an important feature which has to be considered when the oscillator circuit is used for the timing of molecular processes as will be described in section 4.3.

4.2.2 Oscillator tuning and experimental results

The oscillator circuit is relatively sensitive to parameter changes. Therefore, special attention has to be paid during preparation. As DNA strands are ordered single stranded, genelets have to be hybridised in advance of oscillator experiments. The genelets' coding sequences were designed to terminate with a secondary structure hairpin-like loop intended to facilitate the falling-off of the RNAP after transcription and to reduce non-specific extension of the transcribed RNA strand [167]. To ensure that DNA single strands are hybridised into completely double stranded templates, strands are carefully annealed by temperature ramps on an Eppendorf Thermocycler.

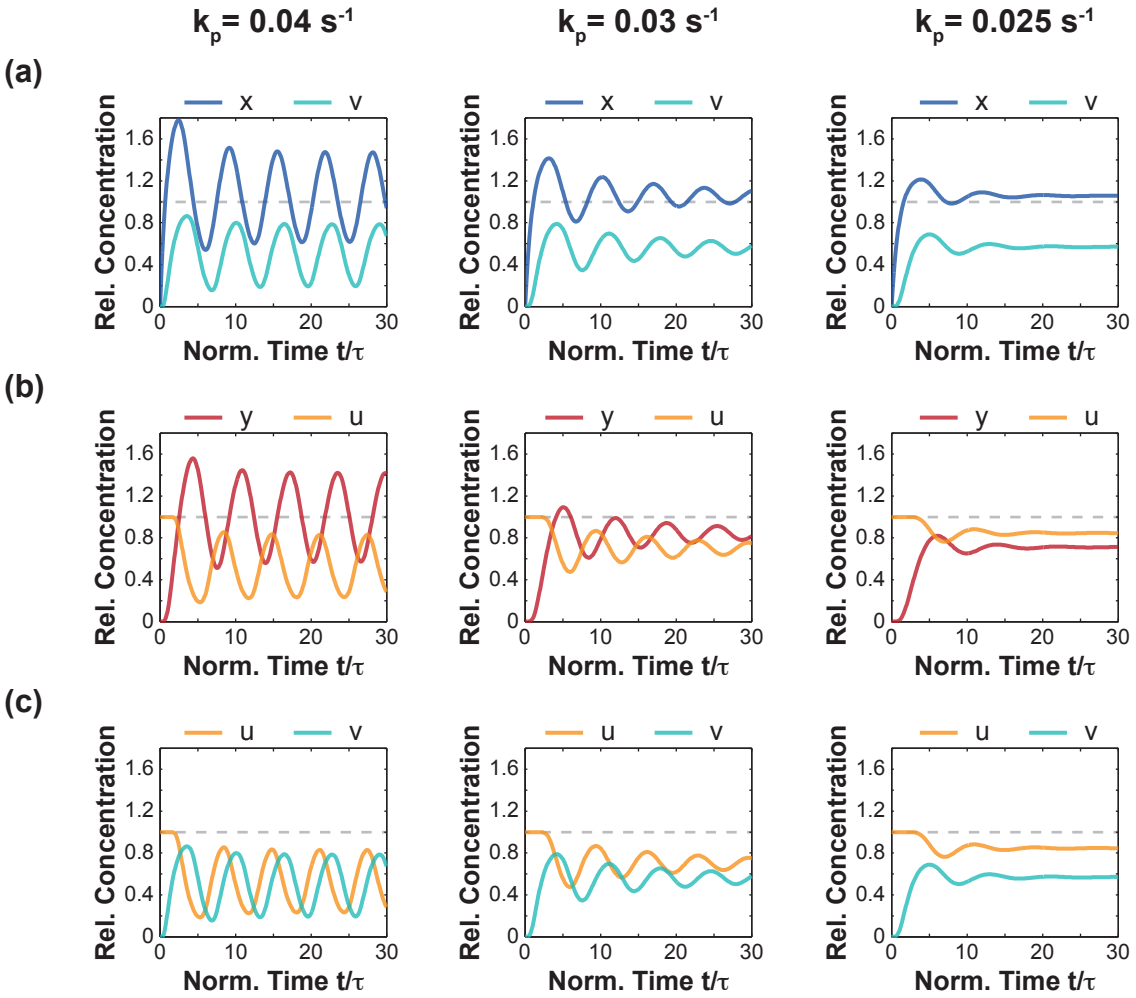

Figure 4.3: Simulation of oscillator dynamics: The simple model. Trajectories from simulations of the simple model for transcription rates $k_p = 0.04\,\mathrm{s}^{-1}$ (left), $k_p = 0.03\,{}^1/_\mathrm{s}$ (centre), and $k_p = 0.025\,{}^1/_\mathrm{s}$ (right). In the simple model, the transcription rate can be associated with the RNAP concentration. Model parameters are $K_A = 1\,\mu\mathrm{M}$, $K_I = 1\,\mu\mathrm{M}$, $k_d = 0.002\,{}^1/_\mathrm{s}$, $[\mathrm{T12}^{\mathrm{tot}}] = 100\,\mathrm{nM}$, $[\mathrm{T21}^{\mathrm{tot}}] = 100\,\mathrm{nM}$, $\tau = 500\,\mathrm{s}$, and $m = n = 5$. The relative concentration (Rel. concentration) $= 1$ is indicated by a grey dashed line. (a) v represents the transcriptionally active fraction of genelet *T21*. x is the ratio of activator *rA1* and its threshold *dI1*. When *rA1* exceeds the threshold concentration ($x > 1$) *T21* is activated thus v. (b) u is the transcriptionally active fraction of genelet *T12*. y is the ratio of inhibitor *rI2* and its threshold *dA2*. When *rI2* exceeds the threshold concentration ($x > 1$) *T12* is deactivated thus u decreases and limit cycle oscillations occur for $k_p = 0.04\,{}^1/_\mathrm{s}$. If the over- and undershoot is not sufficiently strong, damped oscillations occur ($k_p = 0.03\,{}^1/_\mathrm{s}$). For $k_p = 0.025\,{}^1/_\mathrm{s}$ *rI2* does not exceed the threshold thus the circuit falls into a stable steady state and dampens out quickly. (c) Time development of u and v.

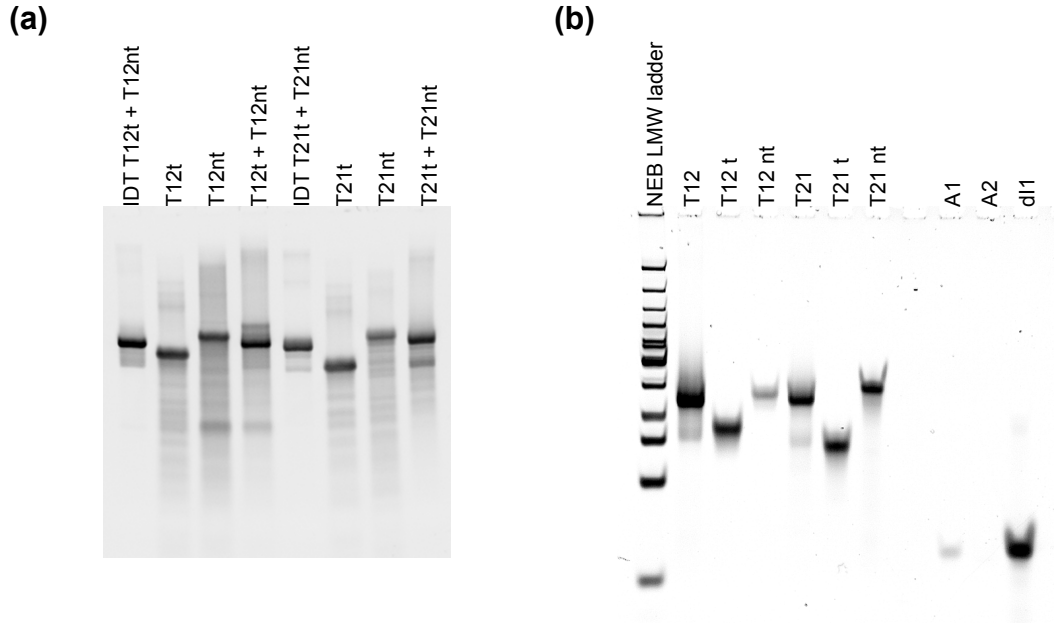

Figure 4.4: 10% polyacrylamide control gels of genelet annealing stained by incubation with 1x SybrGold™ for 30 min. (a) Comparison of PAGE purified strands purchased from IDT at Caltech and HPLC purified strands from biomers.net. The lanes of genelets comprised of strands from IDT don't show any smear for shorter lengths. Genelets "IDT T12t + T12nt" and "IDT T21t + T21nt" were used in oscillator reactions $\#13 - \#20$ shown in Fig. 4.6. Consequently, biomers.net provided newly synthesised template strands purified with PAGE purification particularly developed for dye labeled DNA strands for free. (b) PAGE with DNA oscillator components. Template strands *T12t*, *T12nt*, and *T21t* were purified with standard PAGE purification. *T21nt* was purified with PAGE purification method optimised for fluorescently labeled oligonucleotides. All remaining reactions (except at Riverside $\#21 - \#24$) contained these strands. *T*12 and *T21* are the annealed genelets. Strands *A1* and *A2* were ordered from IDT since Iowa black™ quenchers were only provided by them. Due to the high quenching efficiency *A1* shows only a very faint band and *A2* is not visible in the SybrGold™ fluorescence channel. This observation has been made several times and was also reported in the original publication [4].

The annealing script was written to heat the samples to 95 °C. After 5 min dwell time at 95 °C the samples are cooled down to 20 °C during an 80 min temperature ramp. The result of the genelet annealing is then checked in a polyacrylamide gel as shown in Fig. 4.4. Further information on circuit preparation is given in section 4.6.1.

In the following, a set of oscillator traces is presented to give impression about the different dynamics the oscillator circuit can exhibit. However, the presented traces were not result of an exploration of the parameter space in terms of component concentrations. Our intention was to optimise the oscillator circuit for studies in microemulsions which will be discussed later in section 4.4. Thus, screening of parameters was directed towards sustained oscillations with large amplitudes and low damping. Using the number label of a reaction (#) the tuning for the particular case can be found in the appendix section 6.3. #-labels are not assigned in chronological order regarding date of the experiment. Oscillator traces shown in Figs. 4.5, 4.6, 4.7, 4.8, 4.9, and 4.10 are shown normalised to their maximum fluorescence, typically recorded by a baseline in the beginning, which represents the fraction of *T21* OFF (cf. section 4.6.5). The conversion to concentrations was omitted for the purpose of better comparison of reaction batches with different concentrations of *T21*. The sim-

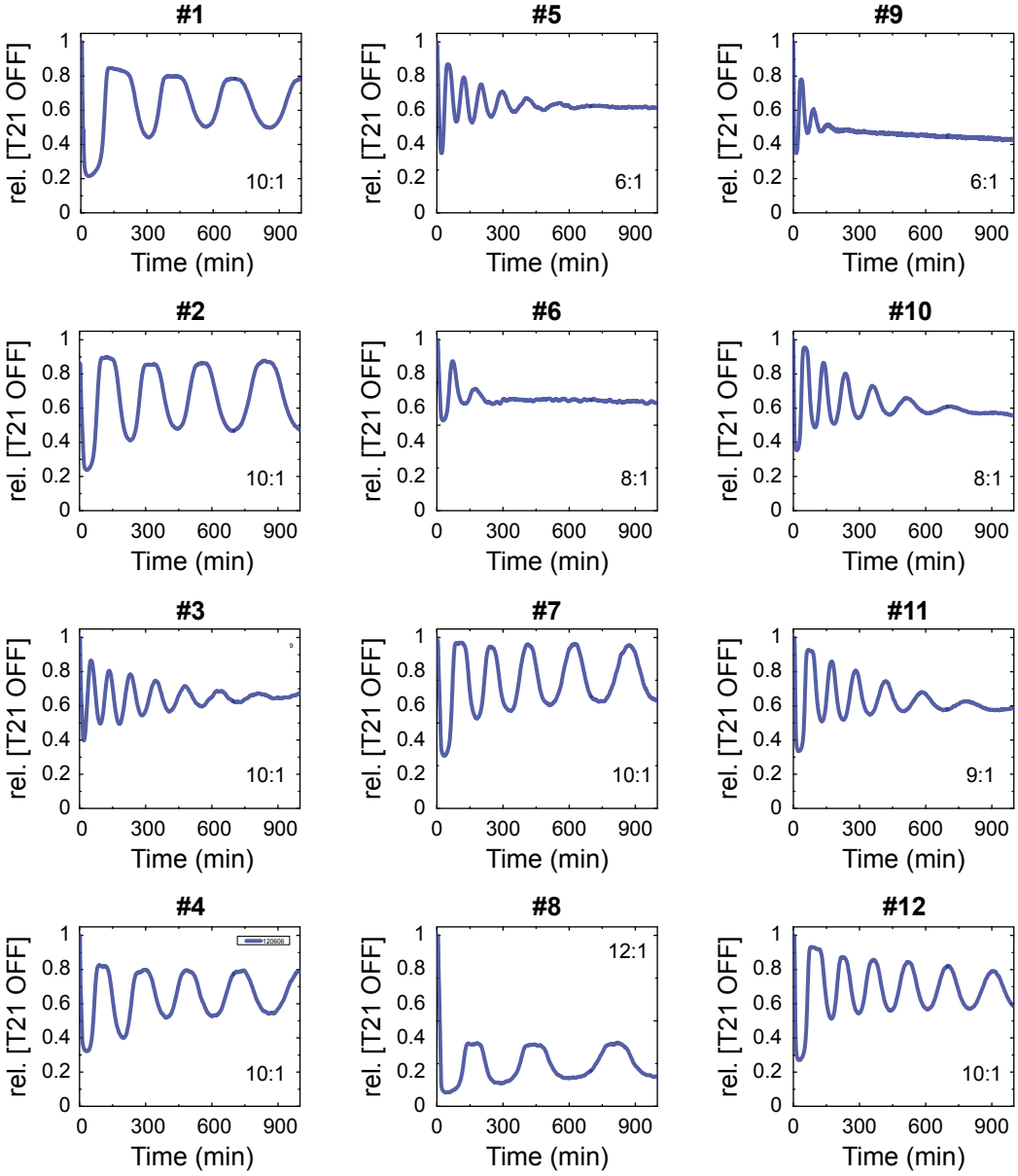

Figure 4.5: Typical oscillator time traces. Reaction #5−#8 RNase H screening for different enzyme ratios while [RNAP] is kept constant. Reaction #9 − #12 RNase H screening for different enzyme ratios with different RNAP batches. At this operating point, decreasing RNase H concentration lowers damping. "rel. [T21OFF]" is the relative concentration of template *T21* in the transcriptional inactive state.

ple model simulation (Fig. 4.3) indicated the influence of varying the ratio between transcription and degradation activity. We therefore tuned the oscillator if necessary predominantly by adjusting enzyme ratios and concentrations or "units". Commercially available enzymes are usually characterised by *unit* concentrations defining a turnover number for a standardised substrate rather than by molar concentrations. The molar concentrations can be derived by means of the "specific activity". The specific activity bears information about the amount of active enzyme and can show strong variations from batch to batch. Hence, it appears more reliable to adjust enzyme concentrations according to their unit concentration. Since the unit definition is obviously different for different enzymes, we used in practice "volume" and "volume ratios". Throughout all data sets a ratio in enzyme volume between RNAP and RNase H of 10 : 1 showed the best results. Results from enzyme ratio adjustment experiments are shown in Fig. 4.5. The volume ratios are printed into the corresponding panels. Molar and unit concentrations for all traces shown are listed in the appendix 6.3. Oscillators $\#5 - \#12$ are results from RNase H screening experiments. As indicated by simple model simulations decreasing the ratio between RNAP and RNase H turns sustained oscillations into damped oscillations. #9, #10, and #12 were studied in microemulsions (section 4.4). Reactions $\#13 - \#16$ in Fig. 4.6 represent results from first attempts of oscillator experiments performed at Caltech in 2011. Reactions $\#17 - \#20$ are traces recorded at TUM with DNA strands brought from Caltech and enzymes purchased in Germany which showed the fastest oscillations we have observed so far. Reactions $\#21 - \#24$ are traces recorded at UC Riverside in 2013. Oscillations typically did not exceed approximately 600 min. Since the vendor where strands were originally ordered stopped to synthesise oligonucleotides with Texas Red label as in genelet *T21*, different dye labels were used. Reaction batch for #24 has identical ingredients as #23 but *T21* was labeled with TYE665 fluorophore. Nonetheless, as mentioned initially, we aimed for optimising the oscillatory behaviour for studies in microcompartments. We therefore started with adjusting the baseline around which the concentration *T21* oscillated in order to maintain sufficiently strong fluorescence for fluorescence microscopy experiments. At this point it has to be noted that although the extended model captures the oscillator dynamics well, predictions can only be made using reliable parameters. Since the parameters are derived by model fits to experimental traces, the extended model could not be used to find the desired behaviour. Due to its non-linearity, it is often not easily conceivable how the circuit will respond to changes in the parameter. However, in terms of the baseline it turned out that often the following considerations are valid. The read-out measure is the concentration of *T21* OFF. A low baseline hence indicates that most of *T21* is ON, and *vice versa*. *T12* ON produces the RNA activator *rA1* which activates (indirectly) *T21* when its concentration exceeds the threshold represented by *dI1*. Hence, increasing the concentration of *dI1* increases the mean fraction of *T21* OFF since *rA1* can not overcome the threshold. Reactions $\#26 - \#32$ nicely show this trend (Fig. 4.7). In Fig. 4.8, the oscillator tuning is shown until oscillatory behaviour was achieved as desired. In reactions $\#37 - \#40$ activator *A2* was screened. The conditions of #37 were then adapted for further adjustment of the enzyme ratio in $\#41 - \#42$. Next, *dI1* was adjusted for enzyme ratio of #42 while enzyme concentrations were increased in $\#44 - \#47$. Slight reduction of enzyme concentrations eventually yielded the oscillator reaction shown in #43. In Fig. 4.9, typical traces

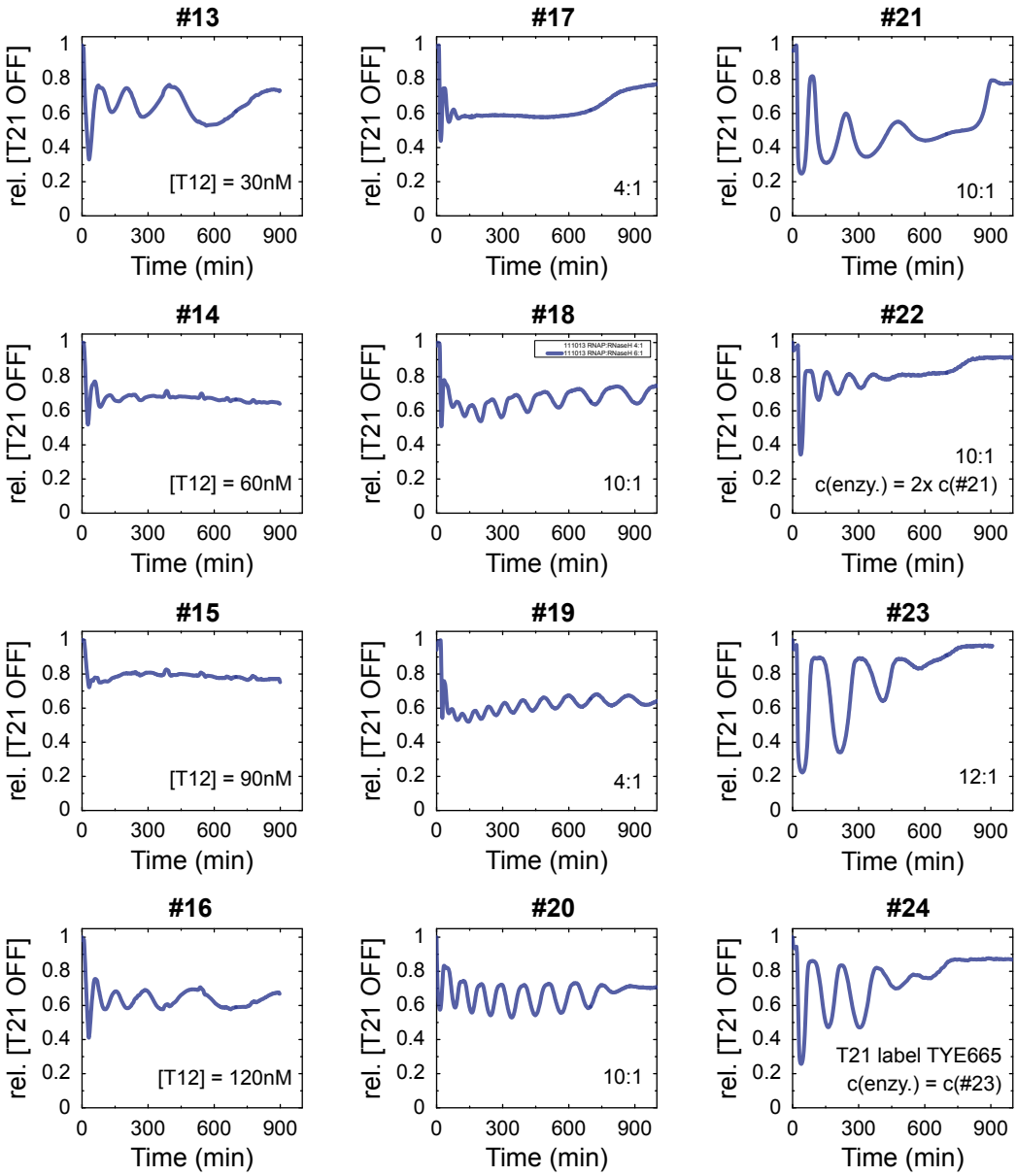

Figure 4.6: First attempts and typical oscillator time traces recorded at Caltech and UCR. Reaction #13 – #16 were the very first attempts to get the oscillator to work. Data was taken at Caltech. Reaction #17 – #20: Oscillator traces of first experiments in Munich using DNA strands from Caltech and enzymes purchased in Germany. Reaction #19 was the fastest oscillator which has been realised so far showing 9 oscillation cycles in 900 min. Reaction #21 – #24: Oscillator traces taken at UCR. Genelet *T21* is labeled with TYE665 instead of TexasRed. Mysteriously, the runtime in Riverside was limited to 12 – 16 h. "rel. [T21OFF]" is the relative concentration of template *T21* in the transcriptional inactive state. Sample compositions are listed in table 6.2.

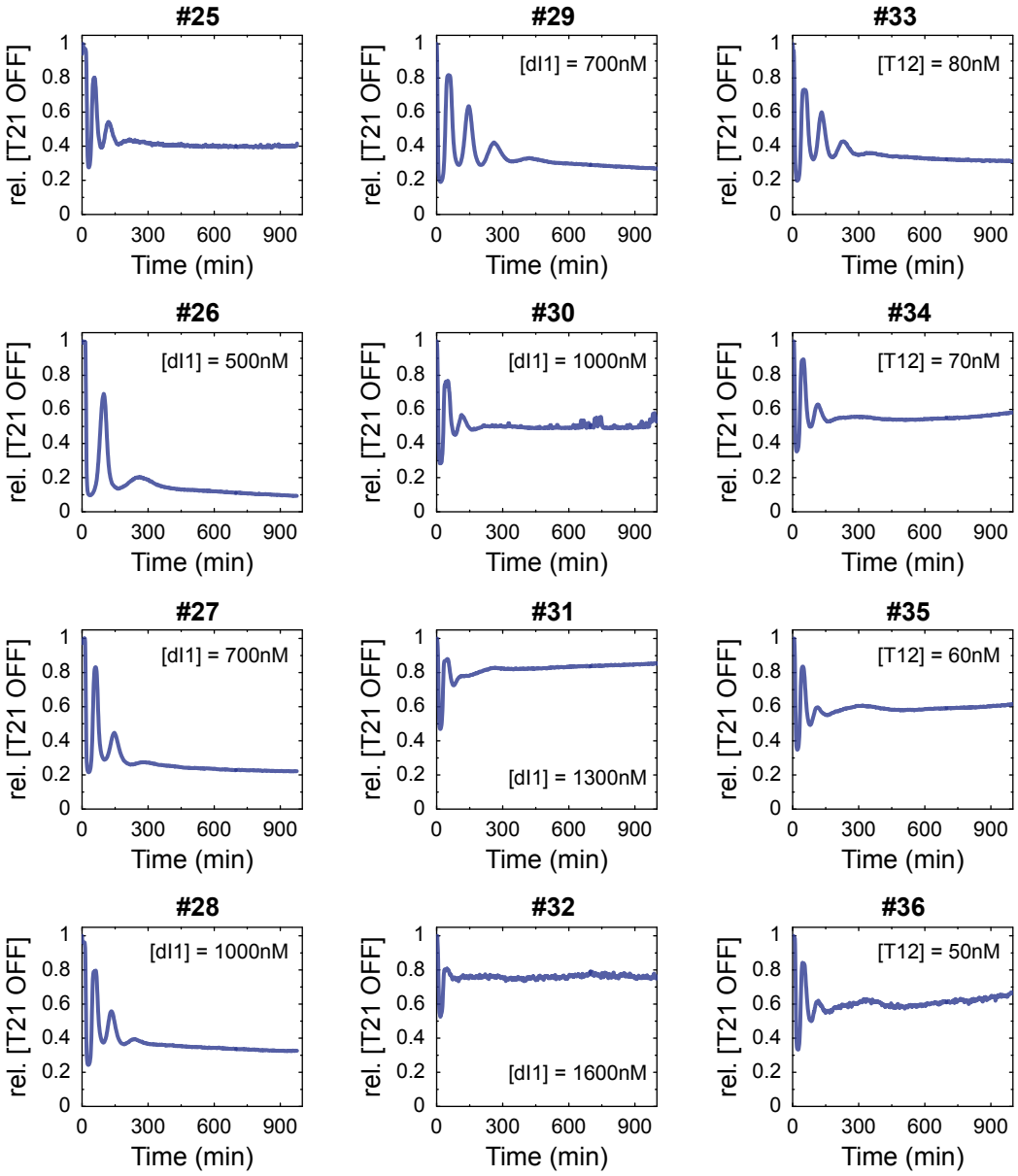

Figure 4.7: Oscillator time traces for variation in [dI1] and [T12] in order to adjust the baseline around rel. [T21OFF] ≈ 0.5. Reaction #26 – #28: Increasing [dI1] raises the baseline which represents the mean amount of $T21\ OFF$. This seems plausible since more inhibitor is available to turn *T21* OFF. Reaction #25: Same composition as reaction #27 but with lower [T12]. Decreasing [T12] influences the dynamics similar as increasing [dI1]. In both cases, in the mean, less *T21* is activated. Reaction #29 – #32: Further variation in $dI1$ with [T12] being the same as in reaction #25 increases the baseline but pushes the operating point out of the oscillatory regime. Reaction #33 – #36: Further decrease of [T12] leads to a higher baseline but does not result in a more sustained oscillating behaviour. "rel. [T21OFF]" is the relative concentration of template *T21* in the transcriptional inactive state. Sample compositions are listed in table 6.2.

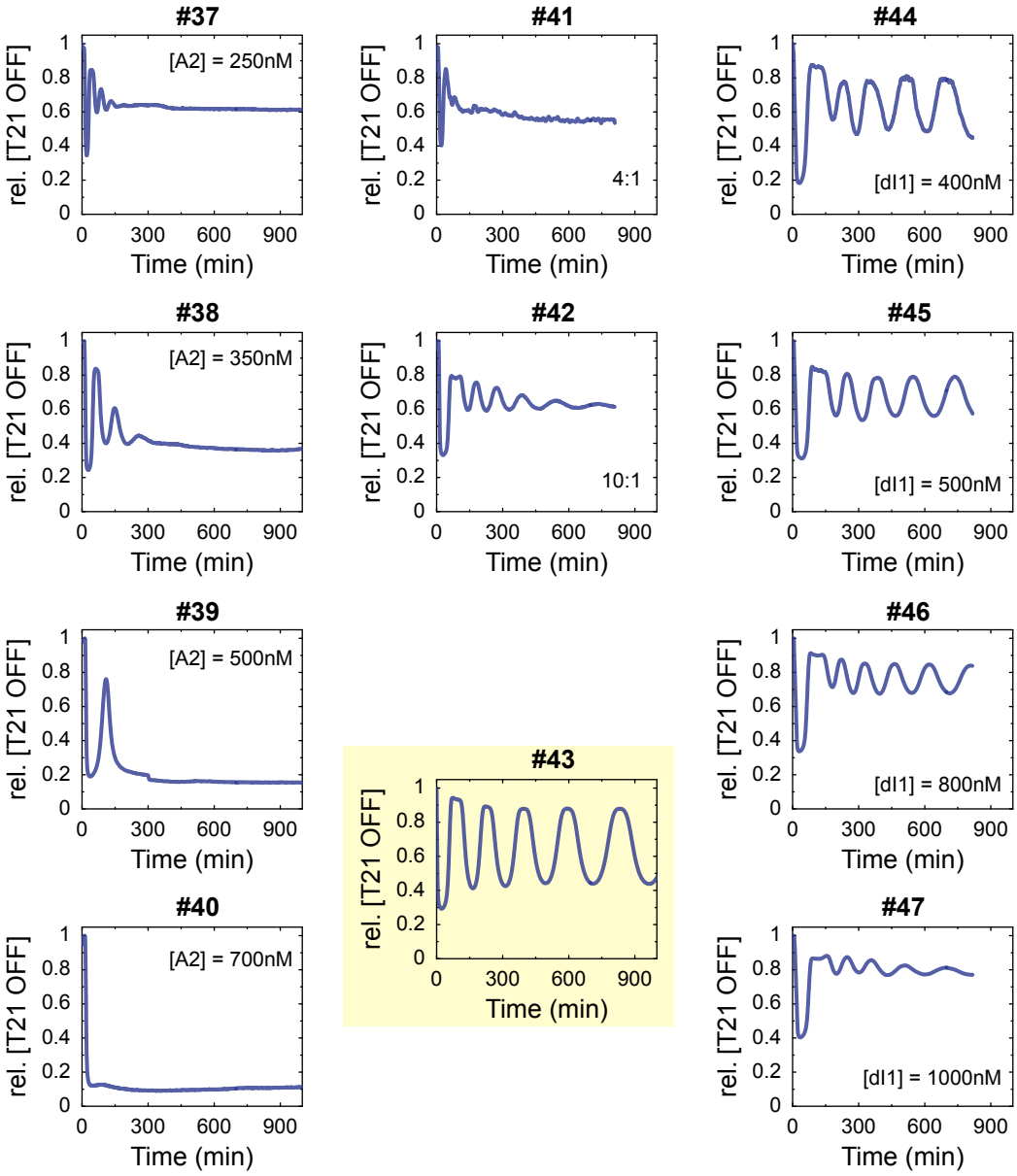

Figure 4.8: Reaction #37 – #40 oscillator time traces for varying [A2]. Reaction #41 is the same composition as reaction #37. Reaction #42 has increased amount of RNAP in respect to reaction #37. Reaction #44 – #47 *dI*1 screening for [A2] as in reaction #39 and enzyme concentration as in reaction #42. Further increase of enzyme concentration turned reaction #45 into #43 which showed sustained oscillations over 25 h. "rel. [T21OFF]" is the relative concentration of template *T21* in the transcriptional inactive state. Sample compositions are listed in table 6.2 and 6.3.

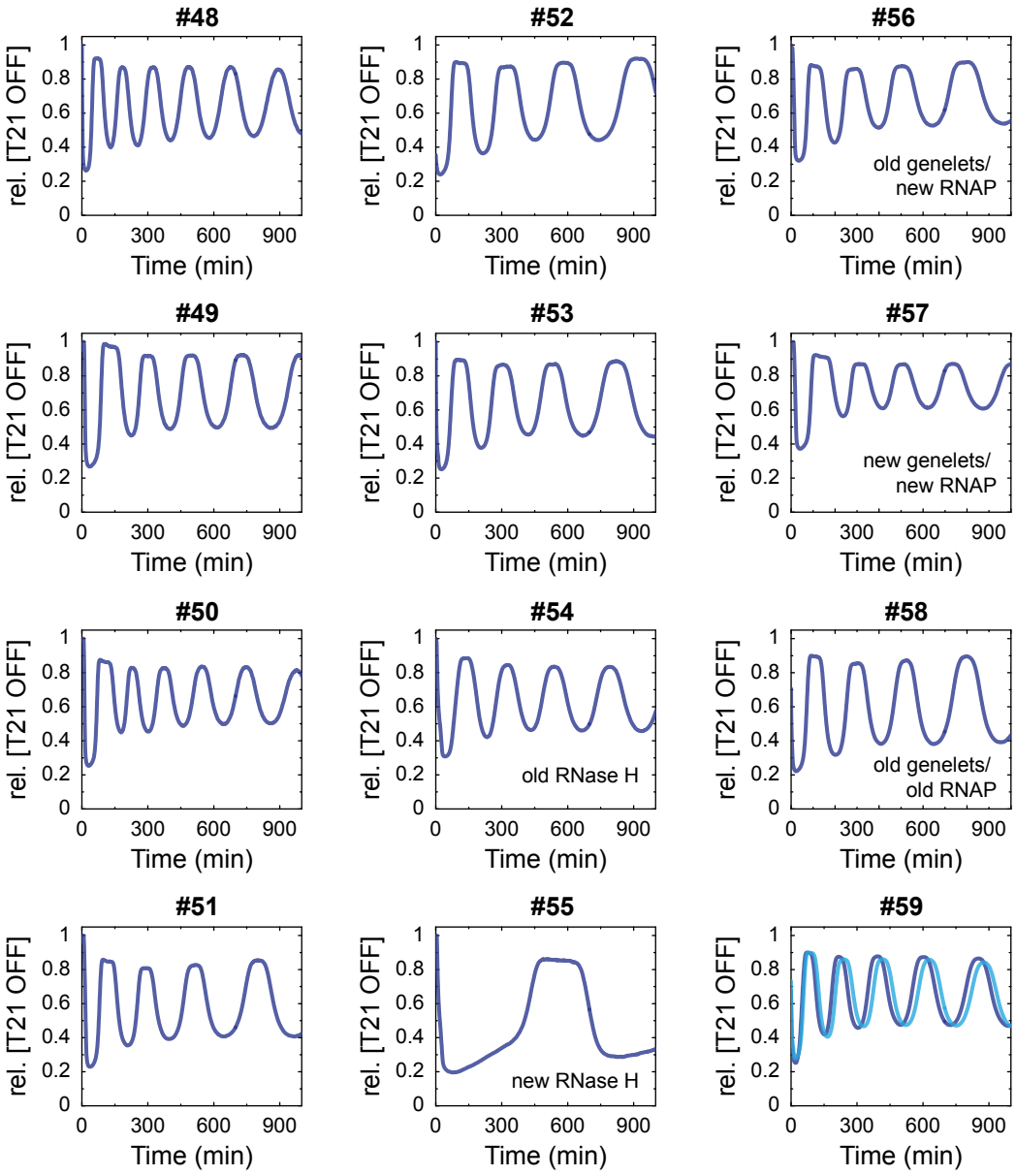

Figure 4.9: Typical oscillator time traces after the optimal operation point was found. Reaction #48 – #50: oscillator time traces for same composition. Differences in the dynamics are very likely due to ageing of enzymes, different reaction batch volumes, and pipetting error. Reaction #51 – #53 is the same as reactions #48 – #50 but using a new RNAP batch. Reactions #54 and #55: Oscillator time traces with old and new RNase H batch. The slower oscillation of reaction #55 is an indicator for lower RNase H activity than the previous batch (#54). This is a very good example for the range of variability in the activity of enzyme batches which in principle should be equal expressed by their equal unit content. Reactions #56 – #58: Comparison between old (#56) and newly annealed (#57) genelets, and old genelets with old RNAP (#58). Reactions #59: Test of reproducibility. Both samples were individually assembled to verify the reproducibility of the oscillator's dynamics. "rel. [T21OFF]" is the relative concentration of template *T21* in the transcriptional inactive state. Sample compositions are listed in table 6.3.

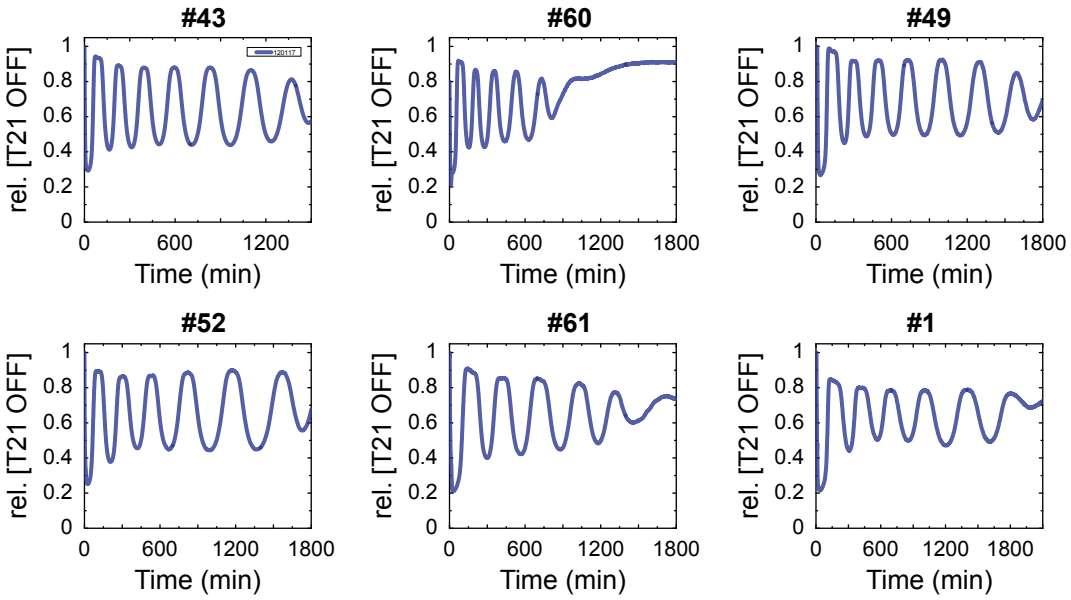

Figure 4.10: Special oscillator time traces. The longest oscillator so far observed: Reaction #1 showed oscillations for almost 35 h. Sample compositions are listed in table 6.2 and 6.3.

are shown for experiments performed after the operating point of #43 was found. As mentioned, an issue occurring for genelet circuits is the difference in activity between enzyme batches. In principle, the unit definition by the vendor is supposed to ensure constant activity between different enzyme batches. While for RNAP this is most of the time the case (cf. 4.9 reactions #48 − #50 and #51 − #53), RNase H can show dramatic variation from one batch to another. For instance, as shown in Fig. 4.9, a several months old RNase H batch (#54) appears to have higher activity than a just recently received one (#55). In Fig. 4.10 a collection of traces with remarkably long life-times is shown. It should be noted that the run-time of the oscillator shows strong variations. Oscillation circuits presented in the original publication did not exceed a 25 h run-time. From the time on when stable reactions were achieved at our lab at TUM, stable oscillations always persisted for more than 20 h and in several cases up to 30 h as shown in Fig. 4.10. We have been looking for reasons by extensive parameter screenings of all components, especially during a research visit at UC Riverside (California, USA) when the oscillator circuit barely reached a life-time of 10 h. In conclusion it can be said that the reactions do not die out due to depletion of NTPs. The influence of different activities for different batches of pyrophosphatase is not completely clear. The only components that were bought at the particular location of the experiments were RNAP, and RNase H. However, the same products were ordered from the same vendors. Thus, a plausible explanation is still missing. As *in vitro* circuits have the purpose to serve as simple model systems with predictable behaviour the cause for this difference appears interesting since reproducibility should not be a matter of the location where an experiment is performed. A global set of parameters were determined for reactions #9, #10, and #12 by a fit of the extended model to the experimental traces shown in Fig. 4.11 (cf. for reactions section 6.1). The range for fit parameters was chosen within a realistic range which is given with the resulting parameters in appendix table 6.1. Amplitude, frequency and damping

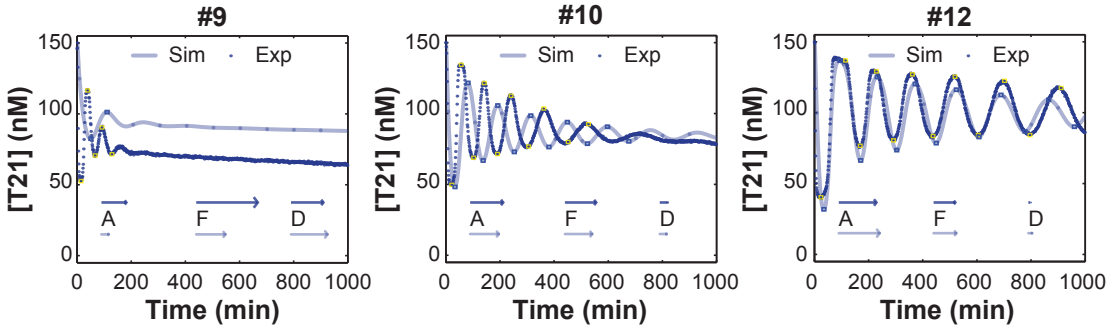

Figure 4.11: Fit of the extended model (solid, light blue line) to experimental oscillator traces #9, #10, and #12 (dark blue dots) which were studied in microemulsion droplets in section 4.4. The magnitudes of amplitude (A), frequency (F), and damping coefficient (D) were rated as originally in [3]. The length of the arrows indicate the magnitude. Same length of arrows indicate good agreement of simulation (Sim) and experiment (Exp). Motivated by the magnitude of the damping coefficient D, the oscillator tunings #9, #10, and #12 were coined "strongly damped", "damped", and "sustained", respectively, in the context of microemulsion studies. A global set of rate parameters was derived for all three tunings (listed in appendix table 6.1). Reactions considered in the extended model are given in section 6.1. Fitting was performed by Jongmin Kim in collaboration for [64].

coefficient are in good agreement for simulation and experiment. The damping coefficients determined by the fitting algorithm motivated us to coin the oscillator tunings #9, #10, and #12, "strongly damped", "damped", and "sustained", respectively, in the context of microemulsion studies in section 4.4. In the following, the dynamics of the transcriptional oscillator are used to control downstream molecular processes.

4.3 Controlling molecular tasks by a biochemical oscillator

A goal in molecular programming is to build up reaction networks with arbitrary scale. Rather recently, an impressive example was demonstrated for enzyme free strand displacement systems - a digital logic circuit comprised of 130 DNA strands which was capable of computing the square root of a 4-bit binary number [168]. While the genelet circuits, in principle, can be wired into arbitrarily complex networks [152] the largest circuits which have been realised so far consisted of only three different genelets. A chemical pH oscillator was previously used to switch the conformation of a sensitive DNA molecule [169]. In order to demonstrate both scalability as well as the capability of the oscillator to act as driving horse for multiple molecular tasks, we adopted a recently published concept of so-called insulator genelets. Franco *et al.* applied the transcriptional oscillator introduced in section 4.2 to control the opening and closing of DNA tweezers as well as the clocked synthesis of a functional RNA aptamer [161, 63]. Portions of oscillator components such as activators and inhibitors were used to change the conformational state of the DNA nanodevices. However, it turned out that the coupling of these devices set a load to the oscillator which had perturbing impact on the dynamics of the core circuit and therefore strongly limited the amount of downstream tasks regulated by the oscillator. As a work around, genelets where pasted between oscillator core circuit and the action of interest whose tran-

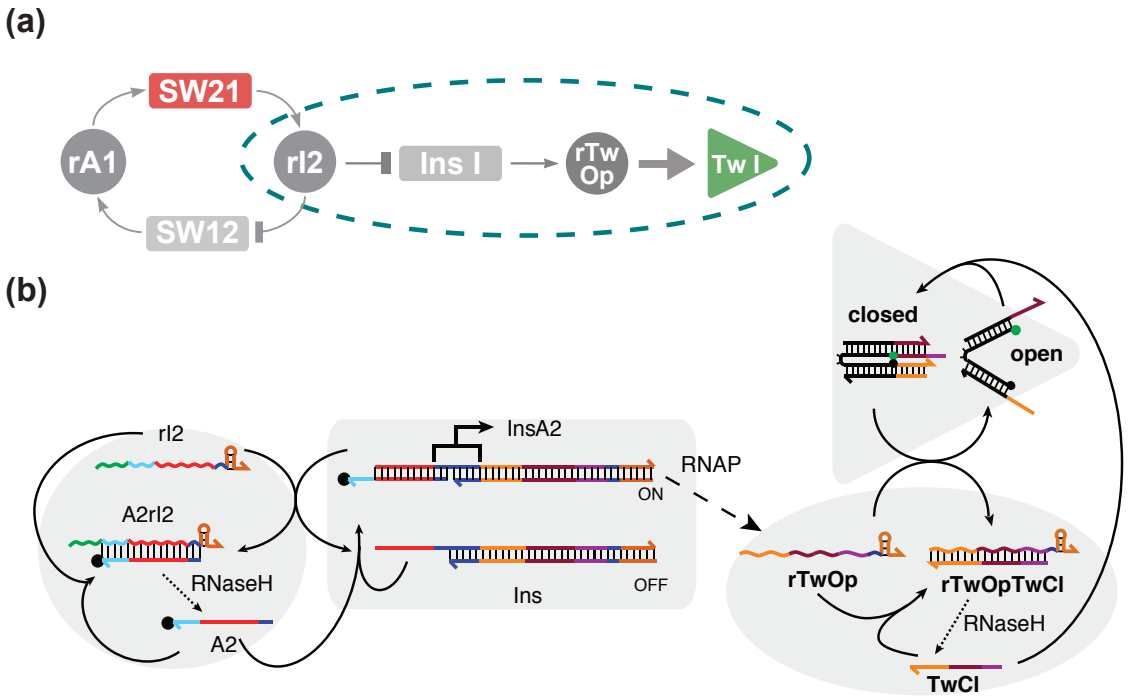

Figure 4.12: Coupling scheme of tweezers set I. (a) Simple scheme of the coupling concept adopted/ reproduced from [63]. Transcription activity of *InsI* is regulated in the same manner as genelet *Sw12* is turned OFF by RNA inhibitor *rI2* and turned ON by DNA activator *A2*. The coding sequence *InsI* encodes for an RNA effector strand that causes tweezers I to open. (b) More detailed scheme of the molecular interactions for tweezers set I coupling. Grey shapes in the background indicate the task in the simplified scheme. The activities of *InsI* and *Sw12* are synchronously modulated by *rI2*. When *InsI* is activated by *A2*, RNA effector *rTwOp* is produced which opens *TwI* by toehold mediated strand displacement of DNA closing strand *TwCl*. RNase H degrades *rTwOp* from the hybrid *rTwOp:TwCl* which releases *TwCl* and resets the initially closed state of *TwI*. The conformational state of *TwI* is monitored by the fluorescence of a dye-quencher (green/ black circles) pair attached to both ends of the base strand. Consequently, fluorescence is low in the closed state and high in the opened.

scriptional activity was switched by the oscillator's activators. Transcription of those genelets amplified the actual activator concentration such that only small amounts of insulating genelets - called insulator - were necessary to drive a considerably larger amount of tweezers or to control the production of an RNA aptamer in a timed fashion with only marginal influence on the core oscillator's dynamics. The tweezers consist of three DNA strands which are two arms with single stranded binding sites for a closing strand and a strand which connects the two arms (Fig. 4.12 (b) and 4.13 (b)). The conformational state is read-out by fluorescence of a dye-quencher pair which is attached at 5'- and 3'-end of the connecting strand - in the following called the base strand. Hence, fluorescence is high in the opened state when the distance between dye and quencher is large and drops when the tweezers are closed. The coupling concept used by Franco *et al.* is shown in Fig. 4.12. The insulator genelet *InsI* is turned transcriptionally active by activator *A2*, which also activates transcription of *T12*. Hence, transcriptional activities of both genelets are in synchrony. As *InsI* is transcriptionally active, an RNA strand is produced which removes a DNA closing strand by toehold mediated strand displacement and renders the tweezers into an opened conformation. The resulting RNA:DNA hybrid *rTwOp:TwCl* serves as substrate for RNase H. Consequently, the DNA closing strand *TwCl* is released due to degradation of its RNA complement, and resets the initially closed state of the tweezers. Inspired by this concept we designed insulator genelets for various molecular tasks and coupled them to the oscillator. First, we used the same insulator-tweezers system as published in [63] which was coupled to the oscillator by activator *A2* (Fig. 4.12). Additionally we designed an alternative set of tweezers (called *TwII*) which was also coupled to *A2*, but in an inverse manner, meaning that the action was counter periodic to the first set. The operation scheme for tweezers set 2 is shown in Fig. 4.13.

In Fig. 4.14 the DNA components of tweezers set 1 and 2 are visualised in polyacrylamide gel electrophoresis. As the gels reveal, the tweezers were formed during the annealing process. To avoid high background fluorescence, the arm strands were used in a 5% excess with respect to the doubly modified base strand. Upon addition of the corresponding closing strand in stoichiometry to tweezers an additional band appears for both tweezers set 1 and 2 indicating that tweezers are closed. This assumption is backed up by fluorescence measurements of the closing (2x excess of closing strands) and opening (4x excess of opening strands) in control experiments in which neither oscillator nor insulator components are present as shown in Fig. 4.15 (a) and (b). Gel studies further show that the closing efficiency is lower than expected from thermodynamic considerations. This could be due to several reasons. First, the excess of arm strands may absorb the closing strands. Second, higher order complexes could be formed since closing strands could also connect multiple tweezers instead of closing individual. Third, the closing dynamics are rather slow thus reactions did not reach equilibrium. Fourth, strands were not mixed in stoichiometry. The first mentioned possibility can be excluded since for both tweezers sets either arm and closing strand bands appear in the "Tweezer + TwCl" lanes. Further, higher order complexes could be formed. However, the band which is suspected to represent the fraction of closed tweezers is the largest complex in the corresponding lanes and also their wells do not show remarkable residuals compared to neighbouring lanes. In fluorescence studies (Fig. 4.15 (a) and (b)), closing and opening dynamics are fast with respect to time scales of gel preparation. Even though closing strands were added in 2x excess the

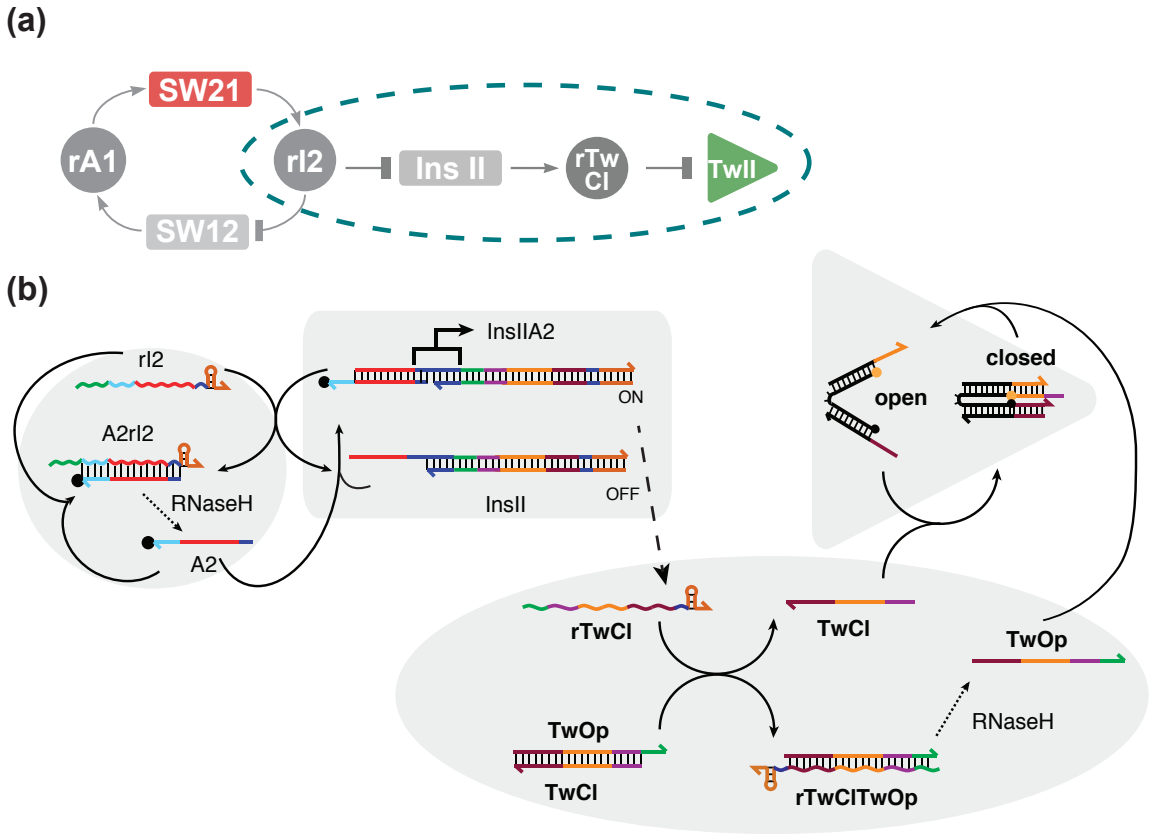

Figure 4.13: Coupling scheme of tweezers set II. (a) Simple scheme of the coupling concept. Transcription activity of *InsII* is regulated in the same manner as for coupling of tweezers set I shown in Fig. 4.12. The coding sequence *InsII* encodes for an RNA effector strand that causes closing of tweezers II. (b) More detailed scheme of the molecular interactions for tweezers set II coupling. As *InsI*, the activities of *InsII* and *Sw12* are synchronously modulated by *rI2*. When *InsII* is activated by *A2*, RNA effector *rTwCl* is produced which releases DNA closing strand *TwCl* by toehold mediated strand displacement. *TwCl* hybridises with the complementary arms of the opened tweezers *TwII* and closes it. RNase H degrades *rTwCl* from the hybrid *rTwCl:TwOp* which releases *TwOp* which opens *TwII* due to toehold accelerated displacement of *TwCl* and resets the initially opened state of *TwII*. The conformational state of *TwII* is monitored by the fluorescence of a dye-quencher (orange/ black circles) pair attached to both ends of the base strand. Consequently, as for *TwI* fluorescence is low in the closed state and high in the opened.

fluorescence for tweezers set 2 drops only to 50% of its maximum in the opened state hence only 50% are effectively closed. The fluorescence time course for tweezers set 2 further indicates a multi-step reaction during closing and opening. This was partly expected for closing of *TwII* by the RNA opening strand *rTwOp* in the presence of the DNA complex *TwOp:TwCl*. *rTwOp* is complementary to the single stranded arm domains of *TwII* in the opened state while it can only bind to the 8 bases toehold of *TwOp* to release DNA *TwCl* from *TwOp:TwCl*. While this appears plausible for closing by *rTwOp* in presence of the DNA complex *TwOp:TwCl*, the two-step closing after addition of DNA *TwCl* remains unclear. As fluorescence measurements (Fig. 4.15 (b)) show, even when effector strands are added in excess, only part of tweezers undergo the desired conformation changes. Despite considerable space for enhancement and the need of more detailed investigations especially in case of tweezers set 2, both tweezers sets were coupled to the oscillator circuit and the reactions were followed in multi-wave length time course mode on a fluorescence spectrometer. 100 nM were coupled to the oscillator's dynamics through 25 nM insulator genelets. The influence of this load on the oscillator performance can be compared by means of the corresponding fluorescence time courses of *T21* OFF as shown in Fig. 4.15 (c). While the difference in dynamics of the bare oscillator and the combination of oscillator with 100 nM *Tw I* is low (for the first three peaks), the coupling of tweezers set 2 seems to have dominating impact on the oscillatory behaviour. Even when both 100 nM *TwI* and 100 nM *TwII* are coupled simultaneously, the fluorescence trace is almost identical to that when only *TwII* is driven. Fig. 4.15 (d) shows the overlaid normalised fluorescence traces of oscillator and *TwI*. As it was expected from enzyme free conformation switching experiments as shown in Fig. 4.15 (a), *TwI* oscillates with nice amplitude and phase shifted. *TwII* coupled to the oscillator is shown in Fig. 4.15 (e). The fluorescence amplitude corresponding to the conformation change of *TwII* is much smaller than for *TwI*. However, the intermediate reaction which inverts the conformation change with respect to tweezers set 1 results in a more synchronous action of *T21* OFF and the tweezers' closing and opening. Finally, 100 nM of both tweezers sets were coupled each through 25 nM insulator genelets. Fig. 4.15 (f) shows overlaid normalised fluorescence time traces for the pace making oscillator circuit, *TwI* and *TwII*. Even though the results are only preliminary and thorough investigation has to be done, this circuit represents the largest *in vitro* transcription circuit which has been realised so far involving 20 DNA strands. As it was already indicated in [63], these results further proof the potential of the oscillator to orchestrate diverse downstream processes and scalability of genelet circuits to build larger regulatory network-like structures.

Inspired by small interfering RNAs [170], further insulators were designed which encoded either the fluorescent aptamers for malachite green [171] or DFHBI [172] extended with an 8 bases toehold, or their complementary sequences - called anti-insulators. Insulators and anti-insulators were coupled to "opposite" activators with respect to the oscillator scheme shown in Fig. 4.2. Consequently, in the oscillator's phase when insulators were active fluorescence decreased while in the phase when anti-insulators were active the functional, fluorescent aptamer was annihilated due to displacement of the dye by the anti-aptamer. This project was previously discussed in detail in [173]. Therefore, only the basic concept as well as preliminary results of proof of concept experiments are presented in the following. A schematic of the

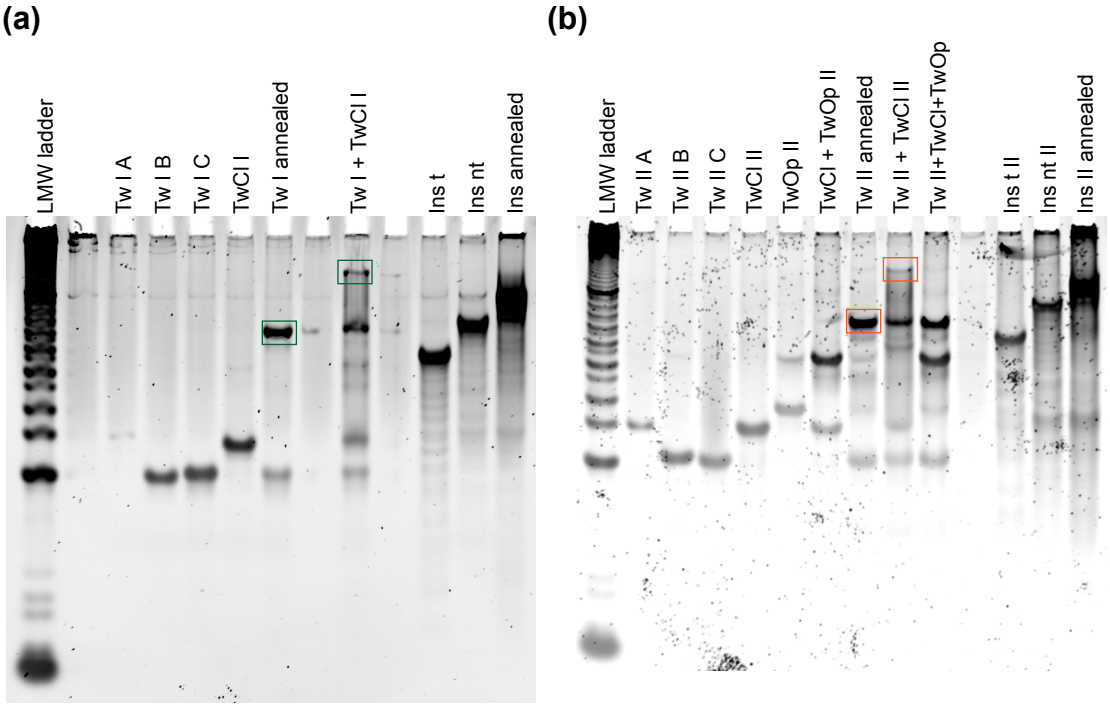

Figure 4.14: 10% polyacrylamide control gels of tweezers annealing stained in 1x SybrGold™ during 30 min incubation. (a) Tweezers set I adopted from [63]. Double modified DNA strand $TwI\,A$ is only barely visible due to its quencher label. To form tweezers I (TwI), $TwI\,A$, $TwI\,B$, and $TwI\,C$ are annealed with 5% excess of arm strands in respect to the basis. The band which corresponds to the annealed complex TwI is indicated by a green box. The excess of arm strands is clearly visible. When closing strand $TwCl\,I$ is added to tweezers, an additional band towards longer DNA strands appears (green box in lane "Tw I + TwCl I"). "Ins annealed" is the annealed duplex of insulator strands $Ins\,t$ and $Ins\,nt$ which encodes for the RNA opening strand. Multiple bands towards shorter lengths indicate that both single strands were subject to some extent of degradation. Reactions involving tweezers set I strands are shown in Fig. 4.15. (b) Tweezers set II. As in (a), basis and arm strands were annealed with 5% excess of arm strands. Lane "Tw II annealed" shows the annealing result for tweezers II ($TwII$). The band associated with the annealed DNA tweezers is highlighted by a orange box. Upon addition of closing strand $TwCl\,II$ an additional band occurs (lane "Tw II + TwCl II"). Further addition of DNA opening strand $TwOp\,II$ causes the previously generated band to vanish and results in an additional band at shorter DNA lengths which can be assigned to the complex as in lane "TwCl + TwOp II". "Ins II annealed" is the annealed duplex of insulator strands $Ins\,t\,II$ and $Ins\,nt\,II$ which encodes for the RNA closing strand. Multiple bands towards shorter lengths indicate that both single strands were subject to some extent of degradation. Reactions involving tweezers set II strands are shown in Fig. 4.15.

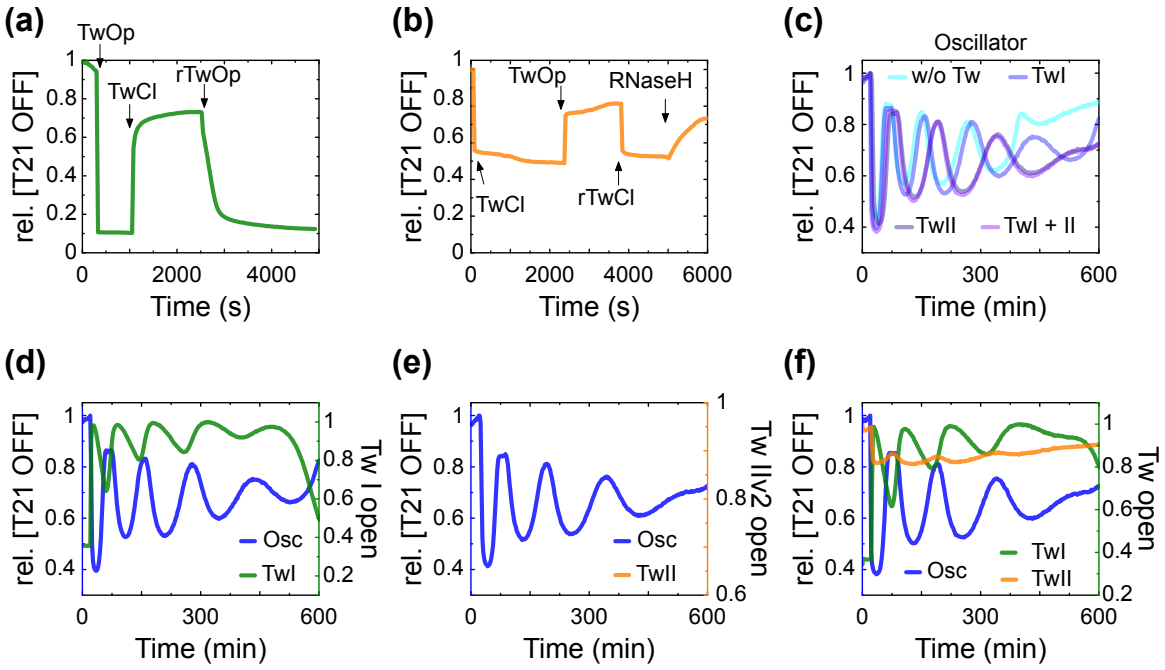

Figure 4.15: Controlling DNA tweezers performance with a transcriptional oscillator. (a) Closing and opening of *TwI* upon addition of DNA closing and opening strands in a reaction where only tweezers strands are present. Tweezers DNA strands, result of annealing and closing reaction is shown in Fig. 4.14 (a). (b) Closing and opening of *TwII* upon addition of DNA closing and opening strands in a reaction where only tweezers strands are present. Upon addition of transcription mix in which RNA closing strand *rTwCl* was synthesised fluorescence drops again indicating the closing of tweezers. Addition of RNase H causes degradation of RNA bound to DNA and leads to the opening of *TwII*. Tweezers DNA strands, result of annealing and closing reaction is shown in Fig. 4.14 (b). (c) Oscillator traces for reactions without tweezers, with tweezers set I, with tweezers set II, and with both tweezers sets. (d) Tweezers set I (green) coupled to the oscillator (blue). (e) Tweezers set II (orange) coupled to the oscillator (blue). (f) Tweezers set I (green) and II (orange) coupled to the oscillator (blue).

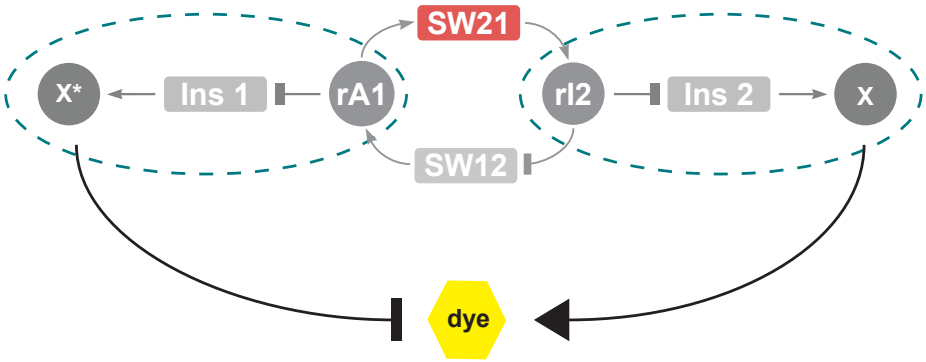

Figure 4.16: Coupling scheme for fluorescence switching by aptamer - anti-aptamer reaction. Aptamer X or anti-aptamers X^* are produced in different phases of an oscillation cycle. A dye which is present in the reaction buffer turns on fluorescence upon specific binding to its aptamer. The aptamer sequence was extended by an eight bases toehold which serves as binding site for an anti-aptamer. An anti-aptamer is the complementary sequence which opens the defined secondary structure of the aptamer and displaces the dye, which leads to a dropping fluorescence.

coupling concept is shown in Fig. 4.16. In contrast to the tweezers coupling discussed above, two different kinds of genelets are coupled to the oscillator's dynamics by two different RNA signals. In analogy to the tweezers insulators, the transcriptional state of the genelet which encodes for the aptamer X is regulated by RNA inhibitor *rI2*. Aptamer X has a secondary structure in which a dye ligand binds to. Upon binding of the dye and aptamer, energy states are stabilised and the dyes fluorescence increases. A second genelet which encodes for the anti-aptamer, a complementary sequence of the aptamer, is transcriptionally regulated by *rA1*. When the anti-aptamer is produced it binds and destroys the aptamer dye complex which causes fluorescence to drop. We successfully used the aptamers and anti-aptamers for malachite green [171] and DFHBI [172]. Fig. 4.17 shows resulting traces for coupling of malachite green/ anti-malachite green (a), DFHBI/ anti-DFHBI (b), as well as the combination of both pairs (c). The operation of both aptamer and anti-aptamer genelets in parallel is constituted by a circuit consisting of 6 genelets and in total 15 DNA strands. We here further demonstrated the possibility to use the oscillator as master clock to time the operation of DNA nanodevices as well as the production of functional RNAs by an *in vitro* transcriptional oscillator. Since the coding sequence of genelet which is coupled to the oscillator can be chosen almost arbitrarily, a variety of RNAs could be produced in a timed fashion. So far, molecular signals were used to regulate either the quantum yield of a dye present in the buffer or to modulate conformational state of DNA-nanodevices. Further reactions which could be coupled are the control of pattern formation or the generation of a cytoskeleton-like dynamical structure. These concepts could be used to equip completely artificial cells with special properties.
However, only simpler biochemical *in vitro* systems have been encapsulated into compartments on the length scale of living cells. To this end, the discussed oscillator is used in the next section to investigate the possibilities and challenges of compartmentalising reaction circuits into cell-sized compartments.

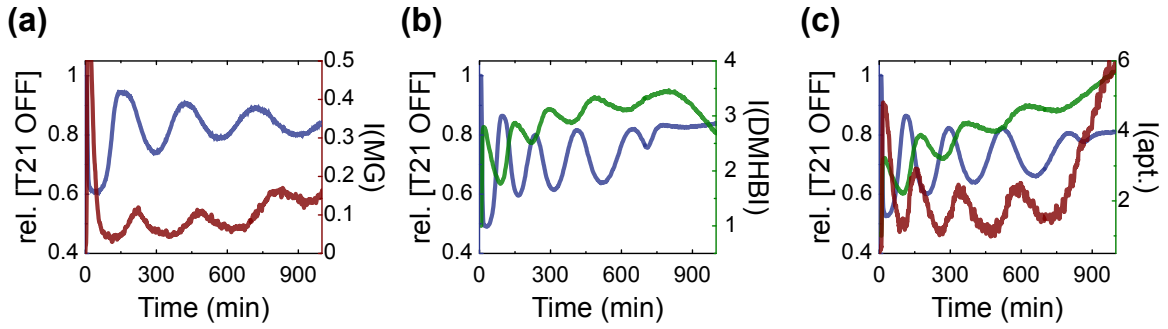

Figure 4.17: Controlling aptamer and anti-aptamer production. "rel. [T21OFF]" is the relative concentration of template *T21* in the transcriptional inactive state (blue trace). Malachite green fluorescence traces are drawn brown, DFHBI traces are green. Malachite green traces are corrected for cross-talk from Texas Red and normalised to maximum and minimum fluorescence values. (a) The oscillator drives the phase shifted production of a malachite green aptamer and its anti-aptamer. The functional aptamer causes the fluorescence intensity of the malachite green dye to increase (I(MG)). (b) Conceptually identical with a but with DFHBI and the corresponding aptamer/ anti-aptamer instead. (c) Simultaneous operation of malachite green and DFHBI aptamer - anti-aptamer pair. The malachite green trace's y-axis range is $[0, 1]$. The influence of cross-talk corrections shown in Fig. 4.35.

4.4 Compartmentalization of a biochemical oscillator

An important step toward engineering compact biomimetic systems is the encapsulation of biochemical circuitry within cell-like microcompartments - the creation of programmable "artificial cells" or "protocells". However, compartmentalization can profoundly influence biochemical reaction kinetics [174, 175]. In particular, typical concentrations of molecules in biomolecular circuits are in the nanomolar range, which for cell-sized reaction containers means that some molecular species are present only at very small copy numbers. A prominent consequence of this is the appearance of biochemical noise, which often is attributed to the inherent stochasticity of chemical reactions as introduced in section 2.1.6 [28]. Understanding the sources and propagation of noise and variability in artificial biochemical systems encapsulated in microcompartments is necessary to provide a foundation for engineering molecular systems that are robust to stochasticity – or that exploit the randomness to advantage. Furthermore, because biological cells face many of the same challenges and opportunities, our understanding of biological systems can be informed by, and can provide insight into, engineering principles for molecular systems at this scale. Cell-free gene expression in lipid bilayer vesicles has been studied in the past, but functional encapsulation of complex mixtures of biochemicals such as the transcription/translation machinery remains challenging [49, 53]. A technologically more developed approach employs water-in-oil emulsion droplets as reaction containers [175, 176, 177], with applications including single-molecule enzymology [111], emulsion PCR [112], and *in vitro* evolution experiments for the selection of ribozymes or functional proteins [113, 114] as introduced in section 2.3. However, thus far the influence of micron-scale encapsulation on *in vitro* biochemical networks with more complex, far from equilibrium dynamical behaviour as the transcriptional oscillator presented in section 4.2 has not been systematically investigated. Compartmentalization of an inorganic chemical dynamical process was

previously demonstrated with the Belousov-Zhabotinsky reaction, where the emergence of oscillations and spatial patterns was studied in microemulsion droplets [178]. As inorganic chemical oscillators operate at much higher concentrations than typical biochemical systems, fluctuations and small number effects were, with few exceptions [179, 115], typically not observed. Recently, a range of synthetic biochemical oscillators have been successfully demonstrated both *in vivo* [1, 180, 2, 181, 40] as well as *in vitro* [5, 182, 4, 63]. There has already been considerable interest in the robustness of naturally occurring biochemical clocks with respect to molecular noise [183, 44, 184] or temperature fluctuations [185] due to their important role in the orchestration of biological processes [186, 183]. So far synthetic *in vitro* biochemical oscillators have been studied only in bulk reactions [5, 4, 63, 59], or encapsulated into emulsion droplets too large to result in considerable dynamic variability [60].
Here, we studied compartmentalization with two different approaches of emulsification. We initially applied a microfluidic technique as introduced in section 2.3 to encapsulate the reaction circuit presented in paragraph 4.4.1. Alternatively, we used simple vortexing to encapsulate the transcriptional oscillator into a population of droplets with a much broader size distribution compared to droplets produced in microfluidics. This method enabled us to analyse the oscillator dynamics in volumes between 33 fL and 16 pL as will be discussed in paragraph 4.4.2.

4.4.1 A biochemical oscillator in microfluidic droplets

In order to study the transcriptional oscillator in cell-scaled compartments we first used a microfluidic flow-focussing device as described in chapter 2.3.2 for encapsulation. Two immediate drawbacks of this technique are that the enzymes had to be added before the encapsulation started such that no baseline could be recorded. For normalisation, we assumed the peak of the first maximum to be at an intensity of 90% of the baseline. This assumption was based on a series of experiments performed with the same oscillator tuning. Furthermore, due to the relatively elaborate technique, it typically took at least 15 min but usually rather 60 min to generate the droplets. To observe which influence the reaction liquid handling in syringe and tubing as well as in general the lag-time during droplet generation on the dynamics of the oscillator had, reference samples were recorded in the fluorescence spectrometer before and after the procedure. In Fig. 4.18 (a) two time traces are shown which correspond to aliquots of the same reaction solution put into the spectrometer before (blue) and after (green) the encapsulation. The time-lag was 72 min. The sample coined "after" was extracted from the syringe-tubing system. Since first attempts failed, presumably due to adhesion or denaturation of enzymes in the syringe-tubing system we passivated these compartments by incubation in 5% $^{w}/_{v}$ BSA solution for at least 24 h. In Fig. 4.18 (b), the trace "after droplet generation" was shifted by 72 min such that it also started at zero for better comparison purpose. The dynamics of the sample which was used for encapsulation showed a slight shift towards slower periods. Since droplet production was performed at room temperature, the initial thought was that the enzymatic reactions may be considerably slower. We performed an oscillator experiment in the fluorescence spectrometer at room temperature to find out what the effect of the reduced temperature is. The corresponding time trace

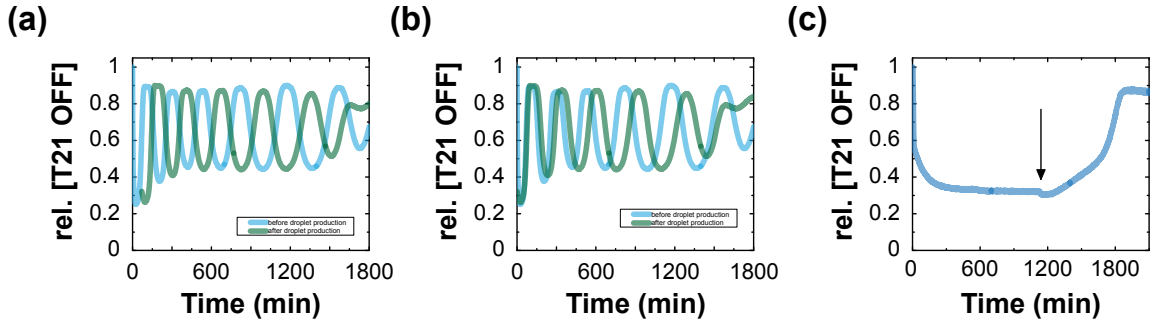

Figure 4.18: Influence of microfluidic droplet generation on the oscillator. (a) Reference time traces of the oscillator taken from the same batch as droplets were produced. An aliquot was taken before the solution was filled into the syringe-tubing system (blue). After droplets were generated and the microscope was started another aliquot was put to the spectrometer (green). The time-lag is 72 min. (b) As (a) but with a 72 min shift of the green trace for better comparison purpose. (c) An oscillator in the spectrometer at room temperature. The heating was turned on at the time point highlighted by an arrow.

is shown in Fig. 4.18 (c). The reaction was started by addition of the enzymes to the cuvette, hence the trace was normalised to its initial baseline. The initially fast dropping fluorescence indicates that RNAP was working with high efficiency while RNase H seemed to be significantly slower than at its optimal operation temperature. The reaction shown in Fig. 4.18 (c) contains identical species concentration as reaction #57 shown in Fig. 4.9. 20 h after start of the experiment, the heating was turned on to 37 °C which caused the signal to increase since RNase H activity was enhanced. From these experiments and further investigations on RNAP activity, we concluded that the quite strong robustness of RNAP against temperature variations and the sensitivity of RNase H may lead to an imbalance in initial conditions which is hard to predict and quantitatively to analyse. Even though the deviation in the spectrometer references was not too large, barely any droplet showed oscillations - at least with larger amplitude than the noise level. However, at a research visit at the Franco Lab at UC Riverside, it was possible due to a heated microscopy stage to produce droplets at 37 °C. In Fig. 4.19 (a) microfluidic droplets containing the transcriptional oscillator are shown. The reference traces were still different (Fig. 4.19 (b)) which could be caused by the fact that the syringe could not be heated but the tubing which was in contact with the heated stage. The lower panel in Fig. 4.19 (b) shows example traces of single droplets. The population mean is overlaid as black trace. Interestingly, the mean frequency in droplets was increased by approximately a factor of two. While reference samples "before" and "after" droplet generation oscillate with periods of about 180 min and 200 min, respectively, the droplet mean oscillates with a period of about 90 min (when determined as mean of the two oscillations between the first and the third peak). An image time series of oscillating droplets and the corresponding time traces is depicted in Fig. 4.20. The accelerated oscillations in droplets may be due to an overall loss of enzymes on the microfluidic channel's walls during the droplet generation process. The channels of the microfluidic chips were not treated with sophisticated surface coating methods. The only modification was due to a second heating step after baking of the PDMS devices, a process which was reported to result in hydrophobic glass surface of the channel closing microscopy slide with a contact angle of 109 ° which is equal to that of native PDMS [151]. Presumably

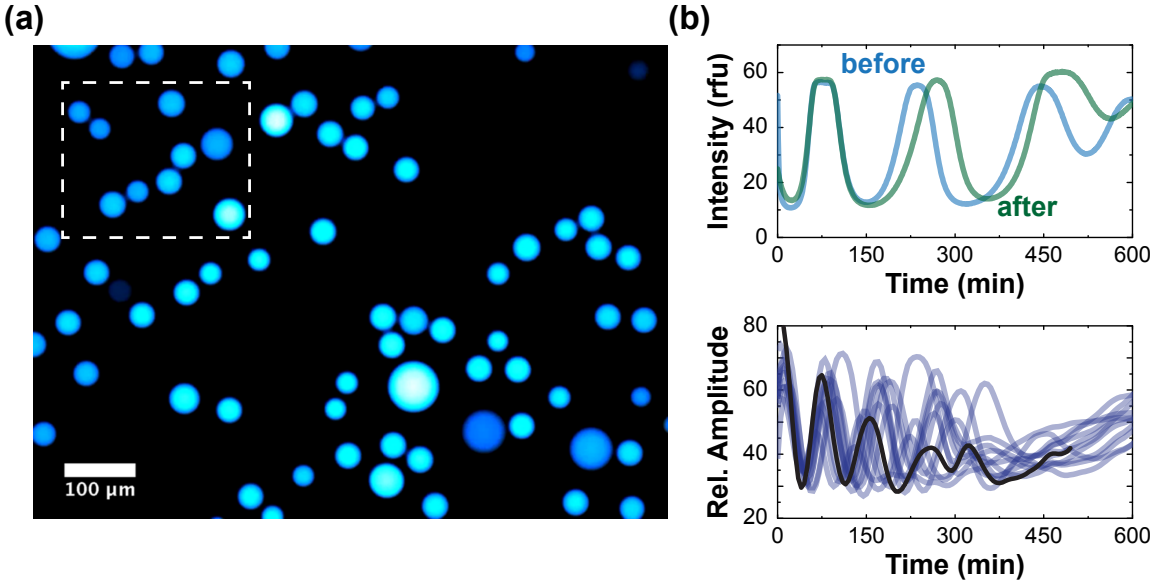

Figure 4.19: The transcriptional oscillator in microfluidic droplets. (a) Fluorescence micrograph of droplets produced in microfluidics containing the transcriptional oscillator. Droplets were produced at 37 °C. The scale bar is 100 µm (b) Upper panel: Fluorescence spectrometer references as described in caption for Fig. 4.18 (b). Lower panel: Example traces of oscillations in microfluidic droplets shown in (a). The black trace represents the population mean.

due to its non-linearity, the oscillator circuit typically responds strongly to changes in enzyme concentrations. Therefore it is quite interesting that on the one hand the oscillations in droplets differ significantly with respect to the spectrometer reference but on the other hand that only few droplets did not oscillate at all. Either for this particular experiment the oscillator was tuned to an especially robust operating point in terms of robustness against variations in enzyme concentrations or the accelerated frequency in small volumes is a result of the confined volume itself.

Droplet traces not only differ from the bulk measurements, there are also quite strong variations in the oscillator dynamics between different droplets. For instance, in Fig. 4.20, even though droplets #3 and #5 are equally sized, #3 dampens out after approximately 150 min while #5 oscillates with quite large amplitude for 300 min. This is in contrast to recently published results where an *in vitro* biochemical oscillator, based on a nicking amplification reaction, was studied in microfluidic droplets [58, 60]. Despite a slight phase shift for later runtimes, no significant difference between individual droplets was reported. This may be either due to the fact that effects by volume constriction are smaller since mean droplet radii were on the order of 50 µm or that the circuit used in this case was more stable against the influences that caused the diversity in dynamics in our case. However, further experiments have to be performed for more substantial conclusions. Instead of going into detail with the microfluidic system, we chose to concentrate on a different droplet generation approach which enabled us to produce a population of droplet oscillators with a wide size distribution ranging from one up to tens of micrometers in a minimal amount of time. The results of these studies will be presented in the next section 4.4.2.

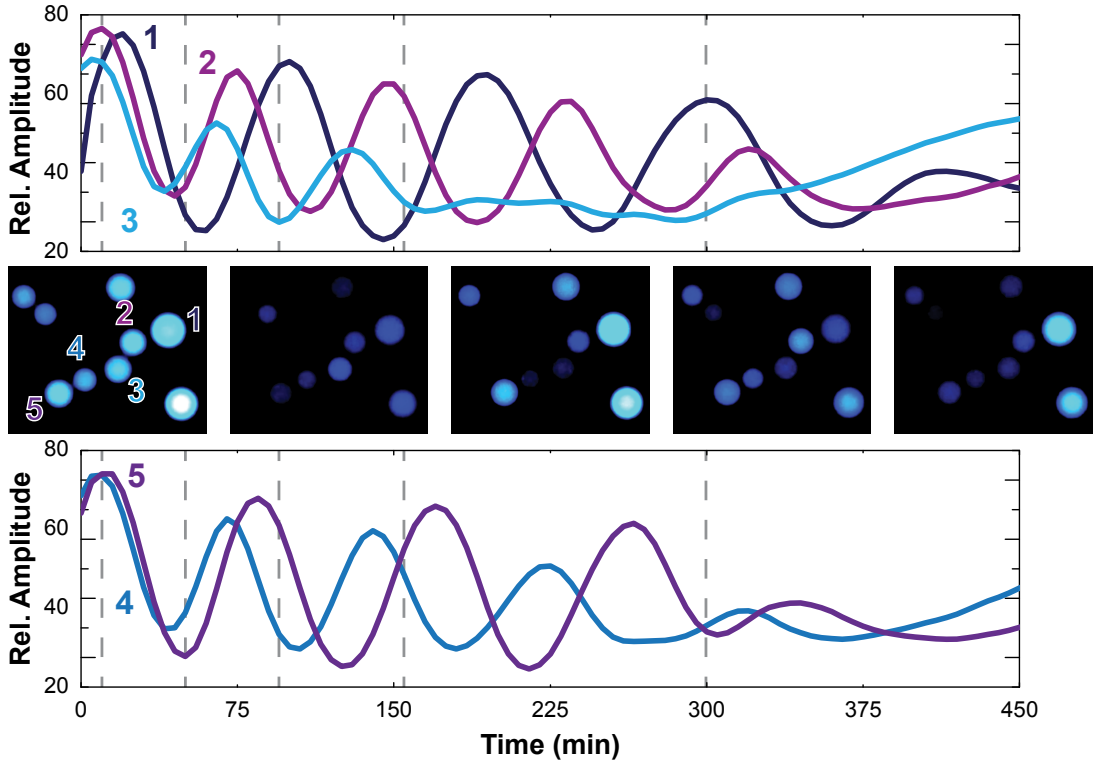

Figure 4.20: Image time series of oscillations in microfluidic droplets. The image series shows the position from Fig. 4.19 (a) highlighted by a dashed white box. Fluorescence time traces which were normalised by their droplets volume (Rel. Amplitude = $I_{droplet,i}(t)/V_{droplet,i}(t)$). The time points of fluorescence microscopy images are indicated by grey dashed lines. Droplet traces are assigned to corresponding droplets by equaled coloured numbers in the first slice. Droplet radii are $R(\#1) = 22\,\mu m$, $R(\#2) = 17\,\mu m$, $R(\#3) = 17.5\,\mu m$, $R(\#4) = 15\,\mu m$, $R(\#5) = 17.5\,\mu m$.

4.4.2 A biochemical oscillator in shaken-not-stirred droplets

Using a simple vortexing technique (see section 3.5.2), emulsions of oscillator containing water-in-oil droplets with volumes ranging from $\gtrsim 16\,\mathrm{pL}$ down to 33 fL were generated. Thousands of individual oscillating droplets were followed simultaneously by optical microscopy, and features such as the period and amplitude of each droplet's fluorescence trace were measured using automated analysis procedures. As anticipated for small number effects, the diversity of behaviours increased in smaller droplets. Surprisingly, however, we found that the diversity could not be explained by the inherent stochasticity of chemical reactions involving small numbers of molecules, which is often used as an explanation for variability within living cells [93, 33]. Instead, our experimental data was more consistent with models in which stochastic partitioning of key low-concentration species (enzymes in particular) introduced diversity in the dynamical behaviour by providing variability in the initial conditions for the oscillator. Thus, the variability observed in our population of oscillators bears more similarities to "cell division noise" that is caused by the unequal distribution of the molecules of dividing cells between their daughter cells [91, 187].
This section presents and discusses results recently published in *Nature Chemistry* [64]. The work had been a collaboration project of a group of people between September 2011 and February 2014. Experimental work, data processing of bulk experiments as well as image analysis of time-lapse movies was done by me. Jongmin Kim was mainly responsible for the simulation part, Korbinian Kapsner for data analysis.

Influence of compartmentalization on subsystems

To validate measurements of droplet-encapsulated molecules such as labeled DNA and proteins, we performed "shaken-not-stirred" droplet control experiments using different fluorescent markers.

First, we ran simple experiments to assess the distribution of dyes inside a single droplet. We examined the distributions of 150 nM (bulk) Texas Red-labeled *T21* genelet, and 2 µM (bulk) green fluorescent protein (GFP), simultaneously encapsulated in the same emulsion following the procedure previously described in section 3.5.2. The two species were suspended in 1x NEB transcription buffer as used for annealing and oscillator experiments. In order to compare fluorescence profiles of dyes with different emission wavelengths without changing the focal plane, we used an Olympus Apochromat 10x objective for this specific experiment. The profile plot of a droplet containing Texas Red labeled DNA switches (*T21* OFF) and GFP are shown in Fig. 4.21 (a). The overlaying fluorescence cross sections indicate that no molecular species is predominantly attached to the oil-buffer interface and the molecular distribution of both in the compartments is similar. As previously discussed in section 2.3.1, the utilised E2K0660 surfactant was invented for biological purposes, and, due to its passive PEG headgroups oriented towards the compartment, we assumed that degradation and adsorption at the inner droplet surface is not a major cause of enzyme loss. We further encapsulated RNAP modified with two different fluorescence labels. Fig. 4.21 (d) shows a fluorescence micrograph in 10x magnification of Alexa488-RNAP (a generous gift from Craig T. Martin and Luis E. Ramírez-Tapia of the University of Massachusetts) in microdroplets. Fig. 4.21 (e) shows bright field and fluorescence microscopy images in 40x magnification of Atto647-RNAP (gener-

ously provided by Evi Stahl from the Dietz lab at TUM) in microemulsion droplets using the same buffer conditions and reaction temperature as in oscillator time-lapse microscopy experiments. As for the oscillator, the sample was prepared in a reaction tube and mixed on a benchtop vortex before enzymes were added - in this case only RNAP. Thereafter, RNAP was added and the solution was mixed by vigorous pipetting. 10 µL were transferred to an Eppendorf Protein LoBind reaction tube containing 45 µL oil-surfactant-mix and encapsulated into microemulsions. Time-lapse movies reveal the presence of fluorescent aggregates in the droplets in the fluorescence channel corresponding to the RNAP dye label but not in the channel for the emission wavelength of a reference dye which was also present in solution. Microscopy images of bulk solution (without encapsulation) on passivated microscope slides did not show aggregation of RNAP. Second, we compared the correlation of fluorescence intensity and droplet size for different molecules encapsulated into microemulsion droplets. We analysed microscopy images of droplets containing either 200 nM (bulk solution) Alexa488-labeled T7 RNA polymerase, 150 nM (bulk) Texas Red-labeled T21 genelet, or 1 µM (bulk) green fluorescent protein (GFP). These molecules were suspended in 1x oscillator reaction buffer (1x NEB transcription buffer, 24 $MgCl_2$, 7.5 mM each NTP, in nuclease free water), and were individually encapsulated in separate emulsions, following the procedure described in section 3.5.2. Brightfield and fluorescence images were recorded over a period of two hours using a NIKON Eclipse TI-E inverted microscope, controlled by Nikon's software NIS Elements, with a CFI Plan Fluor 10x objective. Droplet fluorescence, area, and perimeter data were collected and processed using the same procedures described for the oscillator traces as outlined in section 4.6.5. In addition, fluorescence traces were eliminated if, over the two hour period, the standard deviation of radius measurements exceeded 0.25 µm, the fluorescence standard deviation exceeded 4% the mean, or if the calculated area using the measured perimeter did vary more than 30% in respect to the measured area which indicates that the regarded object is not one circular droplet cross-section. The last criterium was particularly for the analysis of this experiment necessary since no reference dye was present in the droplet content which was used for oscillator droplets to distinguish between real single droplets or clusters of multiple droplets and areas mis-interpreted as droplets, respectively (cf. section 4.6.5).

In Fig. 4.22, measured mean fluorescence intensity versus mean droplet radius is plotted for individual droplets. Fluorescence and radius data for individual droplets were averaged over 2 h for Texas Red-labeled *T21*, and for 10 min for Alexa488-labeled RNAP and GFP which were affected by 5% bleaching on average. Using MATLAB, we fitted our mean fluorescence data to the mean droplet radius R with a power law $y(R) = a \cdot R^n$ as drawn as solid lines in Fig. 4.22. For all three fluorescent molecules, the fit resulted in an exponent of ≈ 3 suggesting that the measured fluorescence intensity scales with the droplet volume, and thus with the total number of molecules in a droplet. The fact that the resulting exponent does not exactly match $n = 3$ is very likely due to an incomplete collection of light emitted by an object that is not planar and situated at the focal plane.

We further performed control experiments with single-enzyme subsystems simpler than the transcriptional oscillator to observe molecular count distributions, and to assess the accuracy and reliability of our experimental and analytical measurement techniques. Specifically, we compartmentalized simple systems for only transcription or

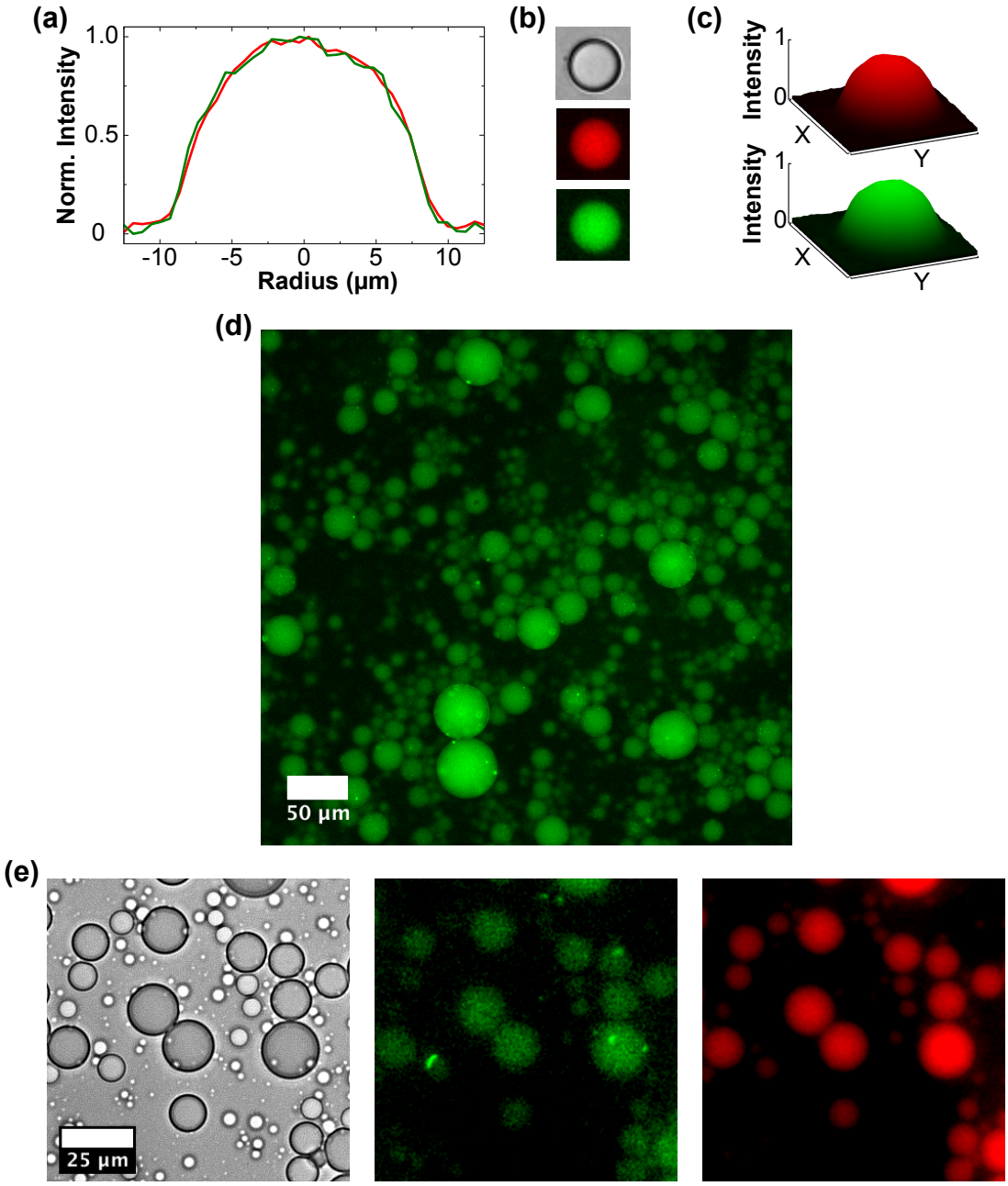

Figure 4.21: (a) Profile plots of normalized droplet fluorescence recorded in the Texas Red (red) and GFP (green) channel. (b) Brightfield (gray) and epifluorescence images in Texas Red (red) and GFP (green) channel of an emulsion droplet containing Texas Red labeled genelet T21 and green fluorescent protein. (c) Surface plots of droplets shown in (b). (d) Fluorescence microscopy image of Alexa488-RNAP in microemulsion droplets in 10x magnification. (e) Microscopy images in 40x magnification of Atto647-RNAP and Alexa488 in droplets in brightfield (gray), Atto647 (green) and Alexa488 (red) channel (false colors are used for better comparison with (d)). Thus, aggregation is observed in the two labeled RNAP samples (Alexa488-RNAP and Atto647-RNAP) but not in the control samples containing GFP, DNA template, or pure dye.

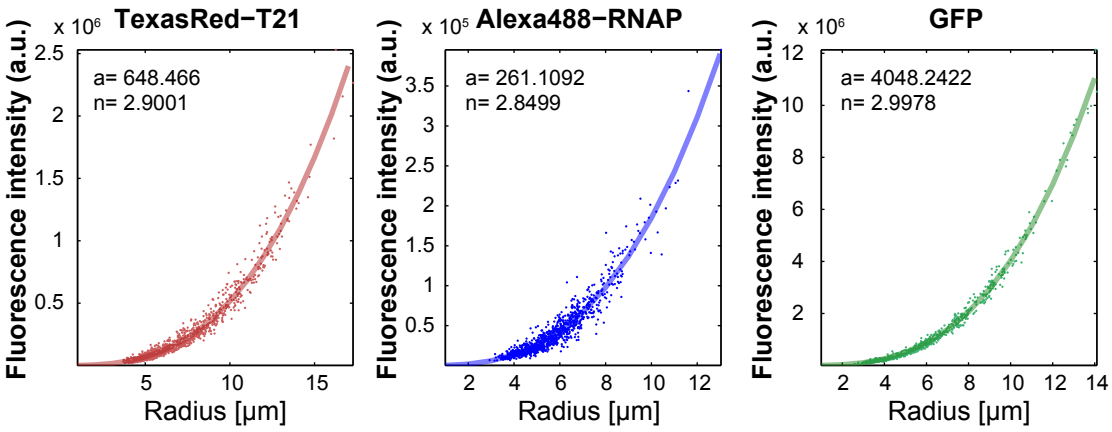

Figure 4.22: Correlation between fluorescence intensity and droplet size for encapsulated Texas Red labeled template *T21*, Alexa 488 labeled RNAP and GFP using the shaken-not-stirred technique. Dots represent single droplet results for their individual mean fluorescence (fluorescence intensity per area) and their radii. Solid lines are derived from fits of the function $y(R) = a \cdot R^n$ to the data points. Parameters a and n derived by the fitting are drawn into the corresponding graphs. Data was analysed and plotted by Elisa Franco.

only degradation of RNA signals. Schemes describing the two enzyme characterisation subsystems as well as results of bulk measurements in a fluorescence spectrometer and results from analysed microscopy time-lapse movies are shown in Fig. 4.23. The RNA transcription system consists of a fully duplex, hence transcriptionally active, template and a preformed reporter consisting of a top strand (labeled with the TAMRA fluorophore at its 5′- end) and a bottom strand (labeled with Black Hole Quencher-2 at its 3′-end) as shown in Fig. 4.23 (a). Sequences are given in Table 6.7. Once the transcription is initiated by RNAP, the RNA transcript displaces the TAMRA-labeled top strand from the reporter complex so that the fluorescence signal increases.
The RNA degradation system contains the aforementioned reporter strands mixed with a separately produced RNA strand such that the TAMRA-labeled top strand is initially free in solution as illustrated in Fig. 4.23 (b). Once the degradation reaction is initiated by addition of RNase H, the RNA strand within RNA-DNA hybrid complex is degraded which allows the fluorophore-labeled top strand to bind to the quencher-labeled bottom strand, resulting in decreased fluorescence. All the other components except for DNA/RNA strands, RNAP and RNase H (i.e., transcriptional buffer components and pyrophosphatase) are chosen as in the oscillator reactions. The experiments were performed in the same manner as later discussed in section 4.4.2. For the characterisation of the RNA transcription subsystem, the fluorescence traces were analysed by determining a rise-time, Δt_{rise}, the time when a fluorescence trace spans 20% to 80% of its maximum intensity value indicated by black horizontal lines in Fig. 4.23 (c). We used this as a surrogate for the speed of transcription and hence the concentration of enzyme molecules. For the normalised bulk trajectory, the rise time was about 12 min.

Since approximately 15 min had passed between the initiation of the reaction in bulk solution and the beginning of observation on the microscope, the initial 15 min of data for microdroplets could not be observed (Fig. 4.23 (c)). Despite normalisation by Alexa 488 dye to control for variations in experimental measurements (as described in section 4.6.5), the experimental traces for microdroplets show a large

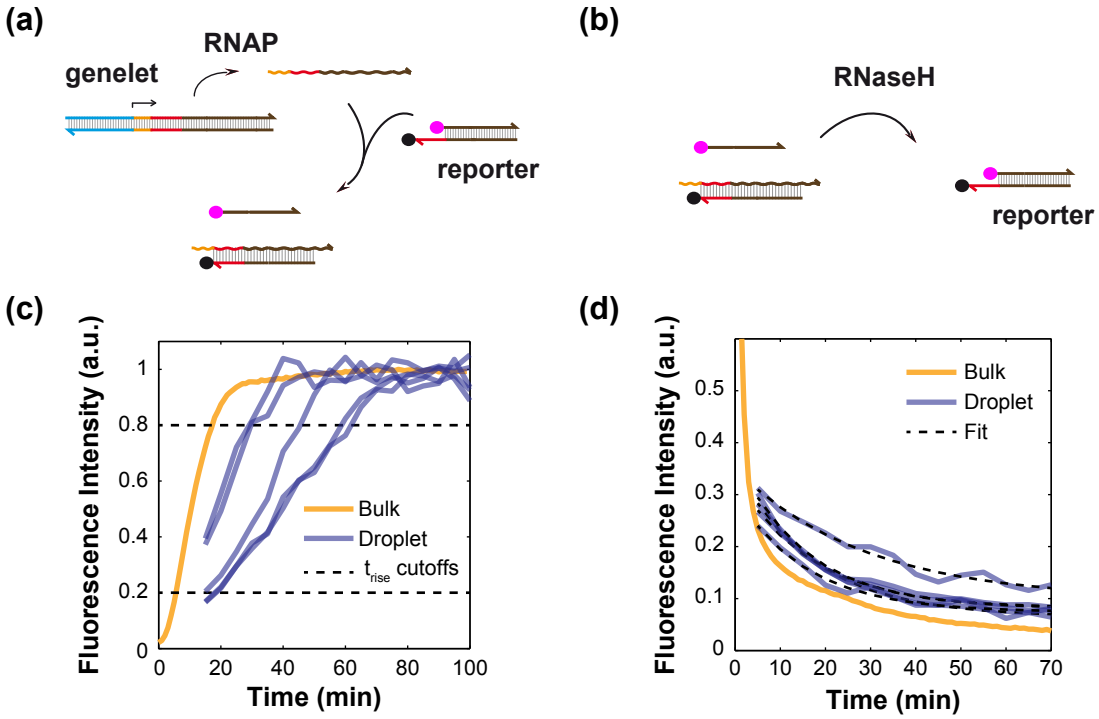

Figure 4.23: Influence of compartmentalization on single-enzyme subsystems. (a) Scheme for the transcription-only subsystem. The system contains two different, initially annealed double stranded DNA molecules, a permanently transcriptionally active genelet without a nicked promoter and a reporter complex. Upon transcription of the genelet's coding sequence into RNA by the RNAP, intensity of the reporter increases since the quencher labeled bottom strand is displaced. (b) Scheme of the RNase H degradation subsystem. The RNase H degradation subsystem consists of the reporter strands also used in (a) and an RNA strand which is complementary to the quencher labeled bottom strand. Before addition of RNase H fluorescence is high since the fluorescently labeled reporter top strand is free in solution. Upon degradation of the RNA strand from the RNA:DNA duplex, reporter double strand becomes available for binding with the dye labeled top strand and consequently causes fluorescence to decrease. (c) Normalised fluorescence time traces of the transcription-only subsystem in bulk (orange) and in droplets (blue). The dashed line indicates the t_{rise} cutoffs. Observation of the reaction in droplets starts with a 15 min offset. (d) Time traces of the RNase H degradation subsystem reaction in bulk (orange) and droplets (blue). Fits to droplet traces are drawn as black dashed lines. The offset between start of the reaction in bulk and start of the droplet observation on the microscope is 5 min.

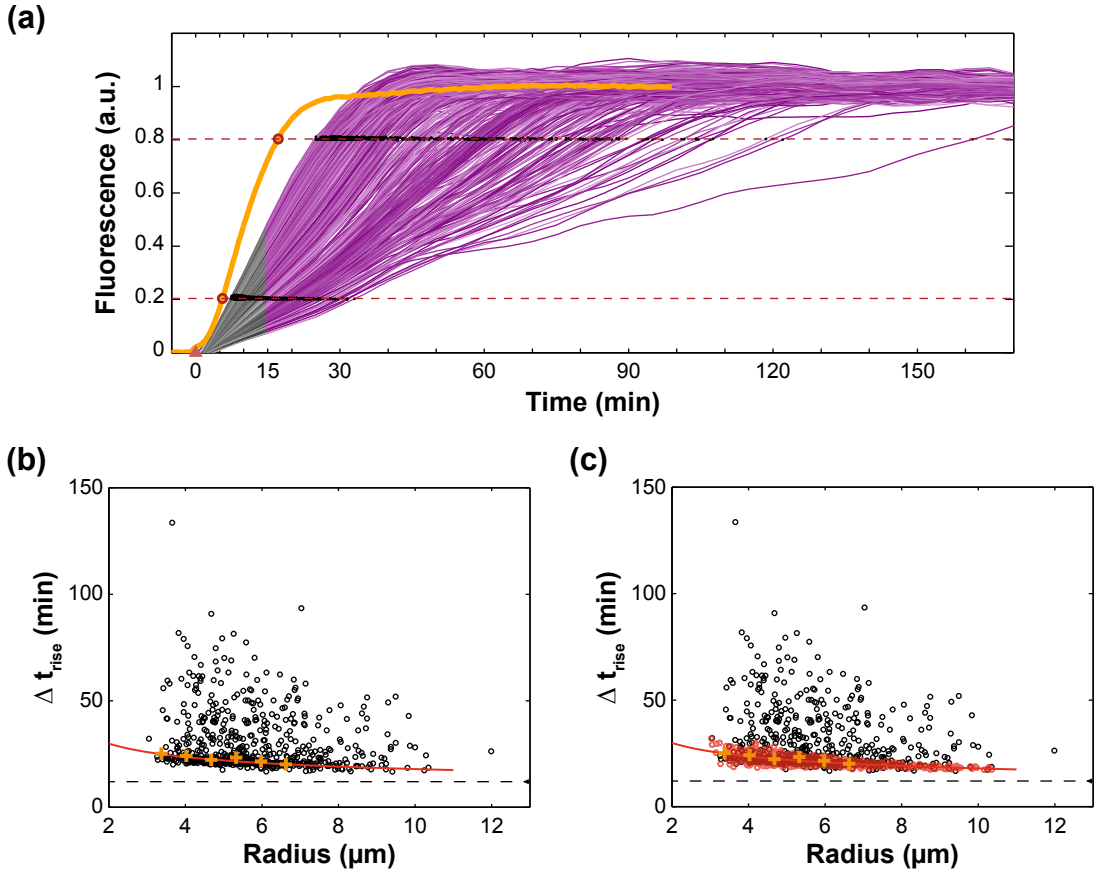

Figure 4.24: Analysis of transcription subsystem in microdroplets. (a) Normalised fluorescence time traces of the bulk reaction and rescaled and normalised droplet traces. The grey area of droplet traces was extrapolated under the assumption that the reaction at $t = 15\,\text{min}$ behaves linearly and that time traces reach 0 fluorescence intensity at $t = 2\,\text{min}$. Rise time cutoffs are indicated by horizontal dashed lines. (b) Scatter plot of single droplet rise times and corresponding droplet radii. The bulk rise time is indicated by a dashed line. The median values (orange crosses) were used to fit a function which describes the number of divisions which a droplet underwent in dependence on its final size. (c) Same as (b) but with simulated rise times assuming initial molecule numbers drawn by a Gamma distribution with $\beta = 100$.

variation in maximum values and presumably in the unobservable minimum values at the beginning. The accurate estimation of maximum and minimum values is critical to determine rise times, and therefore, we decided to perform rescaling and smoothing steps which are explained in detail in the supporting information supplementing [64]. These rescaled fluorescent traces were smoothed by 5 timepoint averages and plotted in 4.24 (a), where grey portions indicate the extrapolated timepoints. Examples of rescaled but not smoothed traces are shown in 4.23 (c). The rise times Δt_{rise} were then determined using the same procedure as for the bulk trace with cutoffs at 20% and 80% of the final mean intensity value as indicated in Figure 4.24 (a). The scatter plot in Figure 4.24 (b) shows that the rise times show larger variability for smaller droplets and that the rise times on average decrease for larger droplets. This latter trend, combined with the observation that the bulk rise time is faster than all observed droplets, suggests some loss of enzyme activity that is more significant for smaller droplets which presumably endured more splitting events.

For characterisation of the RNA degradation subsystem, as shown in Figure 4.23 (b), we measured the speed of RNA degradation. We therefore fitted an exponential decay function to the bulk and individual droplet fluorescence time traces as shown in Fig. 4.23:

$$y(t) = B + A \cdot \left(e^{-(t-t_0)/\tau} - 1\right), \tag{4.4}$$

where B is the baseline, A is a scaling factor, $t_0 = 5\,\text{min}$ is the delay between the initiation of degradation and observed traces, and τ is the degradation time constant. While an exponential decay is not the analytical solution for the RNA degradation reaction, it is a reasonable approximation when the RNA-DNA hybrid substrate level falls below the Michaelis constant of RNase H. Starting from the time t_0, which was 5 min after the mixing of enzyme and reporter complex substrates in the bulk, the droplet observation in the microscope was initiated. We chose a 70 min time window, in which the traces could be fit well by an exponential decay function, to determine the measure τ for the enzyme reaction velocity (e.g. Fig. 4.23 (d)). In Fig. 4.25 the size distribution of droplets is shown which were analysed to obtain τ values, as well as the resulting data points for all individual droplets in a scatter plot.

The calculated τ values are large for slow degradation processes and small for fast degradation processes. Therefore, we can obtain a rough estimate of RNase H activity using $1/\tau$ values. Analogous to the analysis for RNAP data, we chose to explore the fit function using median $1/\tau$ values for radius bins. The resulting fit was as follows:

$$1/\tau = 0.0516\,\text{min}^{-1} + 0.0003\,\text{min}^{-1} \times \ln(R/\mu\text{m}).$$

For comparison, the bulk τ value was 16.9 min: $1/\tau_{bulk} = 0.0592\,\text{min}^{-1}$. Unlike the RNAP data, the dependence of median $1/\tau$ values on droplet radius was negligible. This indicates a stronger sensitivity of RNAP with respect to our droplet production process, which seems to be reasonable in light of the fact that T7 RNAP is a much larger (98 kDa) and more complex enzyme than RNase H (17.6 kDa), and is expected to be more prone to denaturation and loss of function. Consequently, we infer that a global loss of RNase H activity during the transfer of reaction mixture is a plausible scenario. The reciprocal of the above fit function is shown in Fig. 4.25 (b).

The next section will present and discuss results from experiments of the transcriptional oscillator in shaken-not-stirred droplets. However, beforehand, the main findings of the previously discussed control experiments should be summarised. First, fluorescence profile of droplets containing Texas Red labeled template strands *T21* and GFP showed no predominant adhesion of molecules on the buffer-oil-boundary (Fig. 4.21 (a) - (c)). Second, encapsulation of RNAP labeled with either Alexa 488 (Fig. 4.21 (d)) or Atto 647 (Fig. 4.21 (e)) revealed fluorescent agglomerates in the fluorescence channel corresponding to the RNAP-label but not in the channel of a reference dye also present in solution. Independent of the reason for the fluorescent RNAP agglomerates, both a non-Poissonian distribution of the labeled enzymes as well as a slight reduction in overall RNAP enzyme activity seems reasonable to assume. We further investigated the dependence of fluorescence intensity of encapsulated molecules on the droplet radius. As expected from spherical compartments the intensity followed a power law of the radius with an exponent of ≈ 3 suggesting that fluorescence is directly related to the concentration of containing molecules (Fig.

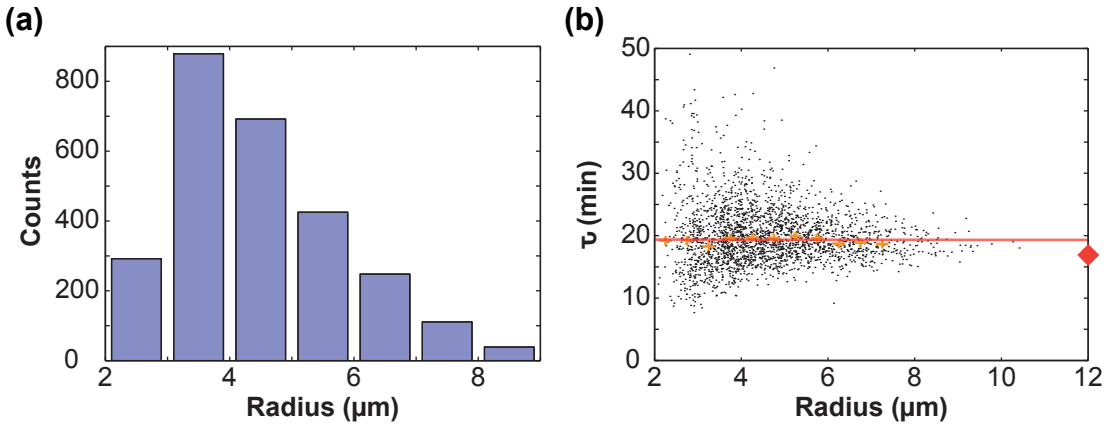

Figure 4.25: Analysis of RNA degradation subsystem in microdroplets. (a) Size distribution of droplets analysed. (b) Scatter plot of τ values for individual droplets and their corresponding radii derived by fits to single droplet traces of the exponential function 4.4. The median values are indicated as red crosses. Using the median values a fit function was derived which ought to describe the droplet size dependence of τ values. A red diamond indicates τ_{bulk} derived from an accompanying spectrometer measurement.

4.22). Third, we studied the influence of our compartmentalization method on simple subsystems containing only one type of the enzymes used in the oscillator reaction (Fig. 4.23 (a) and (b)). While it is difficult to make quantitative statements due to measurement noise, the results suggest several tendencies. Even though high variability in the rise times was observed for the transcription subsystem, no droplet trace showed faster transcription than the bulk indicating partly loss or deactivation of RNAP during the encapsulation process (Fig. 4.23 (a) and (c), Fig. 4.24). This effect is stronger for smaller droplets and the rise time approaches the bulk for larger. For the degradation subsystem (Fig. 4.23 (b) and (c)) loss of activity or material could not be related to droplet sizes but manifested as constant offset of the characteristic degradation time τ suggesting a global inactivation of RNase H compared to the bulk experiment. However, the variability in τ for individual droplets increased for smaller droplets and approached a constant value for larger ones which suggests slower degradation time than in bulk, hence slight deactivation of RNase H (Fig. 4.25).

Oscillations in microdroplets

In order to study the effect of compartmentalization into small volumes on the dynamical behaviour of the transcriptional oscillator described in section 4.2, we encapsulated the oscillator components within microemulsion droplets by the simple vortexing technique introduced in chapter 3.5.2. Due to the non-ionic nature of the surfactant, adsorption of molecules to the droplet boundaries was presumed to be very low. This was validated by previously discussed control fluorescence measurements on emulsions prepared solely with fluorescently labelled oligonucleotides and green fluorescent protein (Fig. 4.21). For initial experiments, we tuned the bulk oscillator into two dynamically distinct regimes by adjusting the RNase H concentration as given in table 4.3. In the first regime (identical to that shown in Figs.4.5 reaction #12, "sustained" oscillations with minimal damping were observed, while in the other regime with 25% higher RNase H concentrations the oscillator circuit produced "damped"

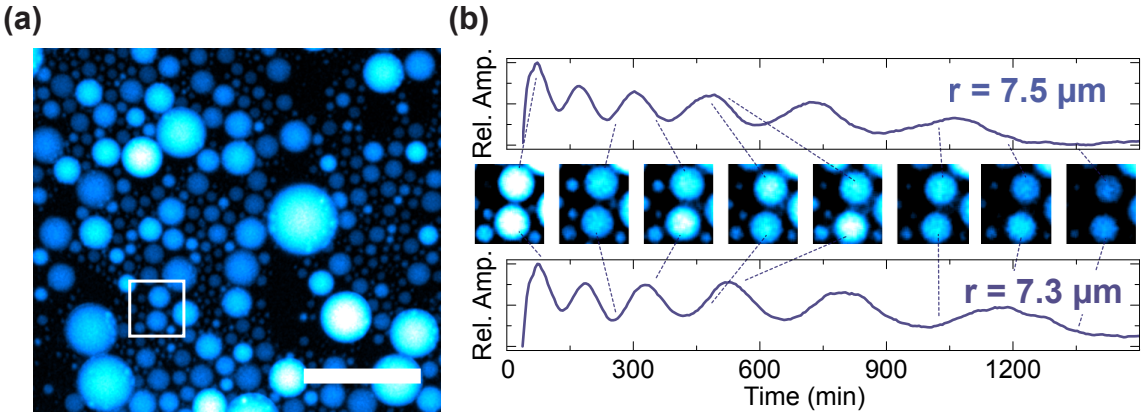

Figure 4.26: Biochemical oscillator in droplets: sustained. (a) False-colour fluorescence microscopy image of droplets containing the *in vitro* transcriptional oscillator. The scale bar is 100 µm. (b) Single droplet time traces of the droplets indicated by a white box in (a). Images of the time series are associated with the corresponding time points in the trajectories by dashed lines.

oscillations, indicating that for this choice of parameters the bulk system was closer to the non-oscillatory region of its phase space (Fig. 4.5 reaction #10). A fluorescence image of microdroplets encapsulating oscillators in the sustained regime is displayed in Fig. 4.26 (a). As indicated by the example traces in Fig. 4.26 (b), oscillations differ in amplitude and frequency from droplet to droplet. In order to obtain a better quantitative picture of the diverse dynamical behaviour within the droplet populations, we tracked individual droplets and evaluated their dynamics as a function of size (section 4.6.5). Bulk traces and corresponding model fits are shown in Fig. 4.11. In Fig. 4.27 (a), 20 example oscillatory traces obtained in the sustained regime are shown in "small" droplets with radii in the range $r = 2 - 5\,\mu\text{m}$ ($V = 33 - 524\,\text{fL}$). Lower panels in Fig. 4.27 represent the population average of the oscillatory signals together with the standard deviation around the mean value. Fig. 4.27 (b) contains example and average traces for the sustained oscillator for "large" droplets with radii $r > 8\,\mu\text{m}$($V > 2\,\text{pL}$). Fig. 4.27 (c) and (d) show the corresponding plots for the damped oscillator. For traces of medium sized droplets as well as example traces which were not classified "oscillating" see Appendix chapter 6.4. Of the traces that passed our filtering criteria (see section 4.6.5), only those with at least five clear extrema were defined as "identifiably oscillating". These may include both sustained oscillations and damped oscillations similar to the bulk traces in Fig. 4.27 (a) and (c). However, strongly damped traces and traces showing sustained but extremely slow oscillations may pass the filtering criteria but would not be deemed "identifiably oscillating". Considering the population mean of all identifiably oscillating traces and their standard deviation, we observe that larger droplets have smaller variation, for both oscillator tunings (Fig. 4.27, lower panels). In the sustained case, large droplets exhibit especially low variance at early times and the oscillating mean value indicates only slow desynchronisation. Smaller droplets have greater initial variance and desynchronise more quickly. The dephasing of the oscillations is less dramatic in the damped case (Fig. 4.27 (c) and (d)), where the standard deviation is considerably smaller than for the sustained oscillator. Due to remarkable diversity in oscillatory behaviour of the droplet traces, we reasoned that encapsulation might also induce

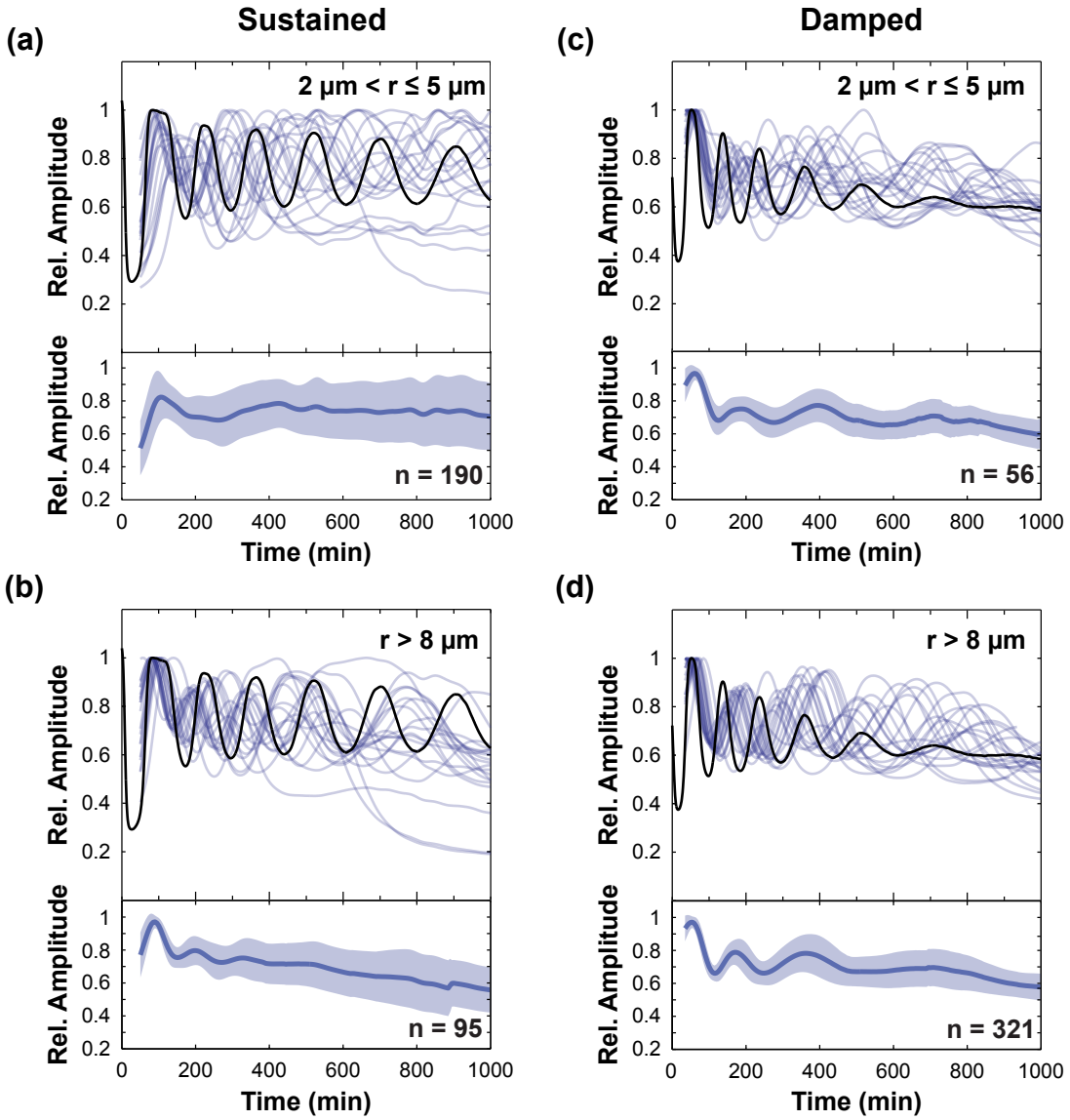

Figure 4.27: Example time traces of oscillating droplets (blue, upper panel) for sustained (a), (b) and damped tuning (c), (d) for smaller droplets with radii between 2 µm and 5 µm and larger ones with radii above 8 µm. The corresponding bulk trace from a spectrometer measurement is shown in black. Only time traces which were classified "identifiably oscillating" are shown as well as the population mean (blue trace, lower panel) of n droplets and their standard deviation (blue shaded area around the mean). For comparison purpose, traces were normalised to their maximum value at the first peak.

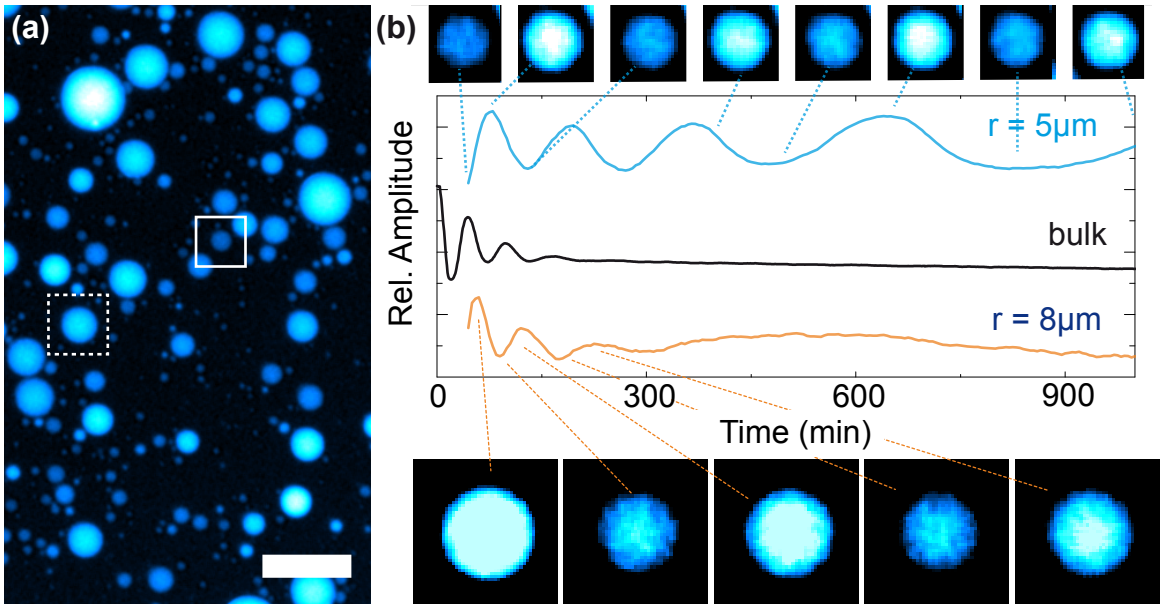

Figure 4.28: Strongly damped oscillator in microdroplets. (a) False-colour fluorescence microscopy image of droplets containing the *in vitro* transcriptional oscillator. The scale bar is 50 µm. (b) Fluorescence time traces for small (solid white box in (a), blue trace) and larger droplet (dashed white box in (a), orange trace) and bulk (black trace). Images of the time series are associated with the corresponding time points in the trajectories by dashed lines. The traces of bulk and larger droplet are classified by our analysis as 'not identifiably oscillating'. For some small droplets, sustained oscillations were observed which persisted for more than 15 h.

oscillations in a system that is barely oscillating in bulk. To test this hypothesis, we tuned the oscillator out of the oscillating region by an appropriate change in enzyme concentrations, resulting in strongly damped behaviour in bulk (Fig. 4.5 reaction #9). Indeed, in this case sustained oscillations were observed in some of the smaller droplets (Fig. 4.28). For traces identified as oscillating, in all three tunings, we determined the oscillation periods as a function of droplet radius r as shown in Fig. 4.29 (a), (c) and (e). The oscillations tended to slow down with smaller droplet size, while the variance of the periods increased. In the strongly damped case, identifiably oscillating droplets were observed only for $r \lesssim 10\,\mu\text{m}$. Our filtering criteria and measurement noise precluded identification of oscillations with periods faster than ≈ 80 min or those with low amplitudes. An analysis of the corresponding amplitudes showed very similar tendencies, with typically increasing amplitudes and amplitude variability for smaller droplets (Fig. 4.28 (b), (d) and (f)). Single droplet amplitudes and periods for all three tunings are shown in Appendix section 6.4.

As potential origins of these behaviours, we considered the effect of stochastic reaction dynamics, the influence of partitioning on the molecule numbers within the droplets, and a potential loss of enzyme activity during compartmentalization. In the following paragraphs, we discuss different scenarios and assess their plausibility by comparison of simulations with the before presented experimental results. We simulated the reactions using the extended model published by Kim and Winfree in [4] with parameters obtained by a least-square fit to data collected from bulk measurements of the three oscillator tunings which we also observed in droplets. The resulting parameters are within a realistic range (see table 6.1) and describe well the dynamics

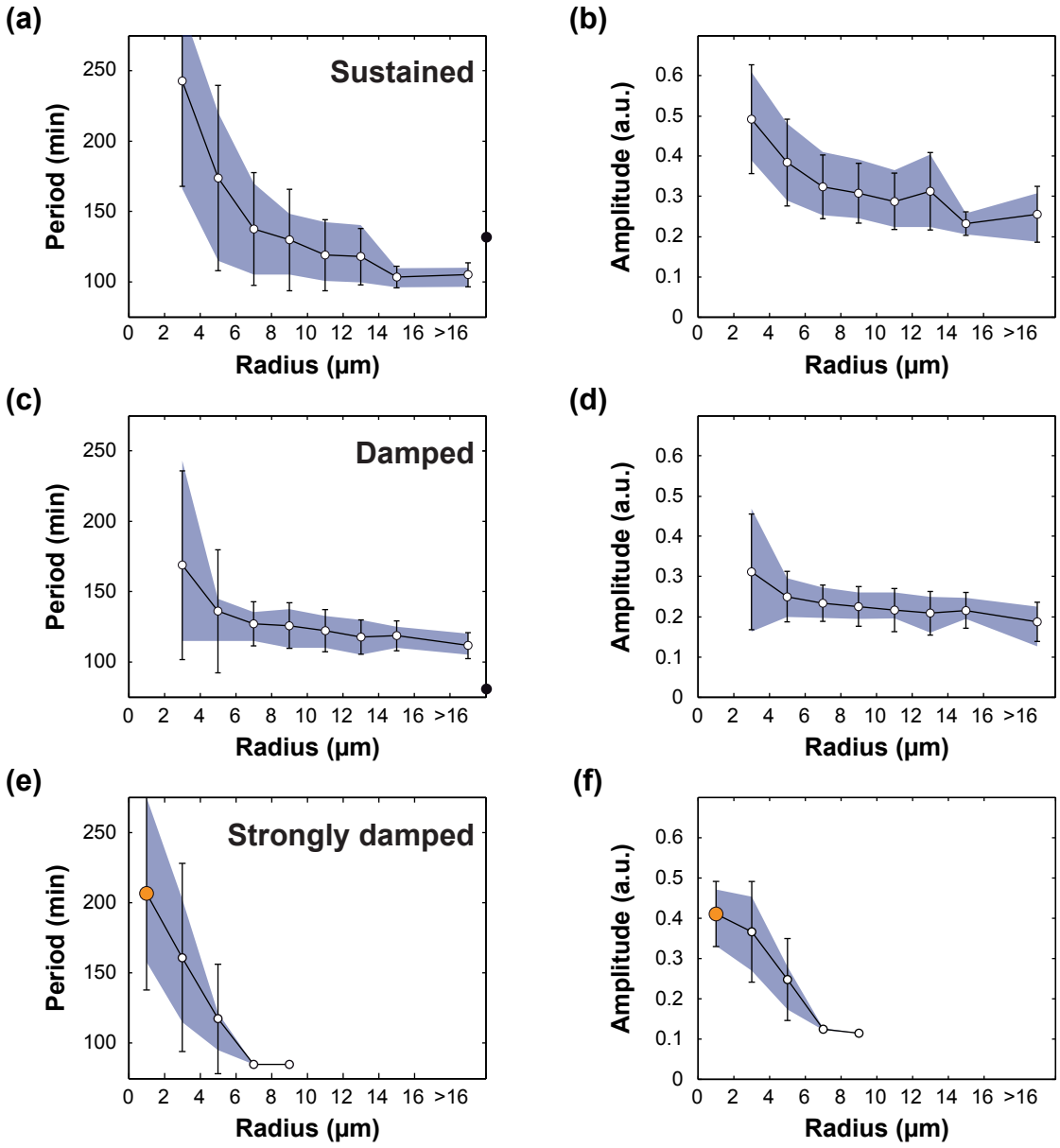

Figure 4.29: Variation of oscillation period and amplitude with droplet radius for the sustained, damped and strongly damped tuning. (a) - (c) Dependence of mean oscillation period on droplet radius. Bars indicate the standard deviation, the blue shaded area represents the range in which 60% of the population resides. The oscillation periods of the corresponding bulk trace are indicated as black dot at the right y-axis. (d) - (f) Oscillation amplitudes in dependence of the droplet radius. 1193, 978, and 644 identifiably oscillating traces were considered for sustained, damped and strongly damped tuning, respectively.

for the sustained, damped, and strongly damped tuning (Fig. 4.11). The enzyme concentrations for the 'sustained' tuning of the oscillator system are [RNaseH] ≈ 5 nM and [RNAP] ≈ 200 nM, and from 80 nM to 500 nM for the oligonucleotide components (table 4.3 and 6.2). In droplets with sub-picolitre volumes, some of the species are present at only relatively low copy numbers. For instance, in a droplet with radius $r = 2\,\mu m$ (corresponding to a volume of $V = 33\,fL$), there will be ≈ 100 molecules of RNase H, ≈ 4000 RNAPs, and the lowest DNA copy numbers will be ≈ 1600.

Stochastic reaction dynamics is expected to play a role for very small reaction volumes and molecule numbers. For instance, the oscillator species with the lowest concentration, RNase H, would be present at a copy number of only ≈ 5 for $V = 1\,fL$. Indeed, stochastic simulations of the dynamics of the transcriptional oscillator using Gillespie's algorithm [28] showed strong fluctuations for $V = 1\,fL$ (Fig. 4.30 (a)). Already at $r \approx 3\,\mu m$ ($V \approx 100\,fL$), however, these fluctuations were strongly reduced (Fig. 4.31 (b)). We analysed our results from stochastic simulations with the same procedure as it was done for droplet traces and plotted the mean period in dependence of the radius (Fig. 4.30 (c)) suggesting that for larger droplets the fluctuations are negligible, indicating that stochastic reaction dynamics is not sufficient to explain the diversity observed in our droplet populations. (Fig. 4.32 (b)).

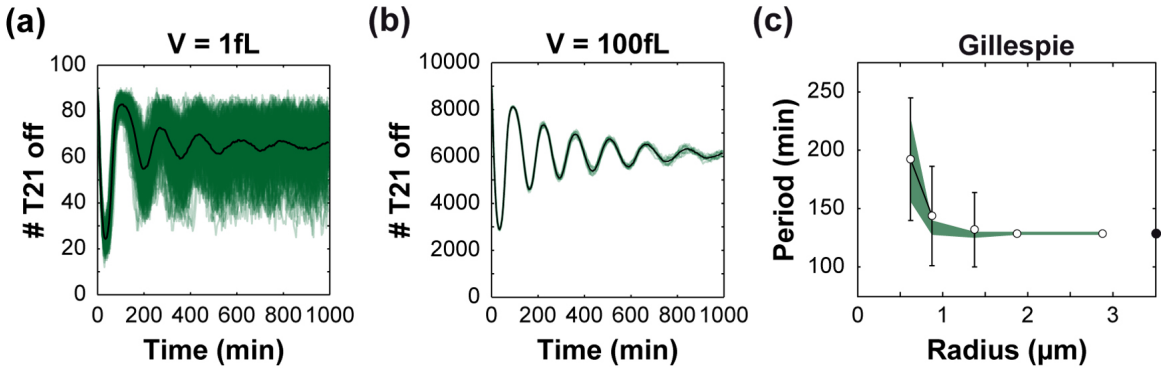

Figure 4.30: Simulations of the extended model using Gillespie's algorithm for the sustained oscillator tuning. Simulation of 200 time traces for $V = 1\,fL$ (a) and 10 time traces for $V = 100\,fL$ (b). Corresponding radii are ≈ 0.6 µm and ≈ 3 µm, respectively. The mean of all simulated traces is shown as black trace. Simulated traces were analysed identically to the experimental traces. (c) Mean oscillation period in dependence of droplet radii (white circles). Bars indicate the standard deviation, the shaded area represents the range in which 60% of the population resides. The bulk period is drawn as black dot at the right y-axis.

A major source of variability in our droplet populations may be **partitioning effects**. Small droplets of size V contain molecule numbers N that are expected to be Poisson distributed according to

$$p(N) = \frac{\lambda^N}{N!} e^{-\lambda}, \tag{4.5}$$

where the parameter $\lambda = c \cdot V \cdot N_A$ is the expected number of molecules for a bulk molar concentration c and N_A is Avogadro's constant. Poisson partitioning also predicts that the standard deviation of N will be $\sqrt{\lambda}$. While number fluctuations and corresponding concentration fluctuations are presumably small for most of the oscillator species (e.g. less than 10% for $\lambda > 100$), the combination of the variations of all 24

species concentrations, amplified by the oscillator circuit, may well lead to considerable variability in the dynamics of the system. In order to assess the influence of such partitioning effects, we performed simulations for varying droplet sizes using deterministic mass-action differential equations, for which the initial molecule numbers of all oscillator species were drawn from a Poisson distribution as exemplary discussed for simple transcription in section 2.1.6. The results of such simulations for a droplet radius of $r = 2\,\mu\text{m}$ are shown in Fig. 4.31 (b). The simulations indicate that partitioning can indeed result in a qualitatively similar variability in dynamical behaviour as that observed experimentally. As shown in Fig. 4.32 (a), the mean period of the oscillations as well as its variance increases with decreasing droplet radius. Quantitatively, however, the effect of Poisson partitioning is significantly weaker than in the experiment, where strong variability is already observed for relatively large droplets. This strong variability could be explained by several factors, such as a small fraction

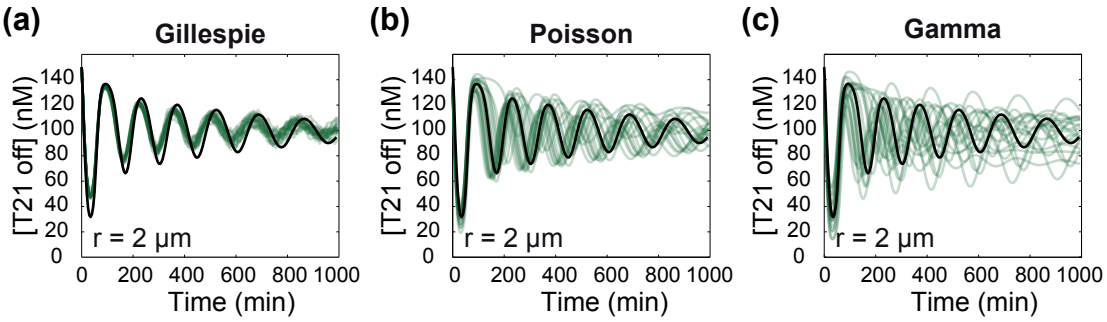

Figure 4.31: Simulations of the sustained oscillator considering different stochastic effects for droplets with $V = 2\,\mu\text{m}$. (a) Time traces simulated using Gillespie's algorithm for stochastic reaction dynamics. (b) and (c) Time traces simulated deterministically but with initial molecule numbers drawn by a Poisson and Gamma distribution, respectively. For the Gamma distribution a shape factor of $\beta = 10$ was used.

of enzymes being active, incomplete mixing of the solution, protein multimerization or aggregation that results in copy number variance greater than the mean [176], or stochastic protein inactivation due to adsorption on the droplet interface and denaturation during generation of the microdroplets [188]. In principle, determination of concentrations within individual droplets could directly assess whether partitioning was Poisson or broader, but our fluorescence measurements did not provide concentration values with sufficient accuracy (Fig. 4.22). Instead, we examined the variability in two 'single-enzyme subsystems' (Fig. 4.23). In the degradation subsystem, the enzyme turnover of ≈ 100 effectively amplifies the signal resulting from partitioning. In both cases, apparent variability in enzyme activity was an order of magnitude larger than would be expected from Poisson partitioning combined with measurement noise (for further details see supporting information supplemented with [64]). Furthermore, independent of the droplet radius the median of RNase H activity seemed to be slightly reduced compared to the bulk case (Fig. 4.25). By contrast, RNAP activity was considerably lower for droplets with smaller radii (Fig. 4.24 (b), (c)), which can be qualitatively explained by a loss of enzyme function during the droplet generation process. This may be caused by enzyme aggregation, which we also observed for fluorescently labelled T7 RNAP encapsulated into the emulsion droplets (Fig. 4.21) [189]. We empirically modelled the larger variability by a Gamma distribution, whose

probability density function for $x \geq 0$ is given by

$$p(x;\alpha,\beta) = [(\beta^{\alpha}\Gamma(\alpha))]^{-1}x^{(\alpha-1)}e^{(-x/\beta)}, \tag{4.6}$$

where $\Gamma(\alpha)$ is the Gamma function, α is the shape factor, and β is the scale factor. This distribution allowed us to independently choose the mean $\langle x\rangle = \alpha\beta$ and variance $var(x) = \langle x\rangle\beta$ of the distribution. The Gamma distribution was previously used in the context of stochastic gene expression, when gene expression bursts were considered together with the influence of cell division and degradation processes [94]. It has been noted, however, that a variety of different scenarios for cellular heterogeneity lead to essentially the same, Gamma-like molecular number distributions [91]. In the revised simulations, the initial molecule numbers for each species in a droplet were drawn from a Gamma distribution with mean $\alpha\beta = c\cdot V\cdot N_A$ and scale factors $\beta = 10$ or $\beta = 100$, which span the range of β-values suggested by the single-enzyme experiments. As shown in Fig. 4.31 (c) and Fig. 4.32 (a), the similarity of simulation and experimental results are much closer. The model qualitatively captures the change of size-dependent variability with changed oscillator tuning: the sustained tuning has a wider range of radii exhibiting significant variability than the damped tuning, the strongly damped tuning yields oscillations only for smaller droplets, and in all three tunings the mean period increases slightly for very small droplets. As also shown in Fig. 4.32 (a), simulations which take the observed enzyme loss into account capture this trend towards larger periods for small droplet radii even better. If we assume that variations in active enzyme concentrations are the dominant sources of variability in our experiments, we can semi-quantitatively understand the experimentally observed trends by regarding two-dimensional 'phase diagrams'. In Fig. 4.32 (b), the oscillator period, amplitude and damping coefficients predicted from our numerical model are shown as a function of [RNaseH] and [RNAP]. As indicated in the plots by ellipses around the bulk operating points, depending on the position of the initial conditions in phase space and also on the width of their distribution, a wide range of oscillations with different periods or amplitudes, and even non-oscillatory dynamics are sampled. This also explains the different behaviour observed for the sustained, damped, and strongly damped oscillator tunings. Whereas variations towards smaller enzyme concentrations can drive the sustained oscillator into a phase space region with large periods, this effect is expected to be less pronounced for the damped oscillator, whose (bulk) initial conditions happen to be in a less sensitive region of phase space. Furthermore, Fig. 4.32 (b) also shows that fluctuations can drive the strongly damped oscillator into the oscillatory region. In addition, an overall loss of enzymes shifts the bulk initial conditions towards lower values as indicated by white arrows.

In our study of a compartmentalized *in vitro* transcriptional oscillator we have found significant deviations from the bulk system and a broad diversity in dynamical behaviour. The dynamics of the encapsulated system can be understood as the result of a broad distribution in initial concentrations amplified by the underlying dynamical system, and can be adequately reproduced in deterministic simulations with statistically varying initial conditions. We found that for the relatively large molecule numbers studied here, the inherent stochasticity of chemical reactions cannot account for the variability, according to our model. Adopting the language of stochastic gene expression, our system appears to be dominated by 'extrinsic' sources of noise such as enzyme number variability and partitioning effects [91, 187] rather than 'intrinsic'

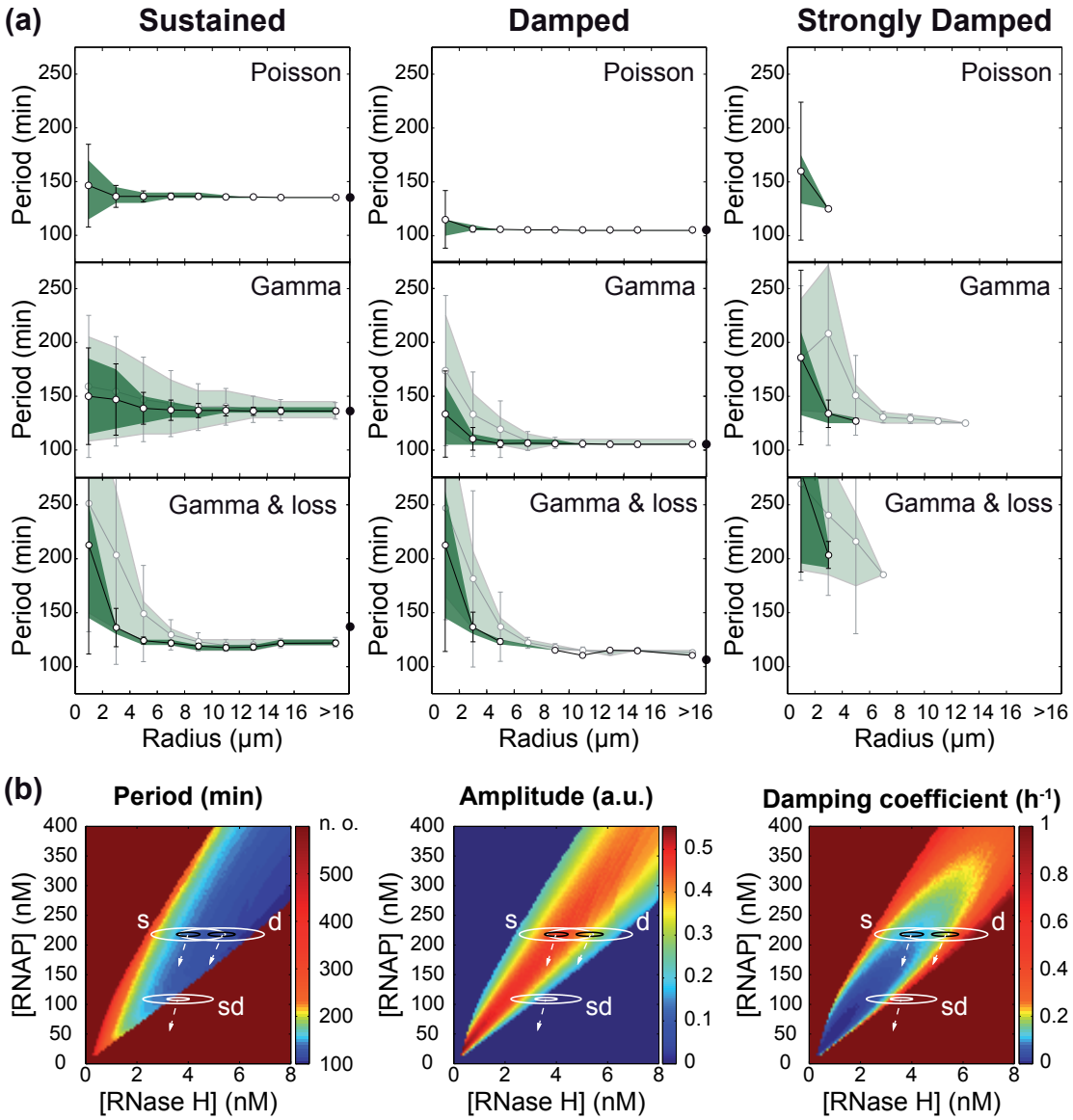

Figure 4.32: Analysis of simulations with partitioning error of initial molecule numbers. (a) Distributions of periods as function of droplet radii for simulations of the sustained (left), damped (centre) and strongly damped (right) oscillator tuning. Simulations were performed for initial molecule numbers following a Poisson distribution (top row), a Gamma distribution (centre row) for $\beta = 10$ (dark green) and $\beta = 100$ (green shaded area), and for Gamma distributions for $\beta = 10$ (dark green) and $\beta = 100$ (green shaded area) with loss of enzyme activity during droplet generation. The mean periods are drawn as white circles, the standard deviations as bars and the green areas represent the region in which 60% of the population resides. (b) Phase diagrams generated by analysis of deterministic simulations for period, amplitude and damping coefficient in dependence on RNAP and RNase H concentrations. The conditions for sustained (s), damped (d) and strongly damped (sd) oscillator tunings are in the centre of ellipses which represent the expected standard deviations due to initial enzyme numbers drawn by a Poisson (small) and a Gamma distribution with $\beta = 10$ (large) for droplets with $r = 2\,\mu\text{m}$. Dashed arrows indicate the effect of enzyme loss on the operation point.

noise caused by stochastic reaction dynamics [93, 33]. Our experiments and numerical analysis suggest that physical sources of variability include stochastic partitioning of molecules and loss of enzyme activity during the droplet production process. These

phenomena yield broader distributions of molecule numbers and thus in dynamic variability among droplet oscillations, which is qualitatively captured by our Gamma partitioning model. Notably, how the variability in initial conditions is propagated through the dynamics, is profoundly affected by the operating point of the oscillator. It has been previously argued for biological cells that the influence of partitioning noise may be under-appreciated compared to stochastic reaction dynamics [91, 187]. Our study suggests that similar 'partitioning effects' play a major role in determining the dynamic behaviour of compartmentalized artificial biochemical circuits: these effects pose a major challenge for cell-scale molecular engineering and the realisation of protocells. It therefore will be important to investigate which synthetic circuits provide better or worse performance in the presence of such noise, and to study and potentially adopt strategies of biological cells for either active noise reduction or increased resilience [183, 44, 184]. An interesting application of compartmentalization-induced variability could be the deliberate generation of widely varying initial concentrations for the exploration and characterisation of the dynamical potential of other non-linear biochemical networks. If the initial concentrations of key species could be directly measured in each droplet, then the resulting behaviour of dynamical species could be directly associated with the initial concentrations thus mapping the phase diagram in a single experiment. Taking our oscillator as an example, if RNAP and RNase H could be individually labelled [190], and if measurement signal-to-noise could be improved, then microscopy of droplet microemulsions could provide a high throughput experimental measurement of phase diagrams analogous to those of Fig. 4.32 (b), as a function of droplet radius. As our programmable biochemical systems can be used to design complex chemical networks with a well-defined topology, an exciting opportunity for quantitative analysis of such networks would be to use the whole range of observed dynamical behaviours for parameter estimation, for example by fitting each trace obtained from an experiment with the same set of rate parameters, but with differing initial conditions. This collective data fitting approach is argued to yield models with more predictive power than if each parameter is individually measured with similar accuracy [191, 192], but is limited by the availability of extensive data sets. Such high-throughput methods for characterising synthetic circuit behaviours - potentially supported by microfluidic liquid handling techniques [177] - are likely to prove essential for the development of sophisticated but stochastic cell-scale artificial molecular systems.

4.5 Conclusion and outlook

Here we have used a biochemical synthetic circuit - a transcriptional oscillator - to coordinate molecular tasks in a timed fashion, and to study effects which arise when systems are scaled down in size. Using the previously demonstrated concept of "insulators", we successfully coupled different DNA nanodevices simultaneously to the oscillator's dynamics and controlled the fluorescence of dye molecules in solution by alternating transcription of fluorescence enhancing aptamers and anti-aptamers. In principle, "insulators" could be designed which encode for different structural units of larger assemblies. This circuit could then be used to produce different units in different phases and therefore establish a sequential assembly in a one-pot-reaction

in a assembly-line-like manner. After the oscillator was tuned to an operating point at which reproducibly sustained oscillations persisted for more than 30 hours we used the oscillator circuit to study effects which appear when systems are scaled down in size. Down-scaling was achieved by encapsulation of the circuit into microemulsion droplets by using microfluidics and a simple vortexing technique, the shaken-not-stirred method. Already in droplet oscillators produced in microfluidics we found considerable droplet-to-droplet variations in the circuit's dynamics. However, we focussed on the shaken-not-stirred method since both smaller droplet sizes as well as broader size distributions could be generated in minimal amount of time. Here, strong variations in the dynamic behaviour of the population of encapsulated oscillators occurred suggesting, as expected, that small ensembles of molecules behave differently than large ensembles in bulk. However, the degree of diversity was far beyond the initially anticipated considering either biochemical noise due to small molecular numbers or errors during partitioning, thus droplets with initial molecule numbers drawn by a Poisson distribution. We could explain the diversity in behaviours assuming initial molecule numbers according to a Gamma distribution which had been used before in the context of beyond-Poisson biochemical noise in gene expression. To this end, our study indicates which challenges have to be overcome by (bio-)engineers when they attempt to get things to work on the micron scale in a deterministic fashion.

In the sense of Richard Feynman's quote "What I cannot create I do not under-

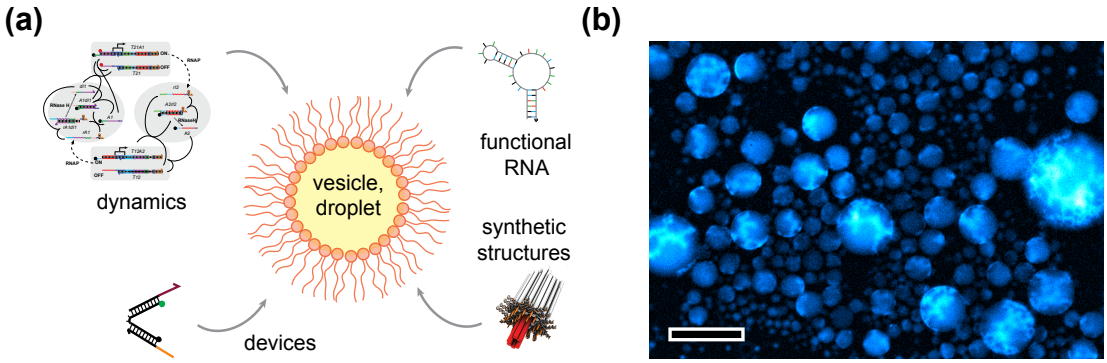

Figure 4.33: Synthetic compartmentalized systems. (a) In future, synthetic components may be used to generate dynamical microscale systems with cell-like properties. Lipid vesicles or microdroplets could be equipped with dynamical systems based on transcription circuits which drive a variety of downstream processes. By coupling RNA aptamers or riboswitches to circuit dynamics the system could respond on small ligand molecules. Integrating DNA-nanostructures, like transmembrane channels could be further used to establish communication channels between neighbouring microreactors or provide supply of "fuel" [193]. DNA nanodevices could be coupled to fulfil tasks on a molecular scale like the generation of larger structures in an assembly-line-like fashion. (b) Fluorescence microscopy image of DsrA RNA molecules, which are known to assemble into larger 2- and 3-dimensional structures, transcribed in droplets [194]. DsrA is a 87-nt non-coding RNA which was found in *E. coli*. The scale bar is 50 nm. RNA structures are visualised by an intercalating dye (SYBR Green II) present in the reaction buffer.

stand", synthetic systems comprised of well-known components may be used to unravel and understand mechanisms of natural systems, like noise robustness. DNA based components, as for instance nanodevices or larger structures could be combined with transcription circuits to create systems with increasing complexity with cell-like capabilities. Lipid vesicles or surfactant stabilised droplets seem to be per-

fect model systems to combine molecular components which are becoming available. In Fig. 4.33 (a), a scheme for an imaginary assembly of such a system comprised of already existing components is shown. Droplet or vesicle microreactors could serve as compartment for dynamical circuits that act as master clock for the coordination of action of nanodevices or transcriptional activity for "genes" which encode for functional RNAs. As it is known from riboswitches [195], RNA molecules could be designed that change their conformation upon binding of a small ligand molecule. Integration of such riboswitch-like molecules into circuits like the transcriptional oscillator may be used to achieve synchronisation by communication through ligands that can change between droplets, as auto-inducers cross cell-walls in bacterial quorum sensing [196]. Larger DNA structures like a previously reported transmembrane channel [193] could be added to the system to achieve communication between droplets when integrated in a lipid double layer vesicle-vesicle-boundary or to enhance circuit life-time by draining waste products into and providing "fuel" from the surrounding medium. RNA molecules which assemble into two- or three-dimensional structures, like DsrA could further be synthesised *in situ* to give microreactors a cytoskeleton-like structure [194, 173]. Fig. 4.33 (b) shows a fluorescence microscopy snapshot of DsrA transcribed in droplets.
Although research on *in vitro* circuits has only just begun achievements in the field indicate their potential to emulate and explore the complexity of biological circuit design. Furthermore, in the future they may be used as control circuits in artificial cells or as model system for biomimetic pattern generation in spatially extended systems.

4.6 Materials and methods

4.6.1 Materials

	T7 RNAP	RNase H	Ppase
Vendor	Epicentre	Ambion	Sigma Aldrich
Art.-Nr.	TM910K	AM2292	I1891-100UN
Molecular weight (kDa)	99	17.6	71
Specific activity $^{\text{units}}/_{\mu\text{g}}$	400/560	$500 - 1000$	1
Optimal temperature (°C)	37	37	25
Stock concentration in $^{\text{units}}/_{\mu\text{L}}$	200	10	1000
Stock concentration in µM	5.05/3.60	≈ 0.646	0.012

Table 4.1: General information on enzymes for T7 RNA polymerase, *E. coli* RNase H, and inorganic Pyrophosphytase (Ppase). For RNAP, the specific activity changed during the project. Information about which experiment was performed with which RNAP batch see Appendix 6.3. Since the vendor refused to disclose the specific activity for RNase H, we estimated the stock concentration based on batches for which the information was available (see Appendix 6.3).

Cuvettes for measurements in the fluorescence spectrometer were cleaned by thoroughly rinsing alternating with nano pure water and isopropanol, ending with water. Thereafter, cuvettes were incubated in a cleaning agent for Quartz cuvettes, 2% Hellmanex III (Hellma Analytics (Art.-Nr. 320.003)) for at least 3 h at room temperature.

Reagent	Concentration
Tris-HCl (pH 7.5)	50 mM
NaCl	100 mM
DTT	1 mM
EDTA	0.1 mM
Glycerol	50%
Triton X-100	0, 1%

Table 4.2: Content of RNAP storage buffer. RNAP was usually added to the oscillator samples in volume ratios of $\approx 5\%$. Storage buffers typically contain salt and agents to avoid freezing like glycerol which might have influence on the reaction dynamics. For instance RNAP is inhibited when the total salt concentration exceeds 50 mM. Glycerol has a high viscosity and may change diffusion properties. Since the volumes added to the samples were roughly constant throughout all experiments, it is not expected that considerations were drawn due to misinterpretations caused by changes in the buffer conditions. However, if for any reason enzyme volumes are changed significantly, these facts has to be considered.

Hellmanex was removed by thoroughly rinsing with nano pure water. Before preheating in the spectrometer, the cuvettes were dried using nitrogen. This procedure removed sediments in the cuvettes' reaction cavity which were even resistant against 5% HCl.

4.6.2 Circuit preparation

Preparations for transcription circuits and all projects which could be sensitive to RNase contamination were done at a bench which was cleaned before and after each usage with a special cleaning agent against RNase contamination (RNaseZap, life Technologies, Art.-Nr. AM9780).

Preparation of reaction components

In advance to experiments with transcription circuits discussed in this chapter some reagents have to be specially prepared.
DNA strands purchased from biomers.net were usually ordered dissolved in nuclease free water (except Texas Red labeled T21 nt) at 100 µM concentration as specified by the vendor which turned out to be a rough reference point with an approximately 15% uncertainty. T21 nt from biomers.net and all strands from IDT were ordered as dry salt. According to the vendors mass content information, the lyophilised powder was dissolved in nuclease free water (IDT) to yield an target stock concentration of 100 µM. The dissolved DNA strands were quantified by measuring the absorbance at 260 nm on a nanophotometer (Implen). Since accuracy in concentration is quite important for the oscillator components at least five measurements for each strand were performed where blanks were reset before each measurement using nuclease free water. The concentrations were then calculated using the extinction coefficients determined with IDT's online tool "Oligo Analyzer" (https://eu.idtdna.com/analyzer/Applications/OligoAnalyzer/). As nice feature this tool corrects for additional extinction of dye or quencher modifications. Single strands were typically further diluted with nuclease free water (IDT) to 10 µM working so-

lutions. DNA species which are present in the reactions as double strands such as genelets, tweezers or reporter systems (except the RNA:DNA complex for the degradation system shown in Fig. 4.23 (b)) were annealed in advance in 500 µL PCR clean DNA LoBind tubes (Eppendorf, Cat.-Nr. 0030108035). For this purpose, the corresponding single strands were mixed in 1x NEB RNAPol reaction buffer (B9012S). Genelet and reporter strands were mixed in stoichiometry to a final stock concentration of 5 µM. For DNA tweezers the arm strands were added in 5% excess in respect to the base strand which was prepared for a target stock concentration of 2.5 µM. The mixed DNA single strands were annealed on a Thermocycler (Eppendorf) using a protocol which first heated up to 95°C, remained there for 5 min before the sample was slowly cooled down to 20 °C over 80 min.
Following the suggested method to extend the lifetime of the transcription reaction [197], we also included $\approx$ 0.001 $^{\text{units}}/_{\mu\text{L}}$ of inorganic pyrophosphatase (Sigma Aldrich, I1891-100UN) in the reaction mixture. **The inorganic Pyrophosphatase** shipped as lyophilised powder was disolved in 20 mM Tris-HCl (pH 7.2 at 25 °C), 1 mM $MgCl_2$ and 50 % glycerol (v/v) yielding a 100 $^{\mu\text{g}}/_{\text{mL}}$ stock solution. According to the vendor, this concentration corresponds to 100x of the optimal pyrophosphatase concentration. Based on the vendor's reported activity $\geq$ 1000 $^{\text{units}}/_{\text{mg}}$ and a molecular weight of 71 kDa, the bulk pyrophosphatase molar concentration is then $\approx$ 12 nM. As a side note: since pyrophosphatase is involved in enhancing RNA transcription and is not directly involved in the dynamics, we do not call it an 'essential enzyme' for the circuit dynamics and neglect this enzyme in our models.

4.6.3 Preparation of reaction mix

For bulk measurements the reaction mix was prepared in 500 µL PCR clean Protein LoBind tubes (Eppendorf, Cat.-Nr. 0030108035). It contained 1x NEB RNAPol reaction buffer (B9012S), additional 24 mM $MgCl_2$ (6 mM are already part of the TX buffer), 7.5 mM each dNTP (Epicentre, RN02825), 1x inorganic Pyrophosphatase (Sigma Aldrich, I1891-100UN, dissolved as described above), and DNA species (cf. table 4.3 for droplet oscillator, for a complete list see Appendix section 6.3). The prepared final volume was adjusted according to its purpose. For only bulk measurements typically 70 µL reaction solution was prepared per cuvette. The reaction mix volume for experiments in shaken-not-stirred droplets was 120 µL and for microfluidic droplets 300 − 400 µL. **For droplet experiments** 500 nM Alexa 488 was added as reference dye which had no influence on the dynamics of the transcription circuits. Except for microfluidic droplet experiments, RNAP (Epicentre, TM910K) and RNase H (Ambion, AM2292, 10 $^{\text{units}}/_{\mu\text{L}}$) were added later in the cuvette after a baseline had been recorded. After assembly, the mix was vortexed thoroughly and incubated for 10 min before it was loaded to the spectrometer. After preparation of the reaction mix, it was added to pre-heated Hellma ultra-micro cell Quartz cuvettes (105.252-QS) without RNAP and RNase H to record the maximum fluorescence intensity value, which corresponds to the maximum concentration of genelet *T21* in the transcriptional off-state. After 10 min of equilibration time, first RNase H and finally RNAP were added to the cuvettes and mixed by vigorously pipetting. For shaken-not-stirred droplet generation, 4 aliquots of 10 µL were extracted and added to Protein LoBind tubes (Eppendorf, Cat.-Nr. 0030108035) filled with 45 µL oil-

surfactant mix (section 3.5.2) before the sample in the cuvette was finally sealed with 35 µL hexadecane (Merk, 8.20633.1000) to prevent evaporation.

Table 4.3: Enzyme concentrations for sustained (s), damped (d) and strongly damped (sd) oscillator tuning discussed in section 4.4.2. DNA concentrations are the same for all conditions, while enzyme concentrations were tuned.

Species	damped (reaction #10)	sustained (reaction #12)	strongly damped (reaction #9)
T21 (nM)	150	150	150
A1 (nM)	150	150	150
dI1 (nM)	500	500	500
T12 (nM)	80	80	80
A2 (nM)	250	250	250
RNAP (nM)	218	218	109
RNase H (nM)	5	4	3.3

4.6.4 Malachite green and DFHBI RNA aptamers

In section 4.3, RNA aptamers were used to enhance fluorescence of dyes present in solution. RNA aptamers are RNA molecules with a secondary structure with a loop to which particular molecules bind specifically. Aptamers which switch on fluorescence of dyes are useful for studying processes without the requirement for specific chemical modification. Malachite green (MG) binds specifically to its aptamer with $K_d = 117\,\text{nM}$. The binding to the aptamer enhances MG's fluorescence by a factor of ≈ 2360, from a quantum yield of $7.9 \cdot 10^{-5}$ to 0.187 [171]. However, due to its toxicity MG is not suited for *in vivo* measurements [198, 199]. Furthermore, its slight hydrophobicity ($K_{ow} = 0.62$ [200]) makes studies in aqueous solutions difficult, especially when solutions are prevented against evaporation by a layer of oil as in our case. However, with high malachite green concentrations and thin layers of oil, the oil phase saturates and measurements are possible. Fig. 4.34 (a) shows the chemical structure of malachite green as well as the RNA aptamer's secondary structure at 37 °C predicted by NUPACK. Rather recently, further dyes were reported for which fluorescence enhancing RNA aptamers exist [172]. One of these dyes, DFHBI (also called Spinach due to its green fluorescence), is commercially available. In contrast to MG, DFHBI is not toxic and has been also used to monitor RNA in live cells and in *in vitro* expression experiments [201, 202]. DFHBI binds to its corresponding aptamer with $K_D = 537\,\text{nM}$. By binding to the RNA aptamer, its quantum yield increases from $7 \cdot 10^{-4}$ to 0.72 [172]. The chemical structure of DFHBI as well as the RNA aptamer's secondary structure are shown in Fig. 4.34 (b).

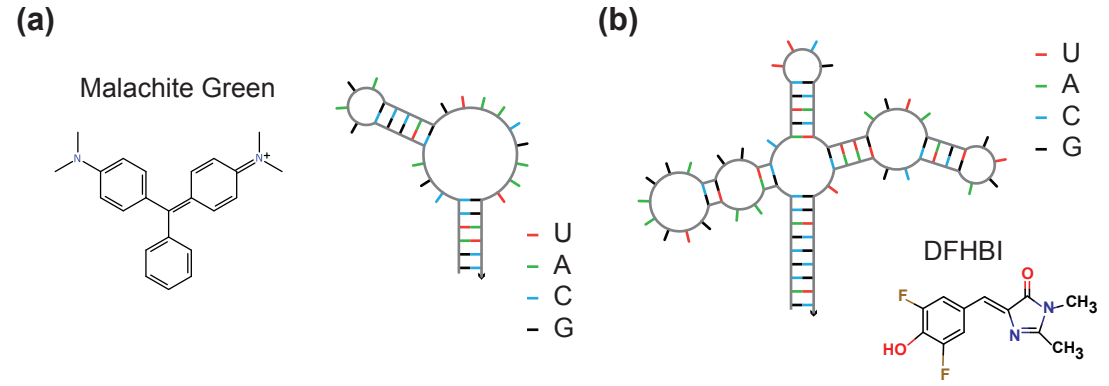

Figure 4.34: Malachite green and DFHBI dye and aptamer. (a) Chemical structure of malachite green dye and the RNA aptamer's secondary structure at 37 °C. (b) Chemical structure for DFHBI and the RNA aptamer's secondary structure at 37 °C. RNA structures were analysed with NUPACK.

4.6.5 Methods

Fluorescence spectrometer measurements and data processing

Bulk fluorescence experiments were performed on a Horiba/Jobin Yvon Fluorolog 3 system in 45 µL Hellma ultra-micro cell cuvettes (105.254-QS). Fluorescence spectroscopy was introduced in section 3.2. Fluorescence emission from labelled DNA strands was recorded every minute. A sample temperature of 37 °C was maintained using a water circulation thermostat. The monochromators were set to excitation/ emission wavelengths:

Dye	λ_{ex}	λ_{em}
DFHBI/ Alexa 488	495 nm	519 nm
Rhodamine Green	504 nm	532 nm
TAMRA	546 nm	570 nm
Texas Red	598 nm	617 nm
Malachite Green	630 nm	655 nm

with slit widths typically set to 5 nm for both excitation and emission. The data obtained from spectrometer measurements was processed and analysed using Origin. Depending on the excitation lamp's age fluorescence time traces showed some noise. This noise was reduced by applying Origin's build in 'Adjacent Averaging' smoothing procedure typically over 5 data points. As its name suggests 'Adjacent Averaging' over n points smoothes data point at index i using n data points (where n is an odd number). The original data point n is then replaced by the smoothed value at index i which is the average of the data interval $[i-(n-1)/2, i+(n-1)/2]$. Before smoothing was applied to traces it was assured by a preview that no features of the oscillator dynamics such as height of peaks or depth of wells were obscured by this procedure. The smoothed data's baseline was then volume corrected depending on the amount of enzymes added to the cuvette assuming that fluorescence intensity depends linearly on concentration of *T21*'s dye label. Smoothed, and volume corrected data was then

normalised and converted to concentrations using

$$C(t) = C_0 \frac{I(t) - I_{min}}{I_{max} - I_{min}}, \tag{4.7}$$

with the time dependent concentration of *T21* OFF, $C(t)$, the total concentration of genelet *T21*, C_0, the measured intensity value $I(t)$, the maximum fluorescence determined by the baseline recording before addition of enzymes when all genelets *T21* are OFF I_{max}, and the fluorescence intensity when all genelets *T21* are ON, I_{min}. As determined in reference measurements $I_{min} = 0.05 \cdot I_{max}$. For better comparison of traces for oscillator samples with different total concentrations of *T21*, the data shown here was just normalised to 1. To convert the shown relative concentration values for *T21* OFF into absolute concentration, the relative concentration has to be multiplied by the total concentration of *T21* in each experiment.

For fluorescence trajectories shown in Fig. 4.17, no scaling factor has been determined so far. Typically, both malachite green as well as DFHBI in absence of the aptamer have very bad quantum yields. For visualisation purposes, the fluorescence data was normalised to their initial values for DFHBI, hence $I_{norm}(t) = I(t)/I_0$ with the normalised intensity I_{norm}, the measured time dependent intensity $I(t)$ and the initial intensity value I_0, which gives a measure of 'fold' increase of fluorescence in respect to the initial state. However, calibration measurements have to be done for later parameter determination by model fits. Malachite green data had to be corrected for cross-talk from the Texas Red label. Reference experiments showed that 0.74% of the Texas Red signal is visible at the malachite green emission wavelength. Even though this factor appears rather negligible, its contribution to the malachite green signal in our case is significant due to the high difference in signal strength between both dyes. The malachite green raw data was therefore cross-talk corrected by

$$I_{cor}(t) = I_{MG}(t) - 7.4 \cdot 10^{-3} \cdot I_{TexRed}(t) \tag{4.8}$$

where$I_{cor}(t)$ is the cross-talk corrected malachite green intensity, $I_{MG}(t)$ is the malachite green raw data, and I_{TexRed} is the Texas Red raw data recorded for the Texas Red emission wavelength. A result of cross-talk correction is shown in Fig. 4.35. After cross-talk correction, the signal goes back to zero intensity as usual for malachite green in absence of functional RNA aptamer indicating that the anti-aptamer efficiently turned off fluorescence. Since after cross-talk correction, slightly negative values also occurred, a normalisation as for DFHBI was not possible. Instead, the malachite green data set was simply normalised to its maximum and minimum value which enabled us at least to compare the dynamics expressed by the different dyes as shown in Fig. 4.17.

Time-lapse microscopy

Time-lapse microscopy measurements were conducted on an Olympus IX81 epifluorescence microscope equipped with an automated x-y-stage (Prior Scientific, Cambridge, UK), an 120 W mercury vapour short arc lamp (Lumen dynamics, X-Cite 120Q), and an incubator box (Okolab) which maintained the operation temperature of 37°C. The setup was controlled with MicroManager 1.4. As sample carrier, sterile hydrophobic ibidi µ−Slides VI$^{0.4}$ (Art.-Nr. 80601) with a channel height of 400 µm were used.

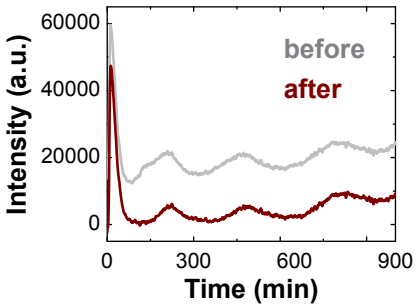

Figure 4.35: Fluorescence traces for malachite green before (grey) and after (brown) cross-talk correction according to 4.8. Even though the dynamics qualitatively seem not to change, the offset as well as a shoulder at the left flank of the first maximum in the not corrected trace show the influence of Texas Red emission at the emission wavelength of malachite green.

The channel inlets were sealed using PCR tape (Biorad) to avoid evaporation. Every 5 minutes, a bright field and fluorescence images in the GFP channel for the reference dye (Alexa 488) and in the Texas Red channel for the *T21* dye label were recorded for at least 1000 minutes. For each observation cycle, a clean position in the centre of a channel on the ibidi slide which was only filled with the oil-surfactant mix was recorded for later correction of the oil's auto-fluorescence. Droplet movies were typically recorded in 10x magnification.

Image analysis and data processing

Time-lapse movies were analysed using the ImageJ plugin Cell Evaluator [203]. Droplets were first segmented in bright-field for each time point and then tracked. For segmentation of objects (droplets), complexity of the grey value original images is reduced by applying the threshold determination method 'Otsu' which is part of the imageJ plugin MultiThresholder written by Kevin Baler, Gabrial Landini and Wayne Rasband. After applying the threshold, the image is converted into a binary image. The tracking software returned a data matrix containing information of interest for each detected object such as centre coordinates, area, perimeter, mean square displacement, and the particular shape. The fluorescence image time series of droplets were background corrected by subtraction of the image time series recorded from a position in the oil channel, pixel by pixel for each time point. The results from segmentation and tracking were then merged with the background corrected fluorescence time series for either Alexa 488 and Texas Red channel which returned the information about the fluorescence of every droplet for each time point. Droplet radii were calculated assuming a circular shape of the segmented objects. The scaling from pixel to µm was 1pixel = 0.641 µm in 10x magnification. The radius then is given by $r = \sqrt{A/\pi} \cdot 0.641\,\mu\text{m}$ with droplet's radius r in µm and area A in pixel.
Since our fluorescence microscopy measurements were rather noisy, a variety of data processing steps were applied in advance to the actual analysis of oscillations. In the following the main steps of processing and analysis will be illustrated on a simplified and certainly incomplete level. The interested reader is referred to the detailed analysis part by Korbinian Kapsner in the Supplementary Information of [64]. **Noise reduction and drift correction:** First, for each droplet the fluorescence intensity is

normalised by the reference dye intensity at the droplets position. Second, a baseline is determined by applying a strong smoothing filter to the normalised droplet intensity time traces. This baseline is subtracted from the normalised trace in order to remove signal drift which could obscure the later determination of oscillation extrema. The baseline corrected trace is then again smoothed but more conservatively in order to reduce short-term fluctuations that are considerably faster than the expected oscillator dynamics. Due to this filtering step periods shorter than 80 min cannot be detected. **Filtering of valid traces:** A number of criteria is used to filter out traces with features that could falsify the later analysis of the oscillations. Traces were discarded with sudden steps, with too less data points (droplet moved to far or out of the observed area), with too low fluorescence of the normalisation dye (background area), where the radius showed strong variations with time (either no droplet or drift in and out of focal plane), with high noise, and a few more. The fraction of droplets which passed these filter criteria were 12.43% for sustained, 6.69% for damped, and 14.39% for strongly damped oscillator tuning. These droplets were further analysed for the detection of extrema and eventually, classified "oscillating" if at least 5 extrema were detected or classified "not oscillating" if less than 5 extrema were detected. For example traces of droplets classified as "not oscillating" see the bulk trace for the strongly damped tuning (Fig. 4.28) or further examples in the Appendix section 6.4.
It is very likely that these criteria also remove some genuine droplet traces that exhibit low amplitude oscillations or non-oscillatory behaviour; consequently, we do not report the fraction of droplets that oscillate, as this measurement would be unreliable. To summarise, in order to filter out artefacts, stringent selection criteria were used to remove anomalous traces deemed likely to result from, for example, failures in droplet tracking or superposition of droplets.

5 Communication and computation of bacteria compartmentalized within microemulsion droplets

In this chapter, communication between spatially separated bacteria will be presented. In section 5.1 the bacterial communication system "quorum sensing" will be introduced. *E. coli* bacteria were equipped with synthetic gene regulatory circuits, which will be presented and characterised in section 5.2. These circuits were used to establish communication of small bacterial colonies in microemulsions. In section 5.3 different droplet communication modes are demonstrated and discussed by means of a simple reaction-diffusion model. A summary of the findings and future prospects of the here presented work will close the chapter in section 5.4.
The work presented and discussed in this chapter was done in collaboration of a group of people. The results were published in the *Journal of the American Chemical Society* [42]. Cultivation of cells and cloning was mainly done by Andrea Mückl. Korbinian Kapsner was involved in data analysis. Characterisation measurements of circuits were the topic of Ronja Berg's bachelor thesis under my supervision. Friedrich C. Simmel simulated the reaction-diffusion system.

5.1 Chemical communication between bacteria

Quorum sensing (QS) is a mode of chemical communication between bacteria mediated by diffusible inducer molecules called autoinducers (AIs). In a typical QS system, autoinducers are produced by an AI synthase, whose production is itself controlled by the presence of AI. Due to the membrane-penetrating nature of the AIs, AI-inducible genes may also be influenced in neighbouring bacteria, which facilitates the induction of genes in a cell-density dependent manner, hence the name "quorum sensing" [196, 204, 205]. QS has been identified both in gram negative as well as in gram positive bacteria, with different classes of autoinducers [206]. Chemically, the AI-1 class autoinducers are N-acyl-L-homoserine lactones (AHLs), which are found in gram negative bacteria. Specifically, the AI-1 of the bacterium *Aliivibrio fischeri* [207] used in the present work – regulating the LuxR-LuxI system – is the amphiphilic molecule N-(3-Oxohexanoyl)-L-homoserine lactone (3OC6HSL). In the context of synthetic biology, bacterial QS systems have been frequently utilised as a means of communication between engineered bacteria. In a seminal work, Weiss and Knight [34] artificially separated the QS system of *A. fischeri* into a "sender" and a "receiver" part, and were therefore able to realise the first synthetic bacterial communication

system. Based on this sender-receiver system, various gene circuits were engineered, such as a population control system [36], pattern-forming systems [32, 208], or distributed bacterial computing [41]. Artificial microenvironments and microfluidics [40] have been previously used to investigate spatial aspects of cell-to-cell communication via QS. For instance, QS between groups of bacteria was studied using inkjet printing [209], fiberoptic microarrays [210] or by microfluidic confinement within small aqueous droplets [211, 212]. Recently, also signalling between bacteria confined in an emulsion droplet to a neighbouring compartment within a microfluidic double droplet trap was shown [213], a mechanism that is also used in our work. In the following sections, we utilise a genetically engineered QS-based sender-receiver system, but also the conventional inducer/repressor module isopropyl-β-D-thiogalactopyranoside (IPTG)/LacI to implement a spatially extended communication system between *E. coli* bacteria encapsulated within large populations of water-in-oil emulsion droplets [42]. We find that both AHL (3OC6HSL) and IPTG partly dissolve in the oil phase. The resulting, slow diffusion of the inducers from compartment to compartment establishes two chemical communication channels between the droplets, which can be directly demonstrated using computational bacteria that process AHL and IPTG signals as inputs of a simple genetic AND gate. Integration of several slowly diffusing signals within a droplet population using synthetic gene regulatory circuits is expected to enable programmable pattern formation [32] and distributed computing [41, 214] on a shorter length scale than in aqueous medium [215].

5.2 Genetic modules and characterisation

In this section, the genetic modules which we used to establish communication between bacteria are introduced and characterised through their specific transfer functions determined in bulk. In section 5.2.1, the Lux-system originally found in *A. fischeri* is presented. The Lac-system, a conventional inducer/repressor module is shown and characterised in section 5.2.2. Transcription regulation mechanisms of both, the Lux- and the Lac-systems were further utilised to create a genetic AND-gate which will be discussed and quantified in section 5.2.3.

5.2.1 The Lux-system

The Lux-system we utilised occurs in nature as part of the quorum sensing system of the gram-negative marine bacterium *A. fischeri*. Main building blocks of the communication system are the regulatory proteins LuxI and LuxR. LuxI is a synthase for acylated homoserine lactones (AHL), the autoinducer molecules in most quorum sensing systems of gram-negative bacteria. LuxR acts as transcription activator when bound to AHL. To be more precise, the AHL used in our work is N-(3-Oxohexanoyl)-L-homoserine lactone (3OC6HSL). Its chemical structure is shown in Fig. 5.1. In the following, 3OC6HSL will be called AHL for simplicity. AHL is an amphiphilic molecule that can freely cross the cell-membrane [196]. Hence, the AI concentration inside an isolated cell will be diluted by the efflux into the surrounding medium. However, as the cell-density increases AHL can accumulate sufficiently to induce tran-

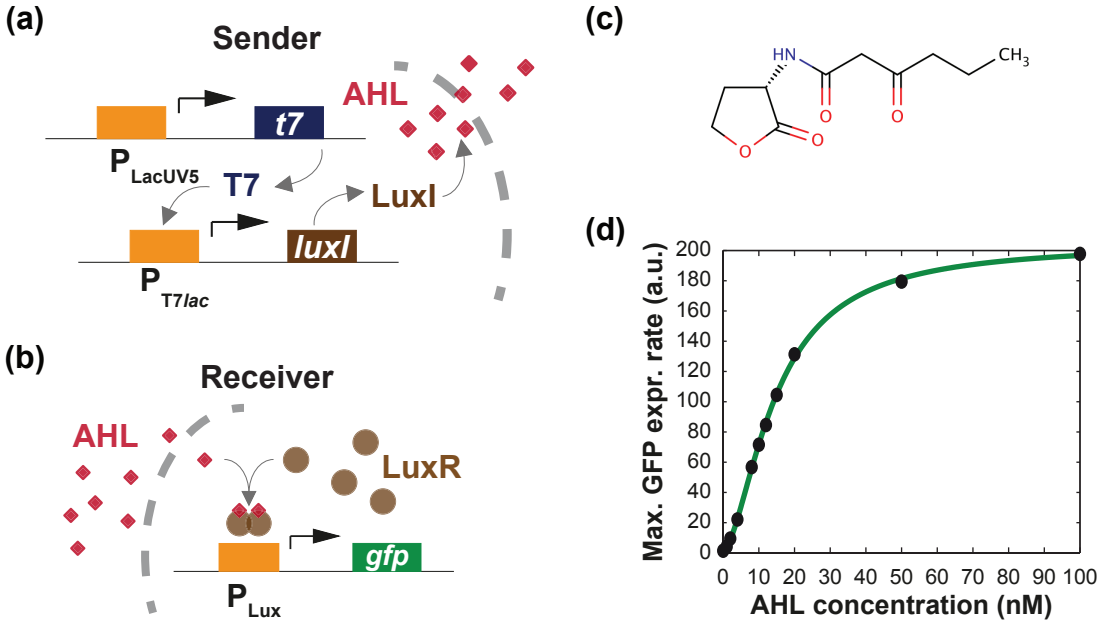

Figure 5.1: The Lux-system can be used to establish chemical communication between bacteria. (a) Schematic of the AHL sender module. Upon induction of the promoter P_{LacUV5} by IPTG (not shown), T7 RNAP (T7) is expressed which transcribes the gene for AHL-synthase LuxI. LuxI synthesises AHL which can freely cross the cell-membrane (gray dashed line). (b) Schematic of the AHL receiver module. The inactive transcription activator LuxR (brown) is expressed constitutively in receiver strains. AHL, which can diffuse through the cell membrane causes LuxR to form dimers which then acts as transcription activator for the promoter P_{Lux} which controls the gene for GFP. (c) The chemical structure of 3OC6HSL. (d) Results of gene induction experiments in bulk. A transfer function was fitted to values of maximal GFP expression rates (max. GFP expr. rate) for varying AHL concentrations. The maximal expression rates were determined by the maximum of the first derivative of fluorescence time courses (typically 2 h after induction) recorded in a plate reader. Black dots are the mean of three identically prepared samples recorded in parallel. The green trace shows the fit of a Hill-function with Hill exponent $n_{AHL} = 1.6$ and threshold $K_{AHL} = 15\,\text{nM}$.

scription of a Lux-promoter upon binding to LuxR. In *A. fischeri*, the expression of LuxI is positively regulated by the AHL-LuxR complex, thus the name auto-inducer. Instead of auto-induction of gene expression, we here used the LuxI/LuxR genes to study communication between bacterial colonies which contained either a gene for the AHL synthase LuxI or a gene under control of a LuxR-inducible promoter. In the sense of communication, bacteria which produce LuxI are in the following called sender, bacteria which produce LuxR and thus respond to AHL are called receiver cells. This system was previously introduced by Weiss and Knight [34]. A schematic of the sender circuit is shown in Fig. 5.1 (a). The gene for LuxI is under control of a promoter for a T7 RNAP P_{T7lac}. Expression of T7 RNAP itself is regulated by the Lac-promoter P_{LacUV5}. In receiver cells, the AHL-binding transcription factor LuxR is produced constitutively. To monitor the occurrence of AHL, receiver cells contain a gene for GFP controlled by the Lux-promoter P_{Lux} (Fig. 5.1 (b)). In absence of AHL, LuxR is inert as transcription factor hence fluorescence of GFP is low. As AHL accumulates by diffusion from a source, it passes the cell membrane and binds to LuxR. Subsequently, LuxR dimerises and activates transcription resulting in an increase in green fluorescence. In order to characterise the receiver circuit, we

performed bulk measurements on a plate-reader where we recorded the time development of fluorescence intensity of GFP for different AHL concentrations. We further built the first derivative of fluorescence trajectories and determined its maximum for each inducer concentration. This gave us a measure for the maximal gene expression level in dependence of induction strength. From this data, we derived the transfer function for receiver bacteria by a fit of a Hill-function (2.85) as introduced in section 2.1.4 with a Hill exponent of $n = 1.6 \pm 0.1$ and an induction threshold of $K_{AHL} = 15 \pm 0.7\,\mathrm{nM}$. Both values are well within the range given by related studies in the literature which are $n_{AHL} = 0.9 - 1.8$ and $K_{AHL} = 1.4 - 100\,\mathrm{nM}$ [216, 217]. The rates of maximum expression and at the basal level were $\alpha_{AHL} = 179.4 \pm 6.2\,^{\mathrm{a.u.}}/_{\mathrm{min}}$ and $\alpha_{0,AHL} = 0.3 \pm 2.4\,^{\mathrm{a.u.}}/_{\mathrm{min}}$, respectively.

5.2.2 The Lac-system

The Lac-system makes use of the *lac* repressor LacI for transcription control. The *lac* repressor was found in *E. coli*'s *lac* operon which is responsible for the regulation of transport and metabolism of lactose. LacI is frequently used for gene induction upon addition of the inducer molecule isopropyl-β-D-thiogalactopyranoside (IPTG). IPTG binds to the tetrameric repressor LacI which results in the release of the operator at the repressed gene, hence induces transcription. The Lac-system, used here to monitor the level of IPTG is schematically shown in Fig. 5.2 (a). These bacteria are in the following called IPTG-receiver cells. In IPTG-receiver bacteria, LacI is constitutively produced. LacI inhibits transcription of the promoter $\mathrm{P_{LacUV5}}$ which controls the gene for T7 RNAP and simultaneously represses the promoter $\mathrm{P_{T7}}_{lac}$, which controls the gene for a red fluorescent protein (RFP). As IPTG accumulates, it sequesters repressors and the promoters are released. Transcription is then activated for both promoters and red fluorescence starts to increase as soon as T7 RNAP appears. The chemical structure of IPTG is shown in Fig. 5.2 b. In analogy to the Lux-system, we studied gene induction in our IPTG-receiver circuit in bulk measurements on a plate reader. A transfer function was derived that characterises the bacterias' response as function of the IPTG concentration. In Fig. 5.2 (c), the experimentally determined maximum gene expression rates are plotted in dependence of the IPTG concentration. A fit of a Hill-curve (2.85) reveals a Hill exponent of $n_{IPTG} = 1.3 \pm 0.1$ and an induction threshold concentration of $K_{IPTG} = 300 \pm 10\,\mu\mathrm{M}$. The rates of maximum expression and at the basal level were $\alpha_{IPTG} = 73.4 \pm 1.5\,^{\mathrm{a.u.}}/_{\mathrm{min}}$ and $\alpha_{0,IPTG} = 0.3 \pm 0.7\,^{\mathrm{a.u.}}/_{\mathrm{min}}$, respectively.

5.2.3 The genetic AND-gate

Chemical signalling between bacteria could be used to build larger network structures of bacterial cell sub-colonies similar to multicellular organisms. For this purpose, either more signalling channels are required or cells have to be introduced that react on specific concentration profiles or different combinations of inducers. The AHL- and IPTG-receiver cells' output is only regulated by AHL or IPTG, respectively. We integrated both mechanisms in a gene circuit that responds with an output only when both inducer molecules are present, thus a genetic AND-gate. A gene circuit was designed as schematically shown in Fig. 5.3 (a). In the native state, LacI is produced

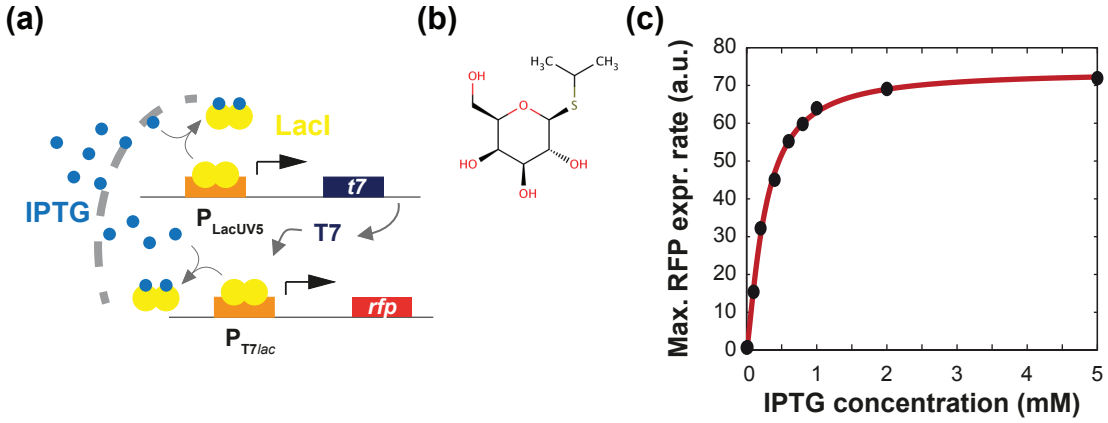

Figure 5.2: The Lac-module. (a) Schematic of the Lac-receiver module. *lac* repressor LacI inhibits transcription of a T7 RNAP until IPTG accumulates. The reaction is monitored by means of RFP fluorescence produced by T7 RNAP. (b) The chemical structure of isopropyl-β-D-thiogalactopyranoside (IPTG). (c) Determination of the transfer function for the Lac-receiver module in *E. coli* in analogy to Fig. 5.1. We determined the maximum of the first derivative of fluorescence time courses of IPTG-receiver bacteria induced with different amounts of IPTG for characterisation, which corresponds to the maximum RFP expression rate (max. RFP expr. rate), typically 2 h after induction. The black dots are the mean of three identically prepared samples recorded in parallel. The red trace shows the fit of a Hill-function with Hill exponent $n = 1.3$ and threshold $K_{IPTG} = 300\,\mu M$.

constitutively and represses the promoter $P_{LlacO\text{-}1}$ which controls a gene for the AHL-binding transcription factor LuxR. The output gene which encodes for GFP is under control of the promoter P_{Lux}. Hence, in absence of either inducers, IPTG and AHL, the *lac* repressor inhibits production of LuxR and GFP expression cannot be activated. If only IPTG is present, LuxR is produced but inert as transcription factor due to lack of AHL. If only AHL is present, GFP expression cannot be activated since LuxR concentration is low. Only when both, IPTG and AHL are present, GFP fluorescence increases since LuxR is produced and rendered into dimerisation after binding to AHL. We characterised the genetic AND-gate *in vivo* by measurements on a plate reader, similar as for the AHL- and IPTG-reporters, but this time with variations in both inducer concentrations. For each pair of inducer concentrations, we determined the maximum in GFP expression rate as the maximum of the first derivative of the fluorescence time traces. A two-dimensional input function was determined by a fit of equation 2.87 to the experimental data (Fig. 5.3). This fit-function gives us Hill-exponents and threshold values for each inducer which were $n_{AHL} = 1.5 \pm 0.2$ and $n_{IPTG} = 1.5 \pm 0.2$ and threshold values of $K_{AHL} = 1.7 \pm 0.2\,nM$ and $K_{IPTG} = 20 \pm 0\,nM$ for AHL and IPTG, respectively. The rates at maximum expression and at the basal level were $\alpha_{AND} = 150.3 \pm 8.1\,^{a.u.}/_{min}$ and $\alpha_{0,AND} = 4.5 \pm 3.0\,^{a.u.}/_{min}$.

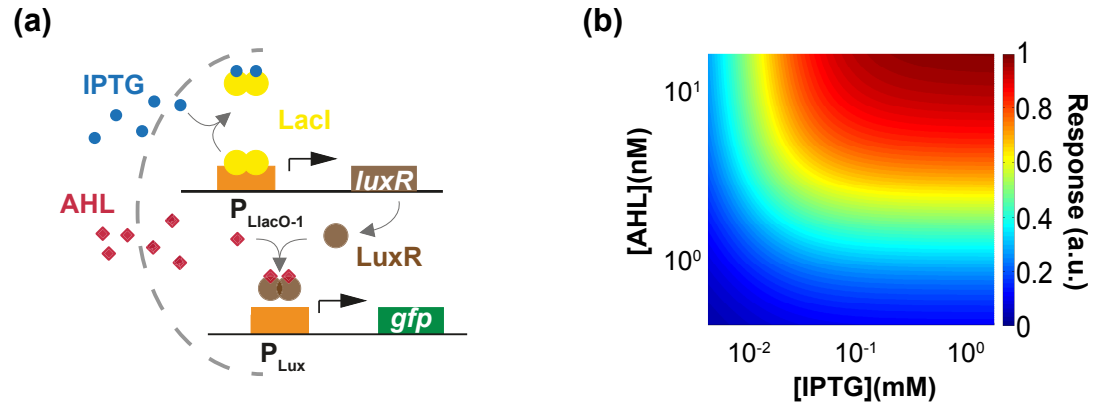

Figure 5.3: The genetic AND-gate. (a) Schematic of the genetic AND-gate module. Repressor LacI is produced constitutively in *E. coli* AND-gate cells. LacI inhibits expression of LuxR in absence of IPTG. Thus, in presence of IPTG, the promoter $P_{LlacO-1}$ is activated and the gene for LuxR is transcribed. However, LuxR as monomer cannot activate transcription of the output gene GFP. When IPTG and AHL are present, LuxR is produced and can form the active transcription factor upon binding of AHL. As a consequence the circuit responds with the production of GFP. (b) Fit of a two-dimensional transfer function determined from bulk measurements of the genetic AND-gate in *E. coli*. The derived fit parameters are Hill exponents $n_{AHL} = 1.5$ and $n_{IPTG} = 1.5$ and threshold values $K_{AHL} = 1.7\,\text{nM}$ and $K_{IPTG} = 20\,\text{nM}$ for AHL and IPTG, respectively.

5.3 Chemical communication in microemulsions

In the previous sections, the genetic circuits used here were introduced and their transfer functions were determined. All transfer functions revealed a sufficiently strong fold change that inducer concentrations can be detected by means of fluorescence read-out. In the following sections, we utilise the genetic modules to establish different communication modes. In section 5.3.2 results of control experiments are presented in which we tested gene induction in the microemulsion environment. We then studied communication modes where encapsulated receiver bacteria and droplets containing inducer molecules were observed for the inducers AHL and IPTG (sections 5.3.3 and 5.3.4). In the following, droplets filled with inducer are called "reservoir". We then followed the response of receiver cell containing droplets in the presence of droplets containing sender cells (section 5.3.5). In order to demonstrate that multiple communication channels can be integrated in our system we finally encapsulated AND-gate bacteria and followed their response in the presence of different reservoir droplets (section 5.3.6).

5.3.1 *E. coli* in microemulsions

E. coli cells encapsulated into microemulsion droplets are constricted by the finite compartment volume which sets an upper limit for the number of bacteria in a droplet. Radii of droplets observed were in average $\approx 12.5\,\mu\text{m}$ corresponding to a volume for a spherical compartment of $\approx 8\,\text{pL}$. Before droplets were generated, we diluted the bacteria in growth medium to an optical density value of $OD600 \approx 0.1$. If one assumes binomial distribution of bacteria into droplets the expectation value for the number of bacteria in a droplet volume of 8 pL would be 0.2 per droplet using as conversion factor from optical density to *E. coli* cell concentration in Luria-

Bertrani medium $OD600 = 0.1 \,\widehat{\approx}\, 2 \cdot 10^7\,{}^{\text{cells}}/_{\text{mL}}$ [149]. In this case only about 20% of droplets would contain at least one cell. However, we found that typically more than 90% of droplets contained bacteria presumably due to cell growth during the droplet production process which usually took more than an hour. Due to space constriction an upper limit for the number of bacteria in a droplet is 8000 *E. coli* cells with a bacterial volume of 1 fL. Yet, droplets are never completely filled with bacteria presumably due to the limited nutrients. Cell growth in droplets typically persisted over $5 - 10$ h, after which a colony had formed with a cross-section smaller than the droplet diameter. We thus assumed that the "carrying capacity" N_{max} is on the order of

$$N_{max} \approx 100 - 1000.$$

We will later use a simple reaction-diffusion model to explain the experimental results. To describe cell growth in droplets we will use the logistic function

$$N(t) = \frac{N_{max} N_0 \mathrm{e}^{\gamma t}}{N_{max} + N_0(\mathrm{e}^{\gamma t} - 1)} \tag{5.1}$$

with the number of cells initially present in a droplet N_0, and the cell division rate γ.

5.3.2 Gene induction in microemulsions

Due to their amphiphilic nature, both inducers used here are capable of penetrating cell membranes. We therefore addressed the question whether the inducer molecules are also partly soluble in the oil-surfactant mix used here. Droplets were produced in microfluidics which contained either 200 nM AHL or 10 mM IPTG in LB medium. We filled the reservoir emulsions into 500 µL reaction tubes and incubated the samples for 3 h at 37 °C. In order to test if inducer partitioned into the oil phase, we carefully extracted the oil phase (situated at the lower part of the tube since water has a lower density as the here used oil) and added the aliquots to observation channels filled with droplets containing the corresponding receiver bacteria. As positive and negative control, we also prepared bacteria filled droplets which were either initially induced with 20 nM AHL or 1 mM IPTG, respectively, or in their native state. Resulting fluorescence microscopy images taken 12 h after start of the experiment are shown in Fig. 5.4 (a) for AHL and (b) for IPTG. In both cases, the initially induced bacteria showed strong fluorescence intensities while bacteria in growth medium without inducer showed no significant fluorescence. However, both AHL and IPTG receiver cells which were mixed with an oil aliquot extracted from the incubated reservoir droplets showed considerable fluorescence in the GFP and the RFP channel, respectively. We therefore figured that to some extent both inducers partition into the oil phase. This is in accordance with literature values which report a slight hydrophobicity of AHL expressed by an octanol/water partitioning coefficient in the range of $0.2-2$ [218, 219]. The partitioning coefficient $\log p$ is defined as

$$\log p = \log \frac{[\mathrm{X}]_\mathrm{o}}{[\mathrm{X}]_\mathrm{w}} \tag{5.2}$$

with the concentration of the regarded species in an octanol-rich $[\mathrm{X}]_\mathrm{o}$ and in a water-rich $[\mathrm{X}]_\mathrm{w}$ phase. We further performed control experiments to test whether the gene

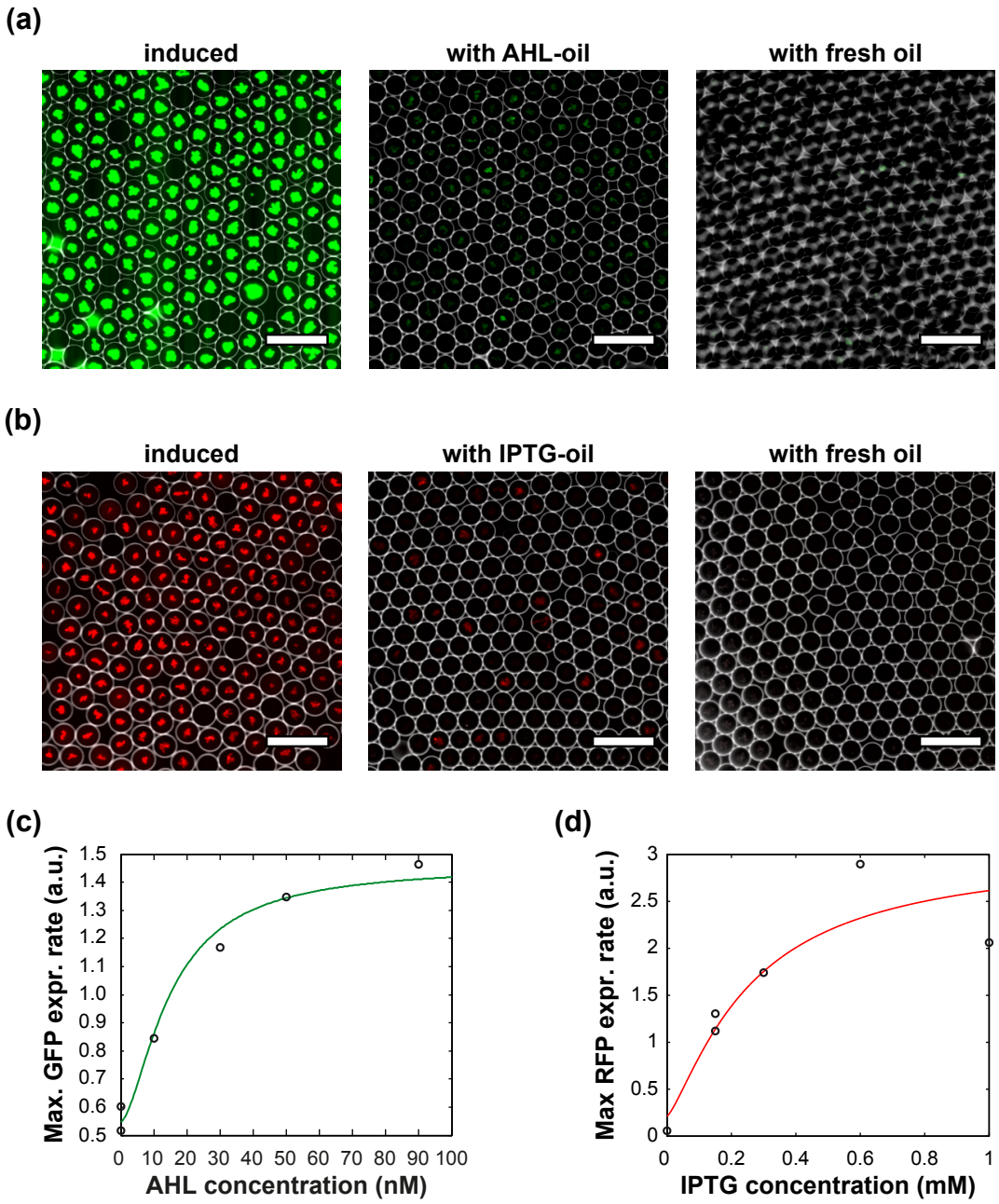

Figure 5.4: Gene induction in microemulsions. (a) Does AHL dissolve in FC-40 Fluorinert oil? To address this question, we produced reservoir droplets filled with 200 nM AHL and incubated the emulsion for 3 h at 37 °C. The left fluorescence microscopy image shows AHL-receiver cells which were induced upon addition of 20 nM AHL before encapsulation. Centre image: AHL-receiver cells were encapsulated into microemulsion droplets. To test whether AHL went into the oil-phase during incubation, we carefully extracted an aliquot of oil and added it to the not-induced receiver bacteria in droplets. The image on the right hand side shows the same bacteria as in the centre but without incubated oil. (b) Same experiment as (a) but with IPTG and IPTG-receiver cells instead. Micrographs shown in (a) and (b) were taken 12 h after the incubated oil was added. Scale bars are 100 µm. (c) Circles represent results from a gene expression experiment of AHL-receiver cells in droplets with different concentrations of AHL added to their medium. The green trace is the rescaled bulk transfer function. (d) Circles represent results from a gene expression experiment of IPTG-receiver cells in droplets with different concentrations of IPTG added to the growth medium. The red trace is the rescaled bulk transfer function.

expression response of bacteria in droplets is significantly different from the behaviour in bulk. To this end, we performed experiments similar to the bulk experiments presented in sections 5.2.1 and 5.2.2. Receiver bacteria were encapsulated in droplets with different inducer concentrations and their gene expression was followed with time-lapse microscopy. We tracked and analysed bacteria containing droplets in time-lapse movies. In contrast to bulk measurements, there was no way to quantify the cell number in droplets. However, droplets were normalised by their size and the maxima of the first derivatives for individual droplets were determined. The mean values in dependence of the inducer concentrations are plotted in Figs. 5.4 (c) and (d). To compare the encapsulated bacteria's response with results from gene expression experiments in bulk, the transfer functions for AHL- and IPTG-receiver cells are drawn into the graph. For both experiments, on the scope of accessibility of droplet data, the results indicate similar gene induction behaviour in bulk and in droplets. Thus, we assume that it is reasonable to use the characteristic values of the transfer curves derived from bulk experiments to explain gene induction in microemulsions.

5.3.3 Communication mode: AHL reservoir - AHL receiver

We first studied communication between AHL filled reservoir droplets and droplets containing the AHL-receiver cells introduced in section 5.2.1 with time-lapse fluorescence microscopy (Fig. 5.5 (a)). Reservoir droplets contained 200 nM AHL which is more than ten times the threshold concentration K_{AHL} determined in bulk experiments. To distinguish between the two types of droplets, 0.5 µM Texas Red was added to reservoir droplets. Fig. 5.5 (b) shows an image series of a time-lapse movie of a reservoir droplet surrounded by receiver droplets. Fluorescence channels for GFP and Texas Red were merged. Droplet boundaries are indicated by white circles which were derived from inverted bright field images. Gene expression in receiver droplets started in a distance dependent manner, first in droplets in close vicinity to the reservoir and later also in droplets further apart. Droplets close to reservoirs which remained dark did typically not contain bacteria. We tracked individual droplets over time and extracted the single droplet data like time courses of position, size and intensity. For analysis, we grouped the 1405 receiver cell containing droplets according to their distance from the closest reservoir (in total 91) and plotted the mean fluorescence intensity as a function of time for individual distances (Fig. 5.5 (c)). A grey dashed line indicates the fluorescence evolution of encapsulated bacteria which were not induced and isolated from reservoir droplets derived in a control experiment. Even though, the fluorescence of this "negative" control also reaches the final fluorescence value for receiver far apart from reservoirs the profile of the time trace is significantly different. While the fluorescence in the "negative" control increases approximately linearly, all droplet traces which shared space with reservoirs show an initial exponential increase and saturate later at a constant value. The shape is presumably caused by the transient induction of inducer which depletes with time due to the finite volume of reservoirs. We further plotted the distance dependent fluorescence profiles for different time points shown in Fig. 5.5 (d). In order to understand the underlying mechanism observed in the experiment, we modelled the system by means of a simple reaction-diffusion reaction which will be introduced in the following.
The bacteria containing emulsion was regarded as homogenous medium with an effec-

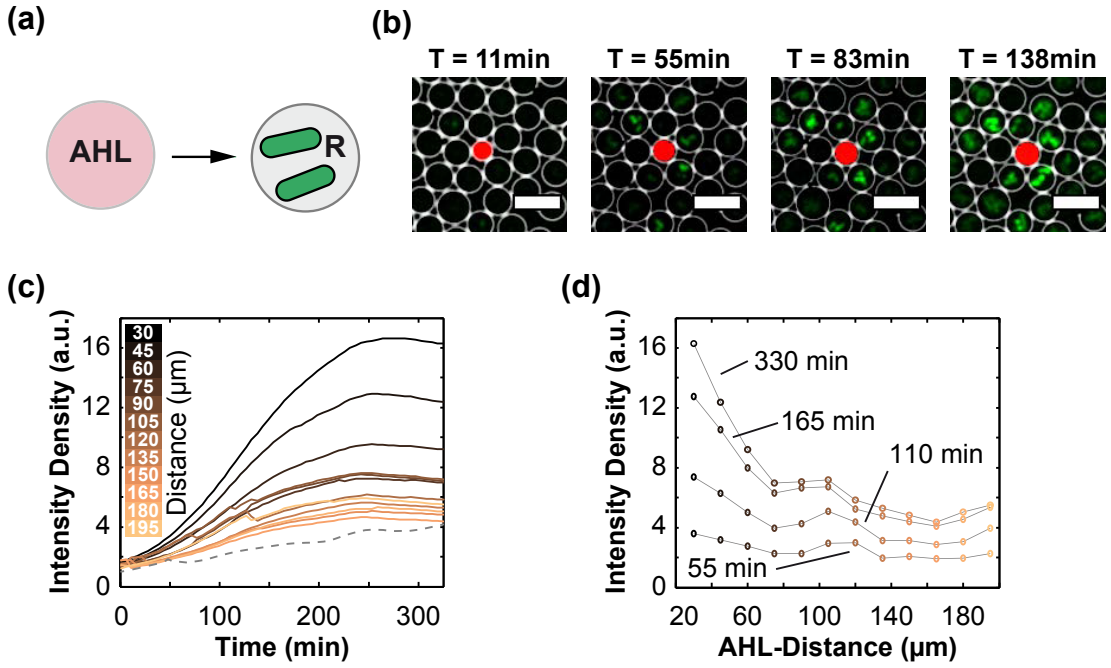

Figure 5.5: Communication mode: AHL reservoir - AHL receiver bacteria. (a) Simple scheme of the experiment. The induction of gene expression of receiver bacteria in presence of AHL filled reservoir droplets is studied. (b) Image series extracted from a fluorescence microscopy time-lapse movie of droplets containing AHL-receiver cells surrounding an AHL reservoir droplet (red). White circles are derived from inverted bright field images and represent droplet boundaries. The scale bar is 50 µm. (c) The expression of GFP was followed in individual droplets. To study the distance dependent GFP induction, we grouped droplets according to their distance to the next AHL-reservoir and plotted the mean fluorescence intensity as a function of time. An appropriately scaled fluorescence time trace from a control experiment with receiver cells isolated from AHL molecules is included for comparison (dashed line). (d) Fluorescence profiles of traces shown in (a) for different times.

tive diffusion coefficient D_{eff}. For better comparison purpose with the experimentally derived results we included cell growth as discussed in section 5.3.1 in the model. The reaction-diffusion model for AHL reservoirs and AHL-receiver cell droplets is described by the following set of differential equations:

$$\frac{\partial[\mathrm{AHL}]}{\partial t} = D_{eff} \cdot \nabla^2[\mathrm{AHL}] - \delta_{AHL}[\mathrm{AHL}] \tag{5.3}$$

$$\frac{d[\mathrm{GFP}]}{dt} = \frac{R\,\alpha_{GFP}}{1 + \left(\frac{K_{AHL}}{[\mathrm{AHL}]}\right)^{n_{AHL}}} - \delta_{GFP}[\mathrm{GFP}] \tag{5.4}$$

$$\frac{d[\mathrm{GFP_{mat}}]}{dt} = \alpha_{mat}[\mathrm{GFP}] - \delta_{GFP}[\mathrm{GFP_{mat}}] \tag{5.5}$$

where [AHL], and [GFP] is the concentration of AHL and nascent GFP, respectively, and [$\mathrm{GFP_{mat}}$] is the concentration of maturated GFP which appears fluorescent. δ_{AHL} and δ_{GFP} are the effective degradation rates for AHL and GFP. α_{GFP} is the maximum GFP expression rate, n_{AHL} and K_{AHL} are the Hill exponent and induction threshold derived from bulk experiments in section 5.2.1 and α_{mat} is the maturation rate of GFP. We further assumed that GFP production rate is proportional to cell growth, which we accounted for with the bacterial cell growth rate

$$R = \dot{N}(N_{max}, N_0, \gamma) \tag{5.6}$$

where $\dot{N}$ is the first time derivative of the logistic function 5.1. GFP concentrations are distance dependent due to the diffusively spreading AHL from a reservoir. Except for the experimentally determined parameters n_{AHL} and K_{AHL}, we used typical values from the literature for model simulations listed in table 5.1. The only unknown parameter in our model is the effective diffusion coefficient D_{eff} in 5.3. In order to

Parameter	Value	Reference
δ_{AHL}	$1 \cdot 10^{-5}\,1/\mathrm{s}$	[205]
δ_{GFP}	$2 \cdot 10^{-4}\,1/\mathrm{s}$	
α_{mat}	$3.3 \cdot 10^{-3}\,1/\mathrm{s}$	[220]
α_{GFP}	1	
γ	$2\,1/\mathrm{h}$	
N_0	1	
N_{max}	100	

Table 5.1: Parameter set for model simulations for AHL reservoirs and AHL-receiver bacteria.

find a fitting value for D_{eff} we simulated the time development of fluorescing GFP_{mat} as shown for the experiment in Fig. 5.5 (c) for varying D_{eff}. The initial conditions used were

$$[\mathrm{AHL}](\vec{r}, t=0) = \begin{cases} 200\,\mathrm{nM}, & \text{if } |\vec{r}| < 12.5\,\mu\mathrm{m} \\ 0\,\mathrm{nM}, & \text{otherwise} \end{cases}$$
$$\text{and} \tag{5.7}$$
$$[\mathrm{GFP}](\vec{r}, t=0) = [\mathrm{GFP}]_{mat}(\vec{r}, t=0) = 0\,\mathrm{nM}.$$

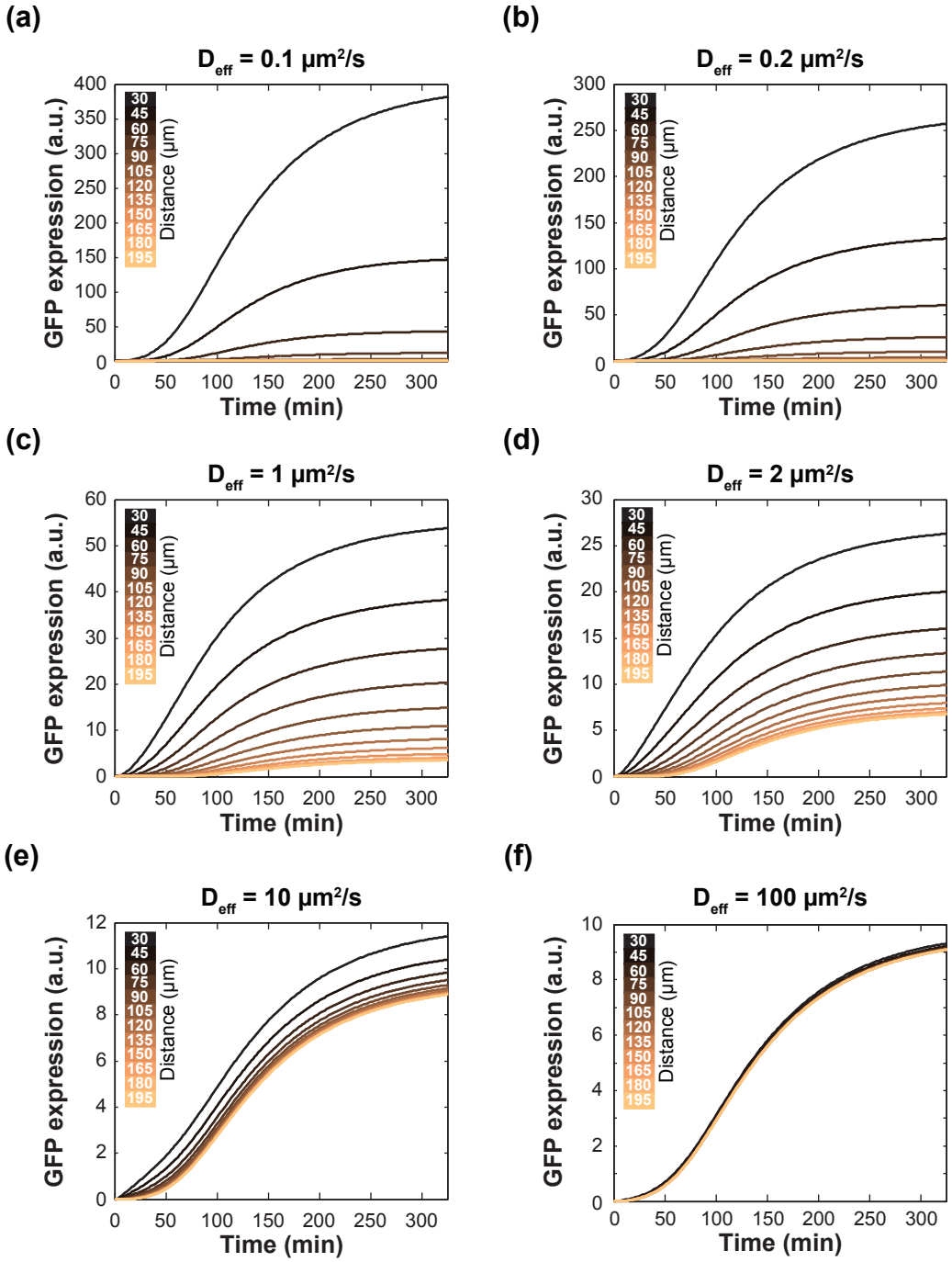

Figure 5.6: Reaction-diffusion simulation: AHL reservoir - receiver bacteria for different effective diffusion coefficients using the transfer function determined in bulk experiments.

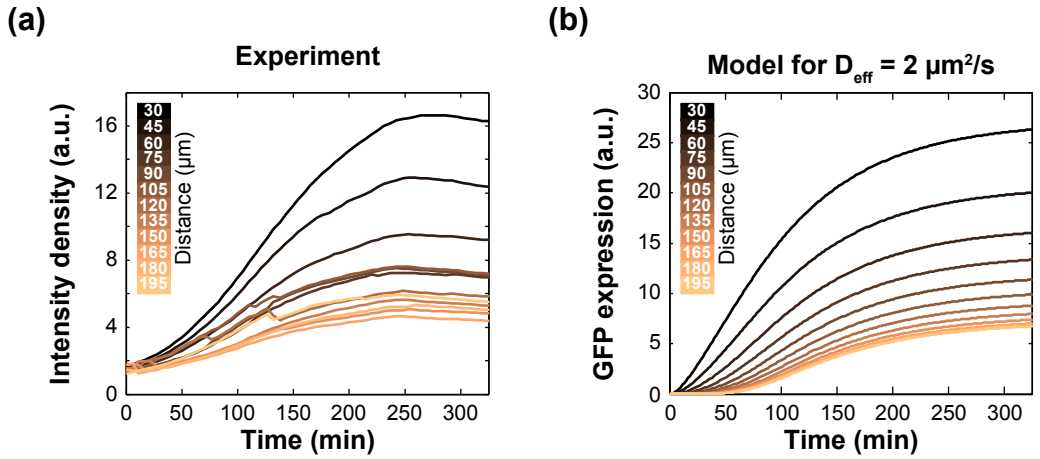

Figure 5.7: (a) Experiment and (b) simulation for AHL reservoirs and receiver bacteria. Experiment and simulation match best for a reduced effective diffusion coefficient of $D_{eff} = 2\,\mu m^2/s$.

The results of model simulations for $D_{eff} = [0.1, 0.2, 1, 2, 10, 100]\,\mu m^2/s$ are shown in Fig. 5.6 (a)-(f). Simulations were performed with Matlab's solver 'pdepe' using cylindrical symmetry. Typical values of diffusion coefficients for AHL in aqueous bulk solutions are in the range of $100 - 1000\,\mu m^2/s$ [209, 214, 215, 221]. However, already for a diffusion coefficient $D_{eff} = 100\,\mu m^2/s$ as shown in Fig. 5.6 (f) no considerable distance dependence of gene induction on length scales as in the experiment is apparent. With decreasing effective diffusion coefficient, the time traces for different distances fan out. The simulation result agrees best with the experiment for $D_{eff} = 2\,\mu m^2/s$ (cf. Fig. 5.7). Since we did not consider basal transcription or leakage expression in absence of AHL in our model there is no offset in fluorescence intensity at $t = 0$ in the simulations. A further difference is that simulations assume an isolated reservoir droplet surrounded by droplets containing receiver bacteria. Thus, there are no next-nearest-neighbour-droplet inducer contributions. In the experiment, in contrast, we observe a 2-dimensional array of randomly mixed reservoir and receiver droplets meaning that the resulting traces are based on gene induction by superimposed inducer wavefronts originating from more than one droplet. The droplets arrange typically in a hexagonally closed packed lattice. Considering the number of reservoir and receiver droplets, the mean distance between reservoir droplets is $\approx 66\,\mu m$. Thus, experimental traces as shown in Fig. 5.7 (a) only approximately reflect the distance from nearest AHL reservoirs. Nonetheless, a reduced diffusion coefficient is plausible in the light of a variety of factors like different diffusion coefficients in the different phases, permeability of droplet-droplet interfaces and geometrical effects.

5.3.4 Communication mode: IPTG reservoir - IPTG receiver

We further reasoned that droplet-droplet induction of gene expression as presented in section 5.3.3 may also work for a conventional inducer molecule like IPTG. As schematically shown in Fig. 5.8 (a), we therefore compartmentalized receiver bacteria containing the Lac-module introduced in section 5.2.2 in microfluidic droplets and followed their gene expression in time-lapse fluorescence microscopy in presence of reservoir droplets containing 10 mM IPTG and 0.5 µM Alexa 488 dye for visualisation.

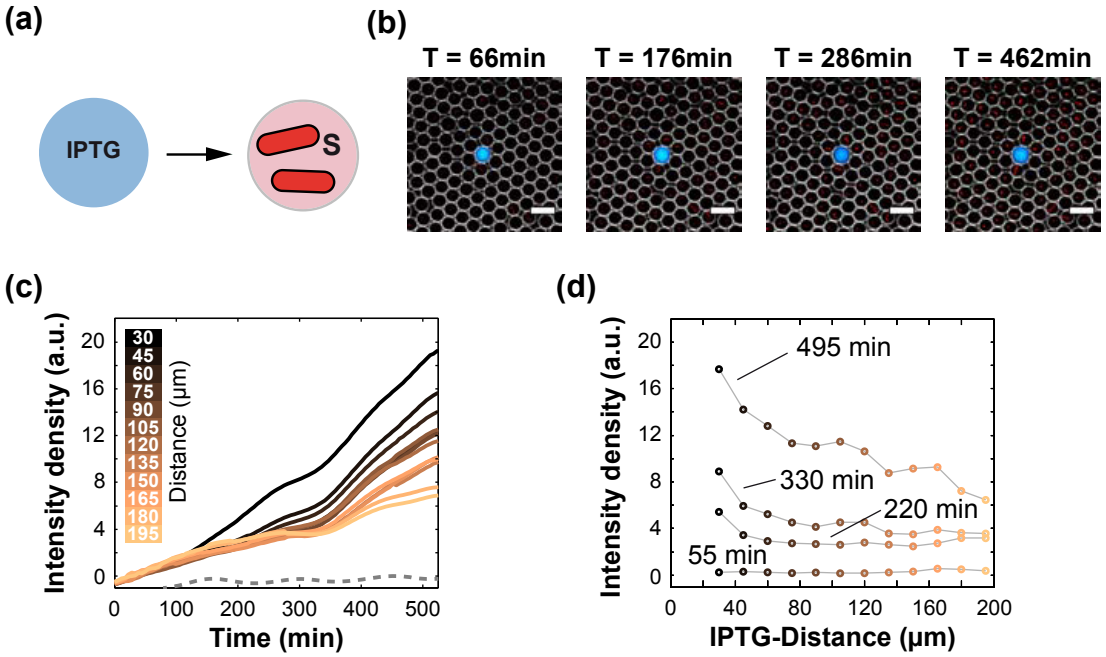

Figure 5.8: Communication mode: IPTG reservoir - IPTG receiver. (a) Schematic of the experiment. The induction of gene expression of IPTG-receiver cells is studied in the presence of reservoir droplets containing 10 mM IPTG. (b) Image series of a time-lapse movie of droplets containing not induced IPTG-receiver bacteria arranged around an IPTG-reservoir. White circles are droplet boundaries derived from inverted BF images. Red fluorescence increases with time in a distant dependent manner from the reservoir droplet. (c) Fluorescence evolution in droplets grouped according to their distance from the next IPTG-reservoir droplet. An appropriately scaled fluorescence time trace from a control experiment with uninduced receiver cells is included for comparison (dashed line). (d) Fluorescence profiles of the traces shown in (a) for different times.

Indeed, time-lapse movies confirmed that expression of RFP starts first in droplets close to a reservoir and later also in droplets further apart. An image time series of merged green, red and inverted bright field images is shown in Fig. 5.8 (b). We tracked individual droplets in time-lapse movies and extracted the data for single droplets which there were 76 IPTG reservoirs and 1358 filled with IPTG receiver. As for AHL-receiver cells, we grouped the IPTG-receiver cell droplets depending on their distance to the next reservoir and determined the mean fluorescence intensity trace for every group. Fig. 5.8 (c) shows the fluorescence time traces for different distances from a reservoir droplet. The fluorescence of encapsulated not-induced IPTG-receiver cells in absence of reservoir droplets remained low as a reference experiment revealed, drawn as grey dashed line. To understand the distance dependent gene induction we generated a similar reaction-diffusion model as for AHL receiver. However, the gene circuit responsible for the detection of IPTG is a bit more complicated. As shown in Fig. 5.2 (a), the RFP output gene is controlled by an IPTG inducible promoter for T7 RNAP. Since expression of T7 RNAP itself is initially suppressed by LacI, RFP is only present at a low concentration level. Consequently, in advance to RFP expression upon release of the promoters by IPTG, T7 RNAP has to be expressed which is accounted for by an additional differential equation. The set of differential equations we used to describe gene expression in droplets filled with IPTG receiver

cells in presence of an inducer reservoir is then:

$$\frac{\partial[\text{IPTG}]}{\partial t} = D_{eff} \cdot \nabla^2[\text{IPTG}] - \delta_{IPTG}[\text{IPTG}] \tag{5.8}$$

$$\frac{d[\text{T7}]}{dt} = R \cdot \alpha_0 + \frac{R \cdot \alpha_{T7}}{1 + \left(\frac{K_{IPTG}}{[\text{IPTG}]}\right)^{n_{IPTG}}} - \delta_{T7}[\text{T7}] \tag{5.9}$$

$$\frac{d[\text{RFP}]}{dt} = R \cdot \alpha_0 + \frac{R \cdot \alpha_{RFP}}{1 + \left(\frac{K_{IPTG}}{[\text{IPTG}]}\right)^{n_{IPTG}}} \cdot \frac{1}{1 + \left(\frac{K_{T7}}{[\text{T7}]}\right)} - \delta_{RFP}[\text{RFP}] \tag{5.10}$$

$$\frac{d[\text{RFP}_{\text{mat}}]}{dt} = \alpha_{mat}[\text{RFP}] - \delta_{RFP}[\text{RFP}_{\text{mat}}] \tag{5.11}$$

where [IPTG], [T7], [RFP], and [RFP_{mat}] is the concentration of IPTG, T7 RNAP, RFP, and maturated RFP, and δ_{IPTG}, and δ_{T7} are degradation rates for IPTG and T7 RNAP, and δ_{RFP} is the degradation rate for RFP and maturated RFP which is assumed to be equal. α_0 is the leak expression rate in absence of inducer and is assumed to be the same for T7 RNAP and for RFP. α_{T7}, and α_{RFP} are the expression rates of T7 RNAP and RFP at full induction. α_{mat} is the maturation rate. For the cell growth R (5.6) we chose $\gamma = 0.5\,1/\text{h}$ which is based on the observation that cell growth of IPTG-receiver cells was considerably slower than for AHL-receiver bacteria. n_{IPTG} and K_{IPTG} are Hill exponent and induction threshold determined in bulk experiments (section 5.2.2), and K_{T7} is the dissociation constant for binding of T7 RNAP and its promoter. The parameters used for model simulations are listed in table 5.2. The boundary conditions were

$$[\text{IPTG}](\vec{r}, t=0) = \begin{cases} 10\,\text{mM}, & \text{if } |\vec{r}| < 12.5\,\mu\text{m} \\ 0\,\text{mM}, & \text{otherwise} \end{cases}$$
$$\text{and} \tag{5.12}$$
$$[\text{T7}](\vec{r}, t=0) = [\text{GFP}](\vec{r}, t=0) = [\text{GFP}]_{mat}(\vec{r}, t=0) = 0\text{mM}.$$

The RFP expression of IPTG receiver cells in droplets in dependence of the dis-

Parameter	Value	Reference
δ_{IPTG}	0	
δ_{T7}	$1 \cdot 10^{-5}\,1/\text{s}$	
δ_{RFP}	$1 \cdot 10^{-4}\,1/\text{s}$	
α_0	$3.3 \cdot 10^{-3}\,1/\text{s}$	
α_{T7}	$0.1\,1/\text{s}$	
α_{RFP}	$10\,1/\text{s}$	
α_{mat}	$3.3 \cdot 10^{-3}\,1/\text{s}$	[220]
K_{T7}	1 nM	
N_0	1	
N_{max}	100	

Table 5.2: Parameter set for model simulations for IPTG reservoirs and receiver bacteria.

tance from an IPTG reservoir droplet was simulated. Comparison of experimental

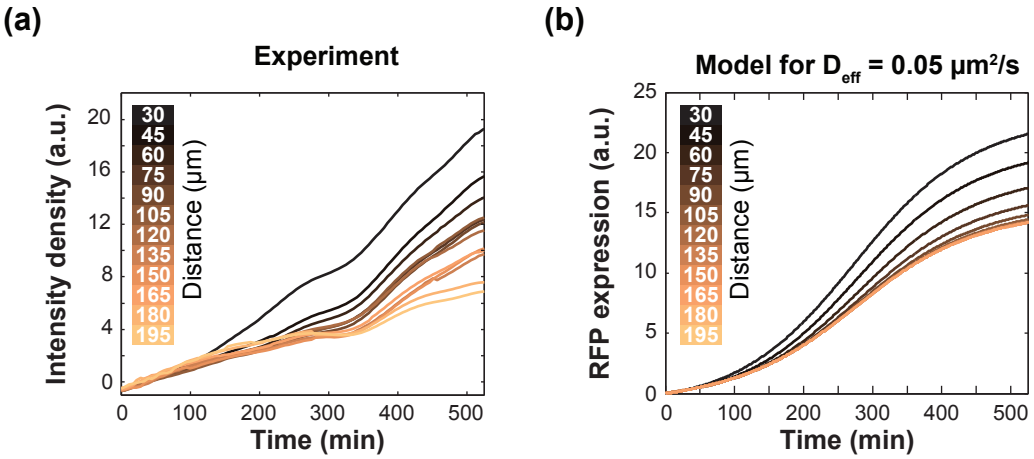

Figure 5.9: Experiment and simulation for IPTG reservoirs and receiver bacteria. Experiment and simulation maths well for a reduced diffusion coefficient of $D_{eff} = 0.05\,\mu m^2/s$. The reaction-diffusion model was simulated using the parameters of the transfer function derived in bulk measurements.

and simulated traces suggests an even lower effective diffusion coefficient as for AHL of $D_{eff} = 0.05\,\mu m^2/s$. Fluorescent time traces for droplets grouped according to their distance to a reservoir droplet are shown in Fig. 5.9 derived experimentally (a) and by simulation of the reaction-diffusion model (b). The lower effective diffusion coefficient for IPTG could be explained by its stronger hydrophilicity expressed by an octanol/water partitioning coefficient of $\log K_{ow} = -1.26$ [200] compared to AHL.

5.3.5 Communication mode: AHL sender - AHL receiver

As communication between droplets was successful for inducer reservoirs and droplets filled with receiver cells we next studied communication between droplets containing AHL-sender cells with droplets filled with AHL-receiver cells. The sender bacteria were equipped with the sender circuit of the Lux-system introduced in section 5.2.1. The receiver bacteria are the same as discussed in section 5.3.3. We therefore simply replaced the AHL reservoir droplets by droplets filled with sender bacteria which produced the AHL synthase LuxI as schematically shown in Fig. 5.10 (a). An image series extracted from a a time-lapse movie shows increasing green fluorescence, first in receiver cell droplets in close proximity to an AHL sender droplet and later also in droplets more distant (Fig. 5.10 (b)). Since the sender bacteria did not express a fluorescent protein, we visualised these droplets by adding 0.5 µM to the growth medium. As for communication modes presented in section 5.3.3 and 5.3.4, we tracked individual droplets in time-lapse movies and analysed the fluorescence evolution in 1679 receiver droplets in presence of 13 sender droplets. Receiver droplets were grouped according to their distance to the next sender droplet. The time development of the mean intensities of individual groups were determined. In this case AHL does not deplete until sender cells run out of nutrients or droplets reach their carrying capacity. Induction of GFP expression starts first in receiver droplets in vicinity of a sender droplet. However, in contrast to communication experiments with reservoir droplets, bacteria are globally induced with time (Fig. 5.10 (c)). Comparison with a fluorescence time trace of encapsulated AHL-receiver bacteria in a reference channel

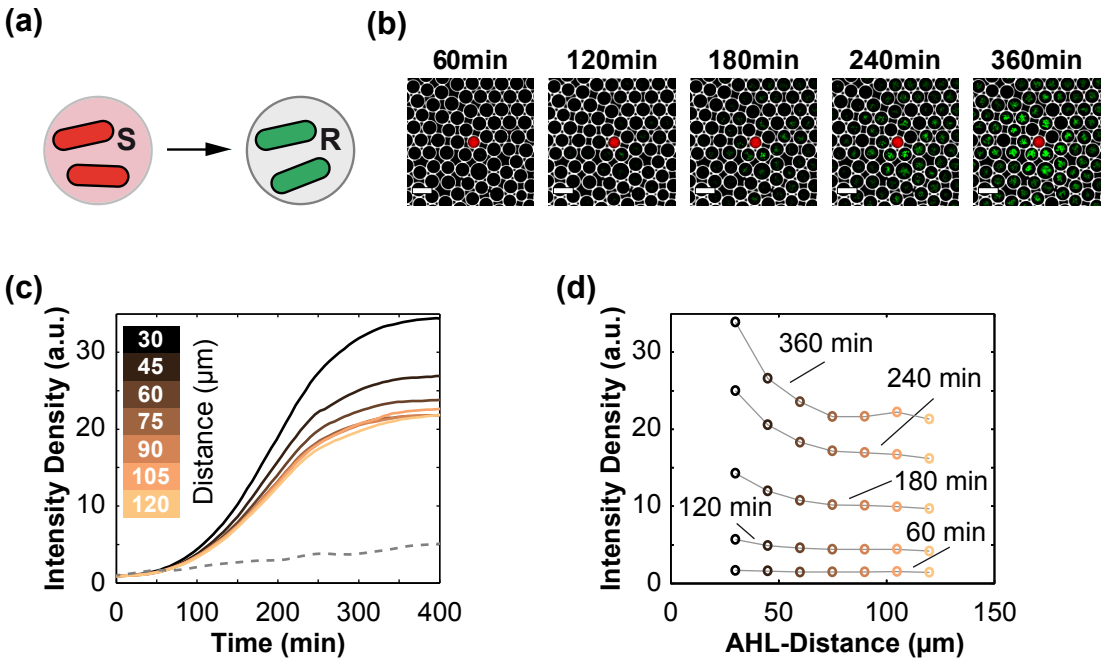

Figure 5.10: Communication mode: AHL sender - AHL receiver. (a) Simple scheme of the experiment. Induction of GFP expression is studied in droplets containing AHL-receiver bacteria in presence of droplets filled with sender bacteria. (b) Image series of a microscopy time-lapse movie of droplets containing receiver cells surrounding an AHL-sender droplet. White circles represent droplet boundaries derived from inverted BF images. The scale bar is 25 µm. (c) Evolution of GFP fluorescence in receiver cell containing droplets grouped according to their distance from the next sender droplet. An appropriately scaled fluorescence time trace recorded in a control experiment with not-induced receiver cells is included for comparison (dashed line). (d) Fluorescence profile of the time traces shown in (c) for different times. The fluorescences increases for all distances since AHL is not depleted due to permanent production by sender bacteria.

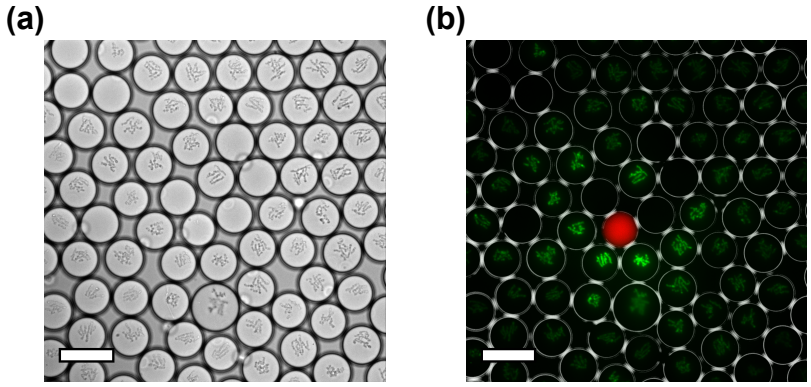

Figure 5.11: AHL sender and receiver bacteria in droplets recorded in 40x magnification in (a) bright field and (b) fluorescence mode 18 h after start of the experiment. White circles representing droplet boundaries derived from inverted bright field images. The scale bar is 50 µm.

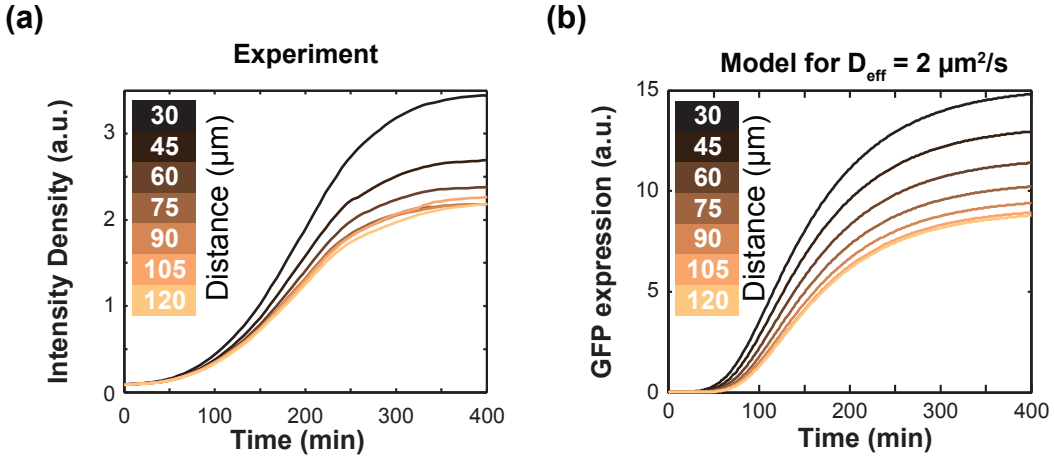

Figure 5.12: Experiment and simulation for AHL sender and AHL receiver bacteria. When inducer reservoirs are exchanged by permanently producing inducer sources, the reaction-diffusion model matches with the experiment assuming a reduced effective diffusion constant of $D_{eft} = 2\,\mu m^2/s$ consistently with previous results in the reservoir-receiver communication mode.

without sender or AHL reservoirs (grey dashed line) shows the extent to which GFP expression is induced when sender are present. We further plotted the fluorescence intensity profiles for different times as shown in Fig. 5.10 (d). A bright field as well as a fluorescence microscopy image in higher magnification of sender and receiver bacteria in microdroplets is shown in Fig. 5.11. As before for reservoir droplets, we modelled the reaction-diffusion model using the the effective diffusion coefficient found for AHL reservoirs and AHL receiver of $D_{eff} = 2\,\mu m^2/s$. The model can be easily derived by adding a term for constant AHL production to equation 5.3. The qualitative agreement between experimental traces (Fig. 5.12 (a)) and simulated traces (Fig. 5.12 (b)) confirms the assumptions made as reasonable. Overlapping fluorescence traces for droplets $75 - 120\,\mu m$ distant from a sender droplet are presumably due to contributions of next-nearest neighbour droplets filled with sender cells (Fig. 5.12).

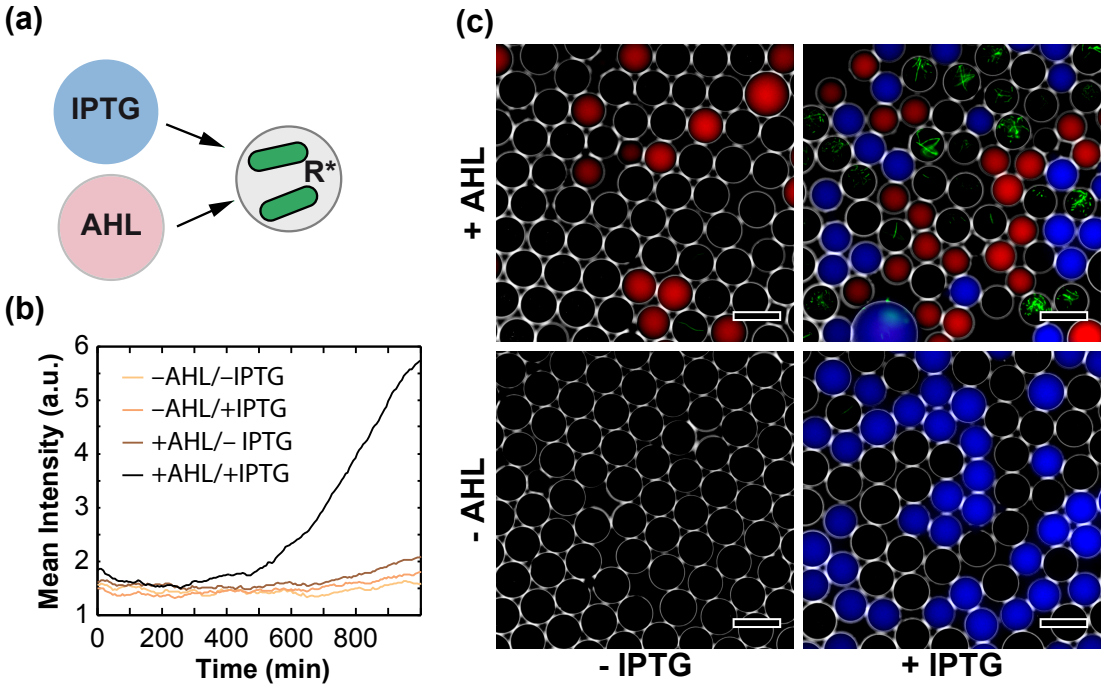

Figure 5.13: AND-gate bacteria in droplets. (a) Simple scheme of the experiment. Induction of gene expression of AND-gate bacteria in droplets is studied in presence of different combinations of AHL and IPTG reservoirs. (b) Fluorescence evolution of AND-gate bacteria in droplets in different observation channels. While AND-gate bacteria do not significantly respond with fluorescence output when no inducer or only one inducer species in droplets is present, fluorescence increases considerably when both are present. (c) Fluorescence microscopy images are overlaid by inverted bright field images to visualise droplet boundaries. The images are arranged in a truth table like fashion for the inputs AHL (+) or no AHL (-) and IPTG (+) or no IPTG (-). Images were taken 18 h after the experiment was initiated. Fluorescence is only high when both inducer droplets are present.

5.3.6 Communication mode: AHL - IPTG - AND-gate receiver

We next encapsulated bacteria containing the AND-gate circuit introduced in section 5.2.3 in microfluidic droplets. Integration of two chemical signals from AHL and IPTG reservoirs was studied (Fig. 5.13). As the transfer function determined in bulk measurements predicted induction thresholds of about one order of magnitude lower than compared to receiver cells, for both AHL and IPTG, we used only 20 nM AHL and 200 µM IPTG to fill reservoir droplets. For visualisation we supplemented AHL and IPTG reservoirs with 1 µM TAMRA and 1 µM ATTO 655, respectively. When droplets were generated, we literally poured the emulsion into observation channels. Consequently, the ratio in numbers of droplet types (AND-gate, AHL, or IPTG) as well as the arrangement were random. Unfortunately, we never achieved a mixing ratio for which a distance dependent evaluation of the gene expression behaviour in droplets was possible. To do so a more elaborate droplet arrangement technique would be required. Instead, we observed the behaviour of a population of droplets which contained AND-gate bacteria in absence of reservoir droplets, in presence of reservoir droplets with only one kind of inducer, and in presence of both kinds of reservoir droplets. AND-gate bacteria droplets in microscopy time-lapse movies were tracked and data of droplets was extracted. Since reservoir droplet densities were too high, a distance dependent analysis was not feasible. Instead, we evaluated the

overall response for absence of reservoirs, presence of reservoirs with only one kind of inducer and presence of both inducer reservoirs. We therefore determined the mean fluorescence intensity time development of droplets for each input combination. In Fig. 5.13 (b) the fluorescence traces for different input combinations are plotted. For AND-gate droplets which were isolated from reservoir droplets as well as for AND-gate droplets in presence of just one type of inducer reservoir fluorescence remained low, hence bacteria did not respond with fluorescence output. However, when reservoirs of both inducers were present, AND-gate bacteria integrated the diffusively transduced signals and responded with a strong increase in fluorescence as desired. Fig. 5.13 (c) shows fluorescence microscopy images taken 18 h after start of the experiment, arranged in a truth table like manner according to the combination of reservoir droplets present. In agreement with time courses, GFP fluorescence in AND-gate bacteria is only high when both AHL and IPTG droplets are present. Distance dependent data evaluation was impossible since almost all AND-gate droplets were in direct contact with reservoir droplets. Preliminary results with lower overall droplet densities hence larger droplet-to-droplet distance but higher inducer concentrations indicated that spatial effects can be observed with the AND-gate bacteria. For this experiment reservoirs were filled with 200 nM AHL and 10 mM IPTG. We characterised the spatial response of 466 droplets containing AND-gate bacteria as a function of their distance to the next nearest AHL or IPTG reservoir droplets (of which there were 265 and 87 in total, respectively). In Fig. 5.14 (a) scatter plots for different time points of the experiment show the fluorescent state of AND-gate bacteria in dependence of their distance to the next AHL and IPTG reservoir droplet. Each dot indicates a droplet's position relative to the next reservoirs. The fluorescence intensity of each droplet is colour coded. The diagrams are sub-divided into four areas which are close to both kinds of reservoirs, close to either AHL or IPTG reservoir and far from both (grey lines). Especially in scatter plots for $t = 495$ min and $t = 660$ min higher fluorescence intensities appear in the region close to both reservoirs. We further analysed these results by evaluating the percentage of droplets for each quadrant whose fluorescence intensity crossed the threshold value 0.1. The percentage of droplets which crossed the threshold value were plotted as a function of time (Fig. 5.14 (b)). This representation reveals that indeed gene expression is first activated in droplets in the region close to both AHL and IPTG. Due to high inducer concentration in reservoirs, the spatial effect is only transient and eventually also bacteria in droplets further apart start to express GFP. GFP expression slows down and the percentage of induced cells saturates at approximately 900 min. At this time almost 90% of droplets close to reservoirs started considerably to fluoresce whereas for droplets close to IPTG, close to AHL, and at remote positions, in only 60%, 50%, and 40% of droplets the threshold was exceeded.

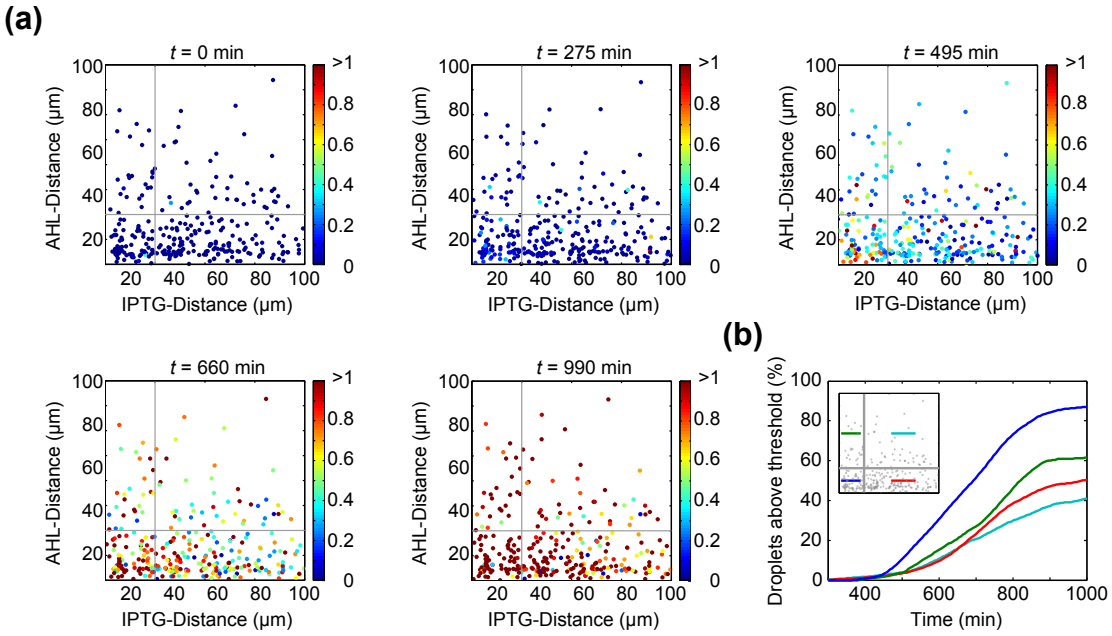

Figure 5.14: Distance dependent gene induction in AND-gate bacteria in presence of reservoir droplets filled with 200 nM AHL and 10 mM IPTG, respectively. (a) Scatter plot of droplets containing AND-gate bacteria in dependence of their distance to the next AHL and IPTG reservoir. Fluorescence intensities of droplets containing bacteria are colour coded. The scatter plot is divided into four sub-areas by grey lines. (b) Induction time course of the percentage of droplets which show a higher fluorescence value than 0.1 for each of the four areas in (a). Fluorescence of droplets which are close to both AHL and IPTG starts earlier to increase than in the other quadrants.

5.4 Conclusion and outlook

We have shown that amphiphilic inducer molecules such as the quorum sensing signal 3OC6HSL or the conventional inducer IPTG can establish chemical communication between inducer reservoirs and small groups of bacteria encapsulated within water-in-oil microemulsion droplets. The distance dependent gene induction in droplet arrays where reservoir or sender droplets were surrounded by droplets filled with receiver cells could be explained by means of a effective diffusion coefficient which is reduced by orders of magnitude compared with typical values of $D_0 = 100 - 1000\,\mu m^2/s$ for small compounds in aqueous solutions. The reduction in diffusivity cannot be explained by the viscosity of the FC-40 oil which is a four times higher value than for water. However, similar problems are known for systems with periodically arranged membranes separated by a medium, like biological tissue. More than 40 years ago, Francis Crick argued for morphogens migrating diffusively through periodically arranged cells will experience an effectively reduced diffusion coefficient which follows

$$\frac{1}{D_{eff}} = \frac{1}{D_0} + \frac{1}{\kappa a}, \tag{5.13}$$

where D_0 is the bulk diffusion coefficient, κ is the permeability of membranes and a is the periodic spacing in between [96, 97]. For low permeabilities and relatively high bulk diffusion coefficients the second term in 5.13 becomes dominating and the expression simplifies to $D_{eff} = \kappa a$. An alternative assumption which leads to a reduced diffusion coefficient is when diffusive exchange of molecules between droplets only hap-

pens through surfactant double layers, thus through the contact area of neighbouring droplets. In this case, the diffusion coefficient in free space would be decreased proportional to r_d/r_c, where r_d is the droplet radius and r_c is the radius of the contact area between droplets [222]. Both scenarios seem reasonable. The effective diffusion coefficient for AHL was larger than for IPTG which is plausible due to their different octanol/water partitioning coefficients. Similar communication scenarios may be presumably realised with other amphiphilic compounds, like antibiotics. Chloramphenicole, for instance, has a partitioning coefficient of $\log K_{ow} = 1.14$ which is on the same order as AHL and may partition partly into the oil phase as well [200]. We showed that engineered computational bacteria can be used to integrate several chemical signals sent out from distinct reservoir droplets in a context-dependent manner. As natural quorum sensing occurs in complex heterogenous environments such as biofilms [205, 221], microemulsions may actually be used as model systems to study these decentralised decision making processes in a topologically more realistic medium. Apart from this, a small, potentially tuneable diffusivity for bacterial communication could be of considerable interest for applications in synthetic biology. Several studies have explored genetically programmed structure formation in the past [32, 214], resulting in patterns on a millimetre length scale. Tuning of the diffusion constant to smaller values would reduce the patterning length scale by a factor $(D_0/D_{eff})^{1/2}$, which is of order ≈ 10 in our case. An additional interesting feature of a droplet-based bacterial communication system is the fact that communication takes place between small and spatially separated groups of bacteria. In principle, different types (even species) of bacteria with potentially different environmental requirements could be prepared in different droplet microreactors, and their interactions studied without physical contact of the cells themselves. Another aspect of compartmentalization is the fact that chemical signals are sent out or received by small groups of bacteria, which is expected to average out fluctuations caused by single cell variability [33]. Previous studies on distributed computing based on spatially separated, communicating microcolonies [41, 223] have emphasised the fact that such averaging makes bacterial computing more robust than in alternative concepts based on intracellular (single cell) computing. A distributed bacterial computer implemented in microemulsions potentially could be faster and operate with reduced space requirements.

5.5 Materials and methods

5.5.1 Materials

Bacterial strains and plasmids

AHL Sender: The gene from *A. fischeri* for *luxI* synthase (BioBrick part BBaC0261) was cloned into pETDuet−1 expression vector (Merck Millipore) between BioBrick cloning sites *XbaI* and *PstI*. The multiple cloning sites (MCS) of the vector precede the IPTG inducible T7 promoter P_{T7lac}. In the compatible host strain *E. coli* BL21(DE3)pLysS (Promega), expression is driven by T7 RNA polymerase.
AHL receiver: The BioBrick part BBaT9002 on pSB1A3 plasmid was transformed to *E. coli*BL21(DE3)pLysS (Promega).
IPTG receiver: The IPTG receiver plasmid was identical to the AHL sender plas-

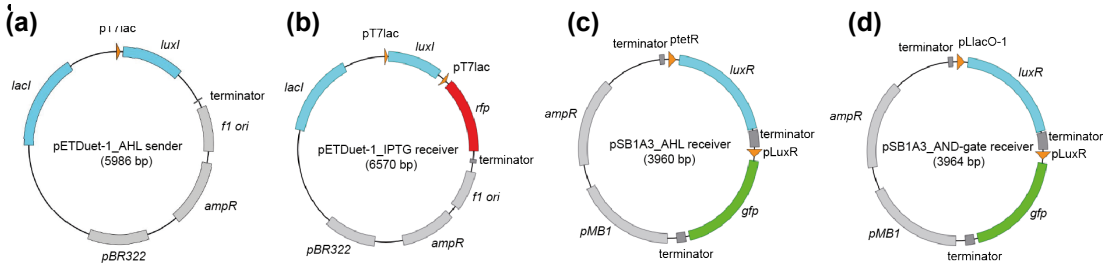

Figure 5.15: Plasmid maps

mid except for an additional gene for RFP. The monomeric reporter gene *rfp* was amplified directly from the template (BioBrick part BBaE1010) by polymerase chain reaction (PCR) with primers (rfp_fwd, rfp_rev) that comprise *NdeI* and *PacI* restriction sites. We performed double digestion of both, the PCR fragment and the plasmid and ligated them according to standard protocols.
AND-gate: T9002 cells were modified to AND receiver cells. Here, we put *luxR* gene under control of pLlacO−1 [224] and *gfp* under control of plux. To achieve the logic operation, T9002pSB1A3 plasmid was linearised by deletion PCR with ptet_del primers followed by circular polymerase extension cloning (CPEC) [225]. CPEC was used to insert placO−1 promoter with the overlapping pLlacO−1 [224] primers.

5.5.2 Methods

Characterisation of genetic circuits

The characterisation of genetic circuits was the topic of Ronja Berg's bachelor's thesis and is described there in detail [226]. This is a brief summary of the procedure. Characterisation measurements were performed on a Fluorostar Omega BMG LABTECH plate reader which measures both fluorescence and absorbance of up to 96 samples in a well-plate. Excitation and emission wave-lengths for fluorescence are selected by appropriate filter sets. Transfer curves were determined by recording time dependent intensities $I_i(t)$ and absorbance at 600 nm $OD600_i(t)$ for individual wells i. The promoter activity is given by

$$\alpha_i(t) = \frac{1}{OD600_i(t)} \left(\frac{dI_i(t)}{dt} \right), \tag{5.14}$$

where $\alpha_i(t)$ is the time dependent promoter activity [148]. We characterised the degree of induction by means of the maximal value of $\alpha_i(t)$. Maximum activities were then plotted as a function of inducer concentration and characteristic parameters, thus Hill-coefficient n and threshold concentration K were derived by a fit of 2.85 for one inducer and 2.87 for AND-gate measurements.

Generation of emulsions

Bacteria and inducer containing droplets were produced in microfluidic chips as described in section 3.5. Droplets with different contents were generated sequentially using new PDMS chips for each "substance" to avoid cross-contamination. Typically, the order of encapsulation was positive controls (initially induced bacteria),

not-induced bacteria, and finally inducer containing droplets to reduce inducer "leakage" into the oil phase to a minimum before measurements were started. During generation, droplets were stored in a "reservoir tubing" connected to the outlet of the microfluidic device. When sufficient emulsion was produced the tubing was disconnected from the chip and droplets were poured into observation channels on an ibidi microslide. The procedure was repeated for a second or third kind of droplet content (dependent on the communication mode). Droplets were then mixed in the observation channels by careful spinning of the ibidi slide. To avoid evaporation, observation channels were finally sealed with PCR tape.

Fluorescence time-lapse microscopy

Fluorescence time-lapse microscopy was performed on an Olympus IX-81 (section 3.3). The fully-automated setup enabled the observation of $10 - 12$ different positions on the observation slide. Depending on the number of positions, images were acquired in $5-6$ min intervals. At each position, images were taken in bright field and fluorescence modes. Fluorescence images of an observation channel filled with oil-surfactant mix were subtracted from droplet images to correct for the oil's auto-fluorescence.

Image analysis and data processing

Tracking of droplets in time-lapse movies and further data processing was done with a custom written algorithm in Matlab by K. Kapsner. The image analysis used a similar concept as described in chapter 4 but with considerable faster performance. Tracked droplets were filtered according to their shape by the measure

$$x = \frac{2 \cdot \pi \cdot r}{S},$$

where x represents the "circularity" which is 1 for a circle, r and S are the droplets radius and perimeter, respectively, in pixel determined by the segmentation function. Objects with $x \leq 0.8$ were discarded since low "circularity" was an indicator that recognised objects were not single droplets. Valid droplets were then classified according to their fluorescence intensity. Constant low fluorescence intensity indicated empty (bacteria) droplets. Bacteria in droplets which were not induced expressed at a basal level high enough to be distinguished from empty droplets. Inducer reservoirs were filled with a reference dye in a different colour than the fluorescent proteins expressed by the bacteria.

6 Appendix: A biochemical *in vitro* oscillator

6.1 Extended model reactions

In this section, only the reactions included in the extended model from [3] are listed. The extended model involves DNA and RNA hybridisation and strand displacement reactions and enzyme reactions as shown in Fig. 4.2 (b). Additionally, it accounts for reactions involving short RNA strands which accumulate with time due to incomplete degradation by RNase H. For a detailed description of the model or the extension with insulator and tweezer, the reader is referred to [3, 64] or [63, 173], respectively.

Hybridisation and strand displacement reactions:

$$
\begin{aligned}
T12 + A2 &\xrightarrow{k_{TA,12}} T12:A2 \\
T21 + A2 &\xrightarrow{k_{TA,21}} T21:A1 \\
A1 + dI1 &\xrightarrow{k_{AI,1}} A1:dI1 \\
rA1 + dI1 &\xrightarrow{k_{rAI,1}} rA1:dI1 \\
A2 + rI2 &\xrightarrow{k_{AI,2}} A2:rI2 \\
T12:A2 + rI2 &\xrightarrow{k_{TAI,12}} T12 + A2:rI2 \\
T21:A1 + dI1 &\xrightarrow{k_{TAI,21}} T21 + A1:dI1 \\
rA1 + A1:dI1 &\xrightarrow{k_{TA,12}} rA1:dI1 + A1 \\
A2 + sI2 &\underset{k_{s,-}}{\overset{k_{s,+}}{\rightleftharpoons}} A2:sI2 \\
T12 + A2:sI2 &\xrightarrow{k_{TA,12}} T12:A2:sI2 \\
T12:A2 + sI2 &\underset{k_{s,-}}{\overset{k_{s,+}}{\rightleftharpoons}} T12:A2:sI2 \\
rI2 + A2:sI2 &\xrightarrow{k_{AI,2}} A2:rI2 + sI2
\end{aligned}
\tag{6.1}
$$

Enzyme reactions:

$$
\begin{aligned}
RNAP + T12:A2 &\underset{k_{-,on,12}}{\overset{k_+}{\rightleftharpoons}} RNAP:T12:A2 &&\xrightarrow{k_{cat,on,12}} RNAP + T12:A2 + rA1 \\
RNAP + T21:A1 &\underset{k_{-,on,21}}{\overset{k_+}{\rightleftharpoons}} RNAP:T21:A1 &&\xrightarrow{k_{cat,on,12}} RNAP + T21:A1 + rI2 \\
RNAP + T12 &\underset{k_{-,off,12}}{\overset{k_+}{\rightleftharpoons}} RNAP:T12:A2 &&\xrightarrow{k_{cat,off,12}} RNAP + T12:A2 + rA1 \\
RNAP + T21 &\underset{k_{-,off,21}}{\overset{k_+}{\rightleftharpoons}} RNAP:T21:A1 &&\xrightarrow{k_{cat,off,12}} RNAP + T21:A1 + rI2 \\
RNAP + T12:A2:sI2 &\underset{k_{-,on,12}}{\overset{k_+}{\rightleftharpoons}} RNAP:T12:A2:sI2 &&\xrightarrow{k_{cat,off,12}} RNAP + T12:A2:sI2 + rA1 \\
RNaseH + rA1:dI1 &\underset{k_{-,H,1}}{\overset{k_{+,H}}{\rightleftharpoons}} RNaseH:rA1:dI1 &&\xrightarrow{k_{cat,H,1}} RNaseH + dI1 \\
RNaseH + A2:rI2 &\underset{k_{-,H,2}}{\overset{k_{+,H}}{\rightleftharpoons}} RNaseH:A2:rI2 &&\xrightarrow{k_{cat,H,2}} RNaseH + A2:sI2
\end{aligned}
\tag{6.2}
$$

6.2 Model parameters

	i=2, j=1	i=1, j=2	Lower limit	Upper limit	Other studies
k_+ (/M/s)	1.90×10^5	–	10^5	10^7	–
$k_{-,ON,ij}$ (/s)	0.0446	0.0634	0.01	0.3	–
$k_{cat,ON,ij}$ (/s)	0.0186	0.0471	0.01	0.1	0.73–1.12
$k_{-,OFF,ij}$ (/s)	0.35	0.388	0.01	1	–
$k_{cat,OFF,ij}$ (/s)	0.0033	0.0131	0.001	0.03	0.11–0.18
$k_{+,H}$ (/M/s)	7.13×10^5	–	10^5	10^7	–
$k_{-,H,j}$ (/s)	0.125	0.0692	0.01	1	–
$k_{cat,H,j}$ (/s)	0.554	0.463	0.01	1	0.02–0.6
$k_{TA,ij}$ (/M/s)	1.59×10^5	9.47×10^3	3×10^3	3×10^5	$0–3\times10^6$
$k_{AI,j}$ (/M/s)	4.69×10^3	5.66×10^4	3×10^3	3×10^5	$0–3\times10^6$
$k_{TAI,ij}$ (/M/s)	1.92×10^4	5.05×10^3	3×10^3	3×10^5	$0–3\times10^6$
$k_{rAI,j}$ (/M/s)	1.50×10^4	–	3×10^3	3×10^5	$0–3\times10^6$
$k_{AIrA,j}$ (/M/s)	1.65×10^4	–	3×10^3	3×10^5	$0–3\times10^6$
$k_{s,+}$ (/M/s)	1.65×10^4	–	3×10^3	3×10^5	–
$k_{s,-}$ (/s)	0.0525	–	0.01	1	–

Table 6.1: Parameter ranges and results for model fits for the extended model

6.3 Oscillator tunings

For sustained oscillations (reaction #12) in droplets discussed in section 4.4.2, we used *E. coli* RNase H and T7 RNA polymerase in a volume relation of 1:10 . T7 RNA polymerase (Epicentre, TM910K) was added to the reaction mix to a final concentration of 12.15 $^{\text{units}}/_{\mu\text{L}}$. For RNAP, the nominal concentration of enzyme stock for the droplet batch was 3.6 µM according to the manufacturer. For RNase H, the nominal concentration of enzyme stock for our specific batch was not available from the manufacturer. By analysing information on several other batches from the same vendor, which reported specific activity for RNase H ranging from 500,000–1,000,000 $^{\text{units}}/_{\text{mg}}$, we concluded that a reasonable nominal concentration of RNase H stock is 0.5–1 µM. We used 0.646 µM as a reasonable estimate for RNase H stock concentration.

Table 6.2: Oscillator tunings for reactions #1 – #30. The RNAP and RNase H contents are either given in their volume ratio as present in the sample which has to be multiplied by the batch's unit concentration 200 $^{\text{units}}/_{\mu\text{L}}$ or converted into nM depending on the availability of the specific activity. For RNase H the vendor refused to provide us with the specific activity. Hence values listed are $^{\text{units}}/_{\mu\text{L}}$ except for reactions #9, #10, and #12 where an estimate based on our model fits and on information available for other batches is given. For reactions #1 – #12 and #57 the RNAP stock concentration was [RNAP] = 3.6μM. For reactions #17 – #20, #25 – #36, #37 – #47, #48 – #56, and #58 – #59 the RNAP stock concentration was [RNAP] = 5.05μM. For reactions #13 – #16 performed at Caltech and #21 – #24 performed at UC Riverside no information about the enzyme stock concentrations is available.

Reaction #	T21 (nM)	A1 (nM)	dI1 (nM)	T12 (nM)	A2 (nM)	RNAP (200 $^{\text{units}}/_{\mu\text{L}}$) / (nM)	RNase H (10^{-2} $^{\text{units}}/_{\mu\text{L}}$)
1	150	150	500	80	250	0.0607/ 218	6.07
2	150	150	500	80	250	0.0572/ 206	5.72
3	150	150	500	80	250	0.0607/ 218	6.07
4	150	150	500	80	250	0.0607/ 218	6.07
5	150	150	500	80	250	0.0572/ 206	9.53
6	150	150	500	80	250	0.0572/ 206	7.14
7	150	150	500	80	250	0.0572/ 206	5.72
8	150	150	500	80	250	0.0572/ 206	5.14
9	150	150	500	80	250	0.0304/ 109	4.3 (*3*)
10	150	150	500	80	250	0.0607/ 218	7.59 (*5*)
11	150	150	500	80	250	0.0607/ 218	6.68
12	150	150	500	80	250	0.0607/ 218	6.7 (*4*)
13	150	200	500	30	350	0.0271	5.429
14	150	200	500	60	350	0.0271	5.429
15	150	200	500	90	350	0.0271	5.429
16	150	200	500	120	350	0.0271	5.429
17	150	200	700	120	350	0.0343/ 173	8.57
18	150	200	700	120	350	0.0367/185	6.14
19	150	200	700	120	350	0.0343/ 173	8.57
20	150	200	700	120	350	0.0367/185	6.14
21	150	150	500	80	250	0.0304	3.04
22	150	150	500	80	250	0.0607	6.07
23	150	150	500	80	250	0.0730	6.07
24	150	150	500	80	250	0.0730	6.07
25	150	250	700	80	350	0.0343/ 173	8.57
26	150	200	500	120	350	0.0343/ 173	8.57
27	150	200	700	120	350	0.0343/ 173	8.57
28	150	200	1000	120	350	0.0343/ 173	8.57
29	150	200	700	80	350	0.0343/ 173	8.57
30	150	200	1000	80	350	0.0343/ 173	8.57

Table 6.3: Oscillator tunings for reactions #31 – #59. The RNAP and RNase H contents are either given in their volume ratio as present in the sample which has to be multiplied by the batch's unit concentration 200 $^{\text{units}}/_{\mu\text{L}}$ or converted into nM depending on the availability of the specific activity. For RNase H the vendor refused to provide us with the specific activity. Hence values listed are $^{\text{units}}/_{\mu\text{L}}$.

Reaction #	T21 (nM)	A1 (nM)	dI1 (nM)	T12 (nM)	A2 (nM)	RNAP (200 $^{\text{units}}/_{\mu\text{L}}$) / (nM)	RNase H ($^{\text{units}}/_{\mu\text{L}}$)
31	150	200	1300	80	350	0.0343/ 173	8.57
32	150	200	1600	80	350	0.0343/ 173	8.57
33	150	200	700	80	350	0.0343/ 173	8.57
34	150	200	700	70	350	0.0343/ 173	8.57
35	150	200	700	60	350	0.0343/ 173	8.57
36	150	200	700	50	350	0.0343/ 173	8.57
37	150	200	700	80	250	0.0343/ 173	8.57
38	150	200	700	80	350	0.0343/ 173	8.57
39	150	200	700	80	500	0.0343/ 173	8.57
40	150	200	700	80	700	0.0343/ 173	8.57
41	150	150	700	80	250	0.0367/185	6.14
42	150	150	700	80	250	0.0571/ 288	5.72
43	150	150	500	80	250	0.0608/ 307	6.07
44	150	150	400	80	250	0.0367/185	6.14
45	150	150	700	80	250	0.0367/185	6.14
46	150	150	800	80	250	0.0367/185	6.14
47	150	150	1000	80	250	0367/185	6.14
48	150	150	500	80	250	0.0608/ 307	6.07
49	150	150	500	80	250	0.0608/ 307	6.07
50	150	150	500	80	250	0.0608/ 307	6.07
51	150	150	500	80	250	0.0571/ 288	5.72
52	150	150	500	80	250	0.0608/ 307	6.07
53	150	150	500	80	250	0.0608/ 307	6.07
54	150	150	500	80	250	0.0571/ 288	5.72
55	150	150	500	80	250	0.0571/ 288	5.72
56	150	150	500	80	250	0.0608/ 307	6.07
57	150	150	500	80	250	0.0608/ 307	6.07
58	150	150	500	80	250	0.0608/ 307	6.07
59	150	150	500	80	250	0.0608/ 307	6.07

6.4 Supplementary Figures

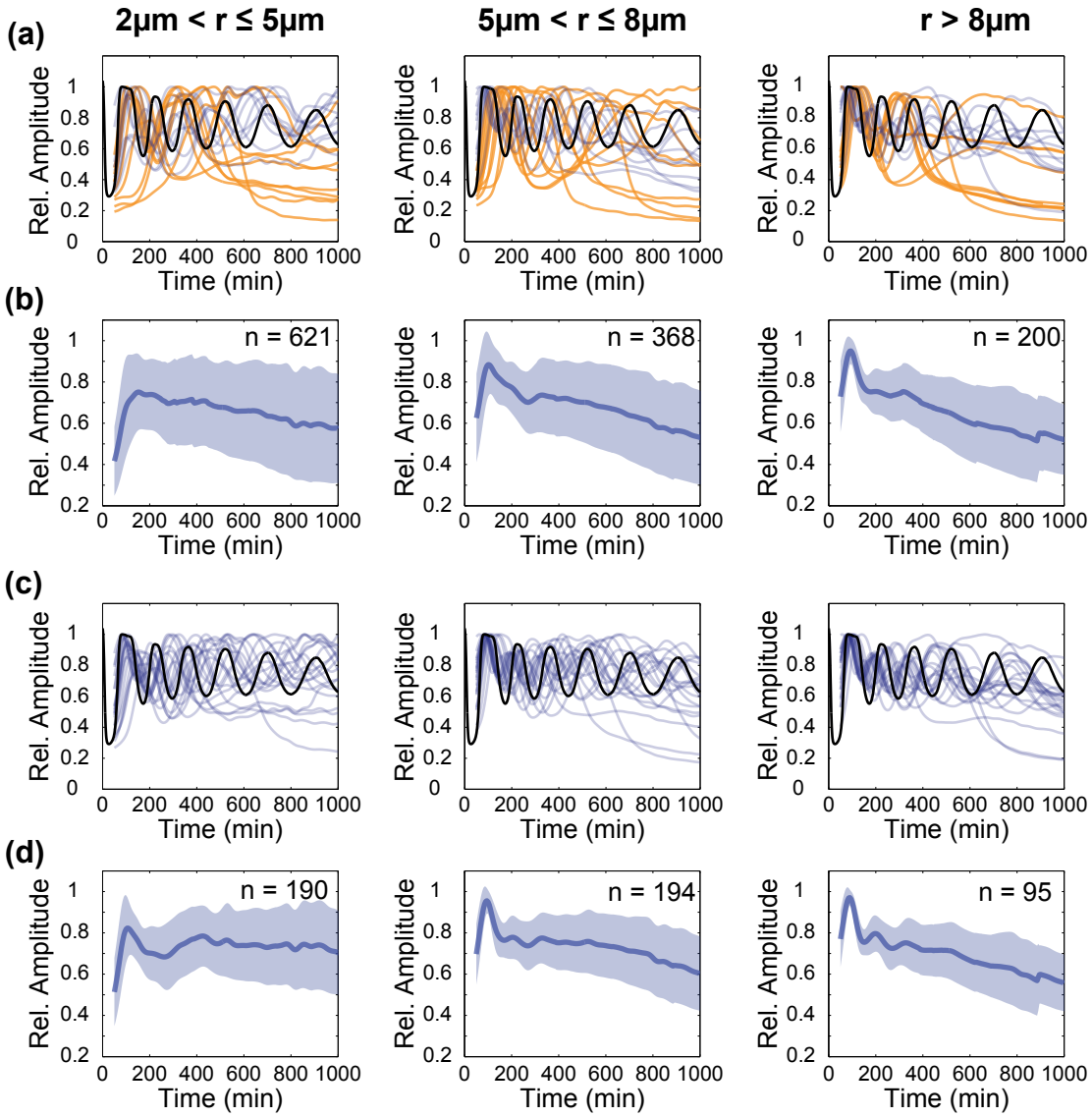

Figure 6.1: Example time traces and population mean for the sustained tuning (a) including traces classified as 'not identifiably oscillating' (orange) and (b) their population mean (blue trace) and standard deviation (blue shaded area), and (c) excluding traces classified as 'not identifiably oscillating' and (d) their population mean (blue trace) and standard deviation (blue shaded area). The numbers of traces considered for population mean and standard deviation n are drawn into the corresponding panels.

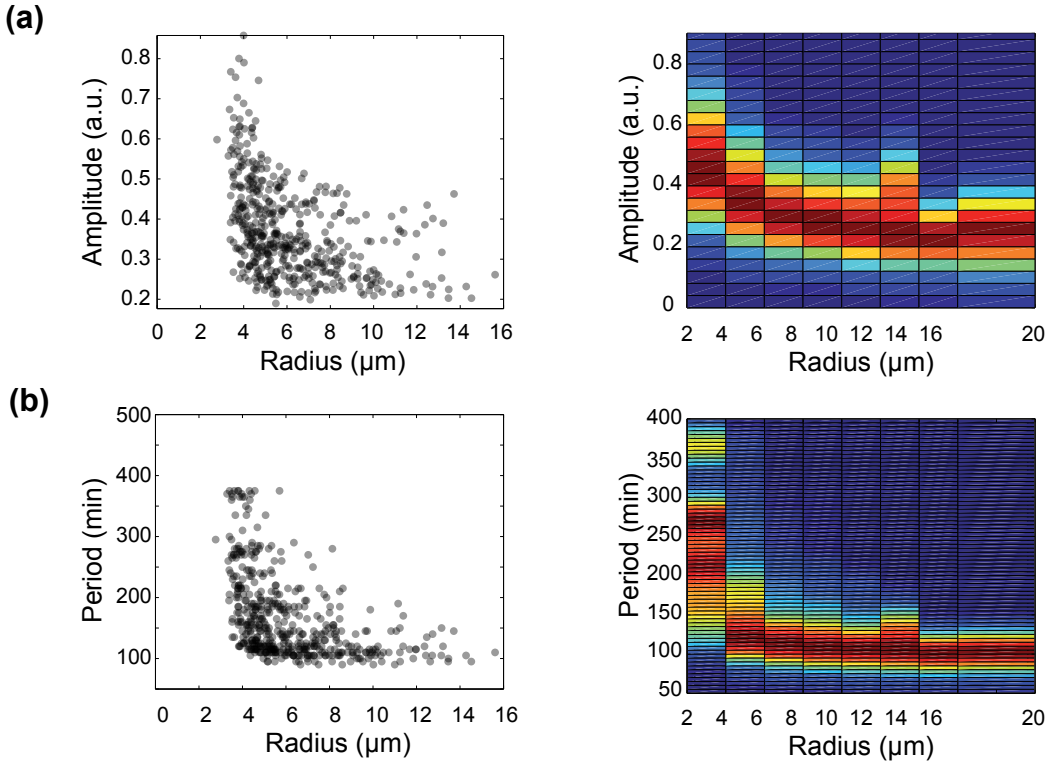

Figure 6.2: Scatter and colour plots for (a) amplitudes and (b) periods of the sustained oscillator tuning. Each dot in the scatter plot represent the analysis result for an individual droplet.

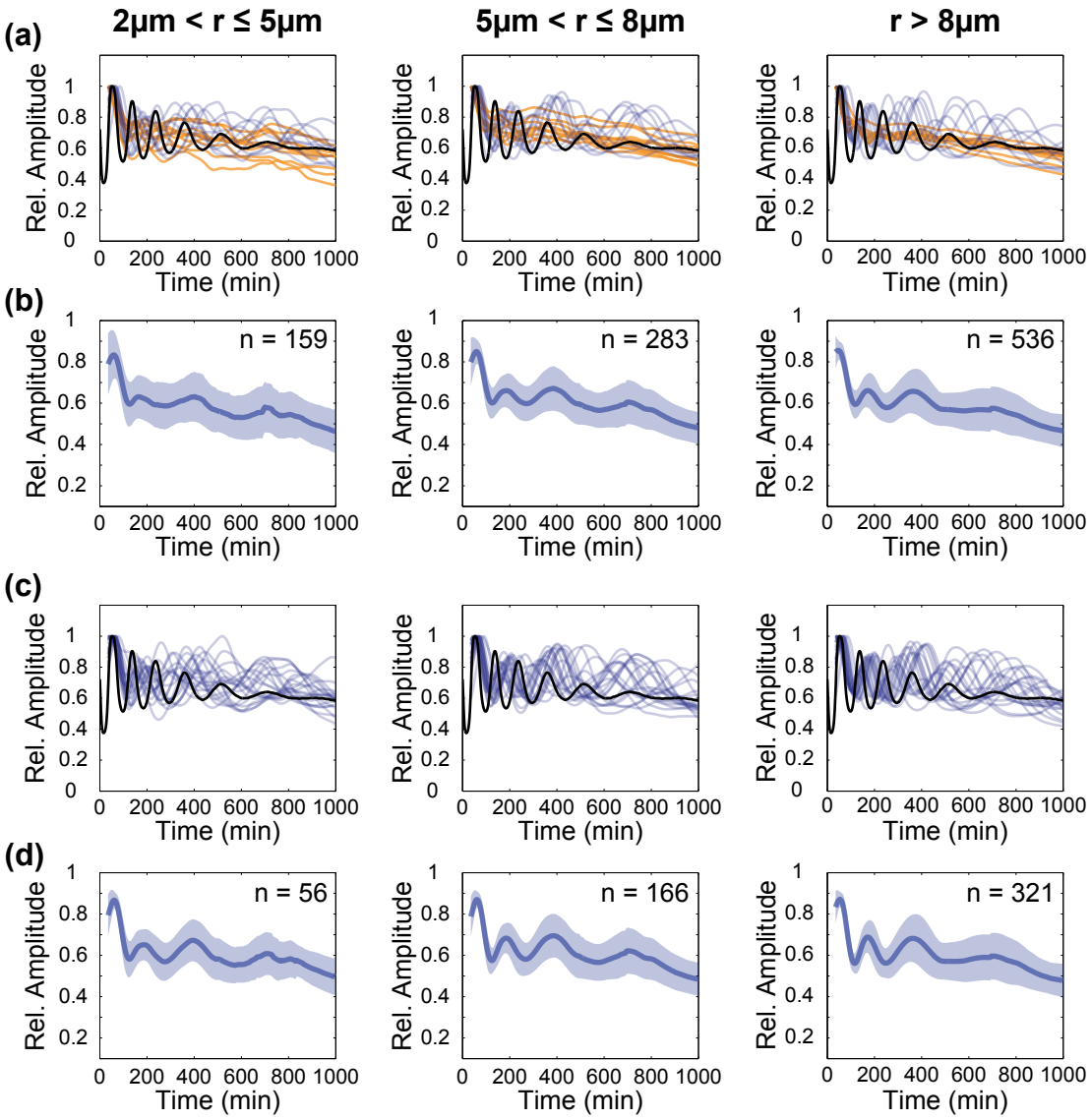

Figure 6.3: Example time traces and population mean for the damped tuning (a) including traces classified as 'not identifiably oscillating' (orange) and (b) their population mean (blue trace) and standard deviation (blue shaded area), and (c) excluding traces classified as 'not identifiably oscillating' and (d) their population mean (blue trace) and standard deviation (blue shaded area). The numbers of traces considered for population mean and standard deviation n are drawn into the corresponding panels.

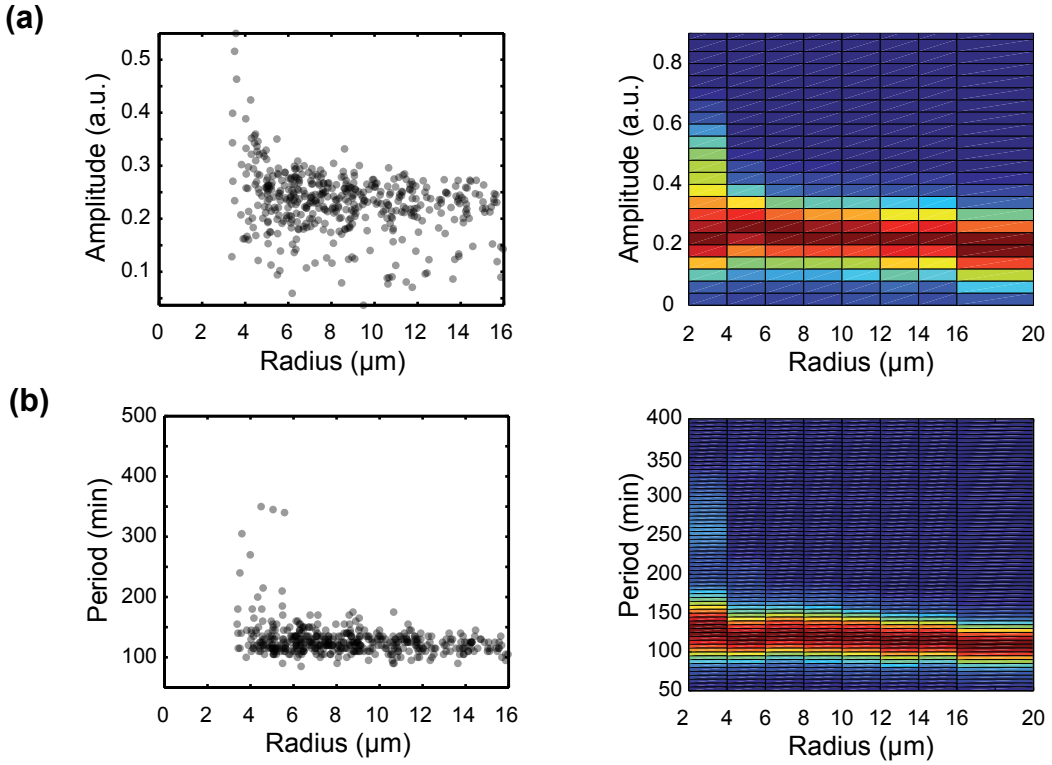

Figure 6.4: Scatter and colour plots for (a) amplitudes and (b) periods of the damped oscillator tuning. Each dot in the scatter plot represent the analysis result for an individual droplet.

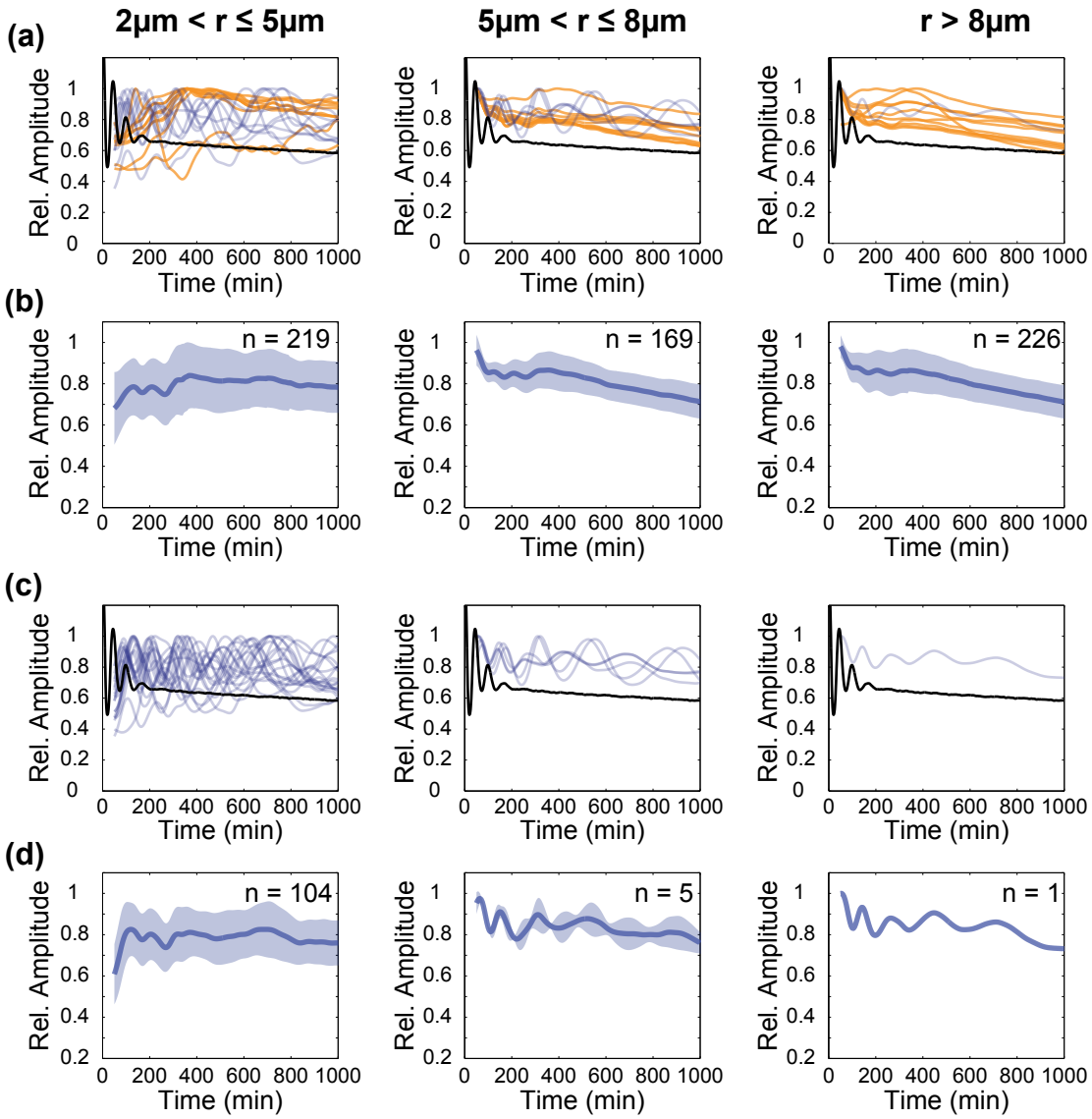

Figure 6.5: Example time traces and population mean for the strongly damped tuning (a) including traces classified as 'not identifiably oscillating' (orange) and (b) their population mean (blue trace) and standard deviation (blue shaded area), and (c) excluding traces classified as 'not identifiably oscillating' and (d) their population mean (blue trace) and standard deviation (blue shaded area). The numbers of traces considered for population mean and standard deviation n are drawn into the corresponding panels.

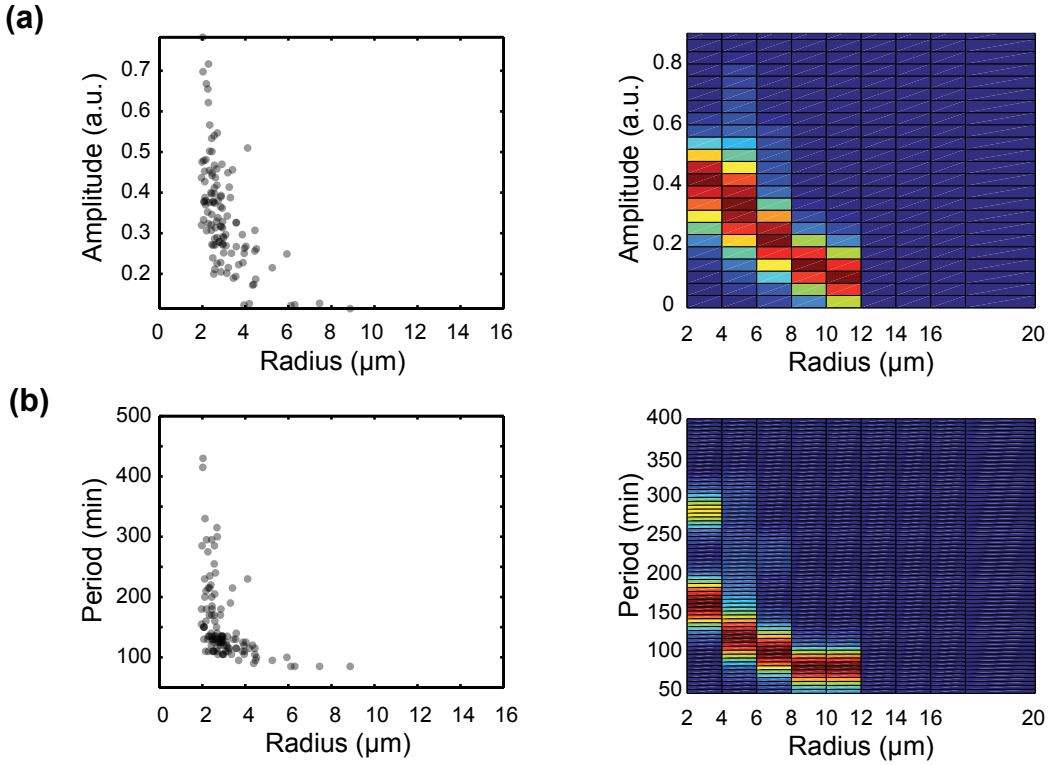

Figure 6.6: Scatter and colour plots for (a) amplitudes and (b) periods of the damped oscillator tuning. Each dot in the scatter plot represent the analysis result for an individual droplet.

6.5 Nucleic acid sequences

Table 6.4: Oscillator sequences

Name	Sequence
T12-t	5'-TTT CTG ACT TTG TCA GTA TTA GTG TGT AGT AGT AGT TCA TTA GTG TCG TTC GTT CTT TGT TTC TCC CTA TAG TGA GTC G
T12-nt	5'-AAG CAA GGG TAA GAT GGA ATG ATA ATA CGA CTC ACT ATA GGG AGA AAC AAA GAA CGA ACG ACA CTA ATG AAC TAC TAC TAC ACA CTA ATA CTG ACA AAG TCA GAA A
T21-t	5'-TTT CTG ACT TTG TCA GTA TTA TCA TTC CAT CTT ACC CTT GCT TCA ATC CGT TTT ACT CTC CCT ATA GTG AGT CG
T21-nt	5'-TexasRed-CAT TAG TGT CGT TCG TTC ACA GTA ATA CGA CTC ACT ATA GGG AGA GTA AAA CGG ATT GAA GCA AGG GTA AGA TGG AAT GAT AAT ACT GAC AAA GTC AGA AA
dI1	5'-GTG TGT AGT AGT AGT TCA TTA GTG TCG TTC GTT CAC AG
A1	5'-TAT TAC TGT GAA CGA ACG ACA CTA ATG AAC TAC TAC-Iowa Black RQ
A2	5'-TAT TAT CAT TCC ATC TTA CCC TTG CTT CAA TCC GT-Iowa Black RQ
rA1 (RNA)	5'-GGG AGA AAC AAA GAA CGA ACG ACA CUA AUG AAC UAC UAC UAC ACA CUA AUA CUG ACA AAG UCA GAA A
rI2 (RNA)	5'-GGG AGA GUA AAA CGG AUU GAA GCA AGG GUA AGA UGG AAU GAU AAU ACU GAC AAA GUC AGA AA

Table 6.5: Tweezer 1 and 2 sequences.

Name	Sequence
TwA	5'-Rhodamine green-TGC CTT GTA AGA GCG ACC ATC AAC CTG GAA TGC TTC GGA T/BHQ 1 -3'
TwB	5'-TCA AAT TTA CAA CGC ATC CGA AGC ATT CCA GGT
TwC	5'-GGT CGC TCT TAC AAG GCA AAC TAA CAT ATA ATC
Ins t	5'-AAG CAA GGG TAA GAT GGA ATG ATA ATA CGA CTC ACT ATA GGG AGA TCA AAT TTA CAA CGC AAC TAA CAT ATA ATC GAA GAC TTA ATA CTG ACA AAG TC
Ins nt	5'-GTG TGT AGT AGT AGT TCA TTA GTG TCG TTC GTT CAC AG
TwCl	5'-AAG TCT TCG ATT ATA TGT TAG TTG CGT TGT AAA TTT GA
rTwCl	5'-GGG AGA UCA AAU UUA CAA CGC AAC UAA CAU AUA AUC GAA GAC UUA AUA CUG ACA AAG UCA GAA A
TwA 2	5'-TAMRA- TAG GCT TCG TAA GGT CCA ACT ACC AGC GAG AAT GTT CCG T -BHQ 2
TwB 2	5'-TTC AGA AGC TAA TAT ACG GAA CAT TCT CGC TGG
TwC 2	5'-TGG ACC TTA CGA AGC CTA ACA ATC AAC GCA ACA
Ins 2 t	5'-TTT CTG ACT TTG TCA GTA TTT TCA GAA GCT AAT ATA CAA TCA ACG CAA CAT TTA AAC TCA CAC ATC TCT CCC TAT AGT GAG TCG
Ins 2 nt	5'-AAG CAA GGG TAA GAT GGA ATG ATA ATA CGA CTC ACT ATA GGG AGA GAT GTG TGA GTT TAA ATG TTG CGT TGA TTG TAT ATT AGC TTC TGA AAA TAC TGA CAA AGT C
TwCl 2	5'-AGT TTA AAT GTT GCG TTG ATT GTA TAT TAG CTT CTG AA
TwOp 2	5'-TTC AGA AGC TAA TAT ACA ATC AAC GCA ACA TTT AAA CTC ACA CAT C

Table 6.6: Aptamer insulator sequences

Name	Sequence
MG th Ins t	5'-TTC ATT AGG ATC CAT TCG TTA CCT GGC TCT CGC CAG TCG GGA TCC TCT CCC TAT AGT GAG TCG
MG th Ins nt	5'-AAG CAA GGG TAA GAT GGA ATG ATA ATA CGA CTC ACT ATA GGG AGA GGA TCC CGA CTG GCG AGA GCC AGG TAA CGA ATG GAT CCT AAT GAA
rMG th	5'-GGG AGA GGA UCC CGA CUG GCG AGA GCC AGG UAA CGA AUG GAU CCU AAU GAA
anti MG th Ins t	5'-GGA TCC CGA CTG GCG AGA GCC AGG TAA CGA ATG GAT CCT AAT GAA TCT CCC TAT AGT GAG TCG
anti MG th Ins nt	5'-TCA TTA GTG TCG TTC GTT CAC AGT AAT ACG ACT CAC TAT AGG GAG ATT CAT TAG GAT CCA TTC GTT ACC TGG CTC TCG CCA GTC GGG ATC C
r anti MG th	5'-GGG AGA UUC AUU AGG AUC CAU UCG UUA CCU GGC UCU CGC CAG UCG GGA UCC
DFHBI th Ins t	5'-AAG CAA GGG TAA GAT GGA ATG ATA ATA CGA CTC ACT ATA GGG AGA GAC GCG ACC GAA ATG GTG AAG GAC GGG TCC AGT GCT TCG GCA CTG TTG AGT AGA GTG TGA GCT CCG TAA CTG GTC GCG TCA AGT AAT C
DFHBI th Ins nt	5'-GAT TAC TTG ACG CGA CCA GTT ACG GAG CTC ACA CTC TAC TCA ACA GTG CCG AAG CAC TGG ACC CGT CCT TCA CCA TTT CGG TCG CGT CTC TCC CTA TAG TGA GTC G
rDFHBI th	5'-GGG AGA GAC GCG ACC GAA AUG GUG AAG GAC GGG UCC AGU GCU UCG GCA CUG UUG AGU AGA GUG UGA GCU CCG UAA CUG GUC GCG UCA AGU AAU C
anti DFHBI th Ins t	5'-GAC GCG ACC GAA ATG GTG AAG GAC GGG TCC AGT GCT TCG GCA CTG TTG AGT AGA GTG TGA GCT CCG TAA CTG GTC GCG TCA AGT AAT CTC TCC CTA TAG TGA GTC G
anti DFHBI th Ins nt	5'-TCA TTA GTG TCG TTC GTT CAC AGT AAT ACG ACT CAC TAT AGG GAG AGA TTA CTT GAC GCG ACC AGT TAC GGA GCT CAC ACT CTA CTC AAC AGT GCC GAA GCA CTG GAC CCG TCC TTC ACC ATT TCG GTC GCG TC
r anti DFHBI th	5'-GGG AGA GAU UAC UUG ACG CGA CCA GUU ACG GAG CUC ACA CUC UAC UCA ACA GUG CCG AAG CAC UGG ACC CGU CCU UCA CCA UUU CGG UCG CGU C

Table 6.7: Subsystem DNA sequences

Name	Sequence
Genelet-t	5'-TAG TTG TGA GTC GTA TTA ATT GAA TGG TTA TGT GGA TCT ATA GTG AGT CGT ATT AAT TGA AT
Genelet-nt	5'-ATT CAA TTA ATA CGA CTC ACT ATA GAT CCA CAT AAC CAT TCA ATT AAT ACG ACT CAC AAC TA
Rep-Q	5'-GTG AGT CGT ATT AAT TGA ATG GTT ATG T-BHQ2
Rep-F	5'-TAMRA-ATT CAA TTA ATA CGA CTC AC

Bibliography

[1] M. Elowitz, S. Leibler, *Nature* **403**, 335 (2000).

[2] J. Stricker, *et al.*, *Nature* **456**, 516 (2008).

[3] J. Kim, K. S. White, E. Winfree, *Molecular Systems Biology* **2**, 68 (2006).

[4] J. Kim, E. Winfree, *Molecular Systems Biology* **7** (2011).

[5] K. Montagne, R. Plasson, Y. Sakai, T. Fujii, Y. Rondelez, *Molecular Systems Biology* **7** (2011).

[6] A. C. Forster, G. M. Church, *Molecular Systems Biology* **2** (2006).

[7] S. Leduc, La biologie synthétique (1912).

[8] S. L. Duc, W. D. Butcher, *The mechanism of life* (Rebman Company, 1914).

[9] J. Loeb, *The dynamics of living matter*, Columbia University biological series. 8 (Columbia University Press ; The MacMillan Company, 1906).

[10] J. B. B. Burke, *The origin of life: its physical basis and definition* (Chapman and Hall, Limited, 1906).

[11] E. H. Szybalska, W. Szybalski, *Proceedings of the National Academy of Sciences of the United States of America* **48**, 2026 (1962).

[12] W. Szybalski, *Control of Gene Expression*, A. Kohn, A. Shatkay, eds. (Springer US, 1974), vol. 44 of *Advances in Experimental Medicine and Biology*, pp. 23–24.

[13] R. L. Rawis, *Chemical & Engineering News* **78**, 49 (2000).

[14] S. A. Benner, A. M. Sismour, *Nature Reviews Genetics* **6**, 533 (2005).

[15] F. R. Blattner, *et al.*, *Science* **277**, 1453 (1997).

[16] G. Bokinsky, *et al.*, *Proceedings of the National Academy of Sciences of the United States of America* **108**, 19949 (2011).

[17] D. R. Georgianna, S. P. Mayfield, *Nature* **488**, 329 (2012).

[18] C. A. Rabinovitch-Deere, J. W. K. Oliver, G. M. Rodriguez, S. Atsumi, *Chemical Reviews* **113**, 4611 (2013).

[19] M. H. Medema, R. Breitling, R. Bovenberg, E. Takano, *Nature Reviews Microbiology* **9**, 131 (2010).

[20] W. Weber, M. Fussenegger, *Nature Reviews Genetics* **13**, 21 (2011).

[21] J. D. Keasling, *Metabolic Engineering* **14**, 189 (2012).

[22] J. R. van der Meer, S. Belkin, *Nature Reviews Microbiology* **8**, 511 (2010).

[23] S. Jouanneau, M.-J. Durand, P. Courcoux, T. Blusseau, G. Thouand, *Environmental Science & Technology* **45**, 2925 (2011).

[24] F. Jacob, J. Monod, *Journal of Molecular Biology* **3**, 318 (1961).

[25] J. Monod, F. Jacob, *Cold Spring Harbor Symposia on Quantitative Biology* **26**, 389 (1961).

[26] D. Bray, *Nature* **376**, 307 (1995).

[27] T. S. Gardner, C. R. Cantor, J. J. Collins, *Nature* **20**, 339 (2000).

[28] D. T. Gillespie, *The journal of physical chemistry* **81**, 2340 (1977).

[29] M. B. Elowitz, A. J. Levine, E. D. Siggia, P. S. Swain, *Science* **297**, 1183–1186 (2002).

[30] U. Alon, M. G. Surette, N. Barkai, S. Leibler, *Nature* **397**, 168 (1999).

[31] A. Becskei, L. Serrano, *Nature* **405**, 590 (2000).

[32] S. Basu, Y. Gerchman, C. H. Collins, F. H. Arnold, R. Weiss, *Nature* **434**, 1130 (2005).

[33] A. Raj, A. van Oudenaarden, *Cell* **135**, 216 (2008).

[34] R. Weiss, T. F. j. Knight, *In DNA Computing, 6th International Workshop on DNA-Based Computers, DNA6* (2000).

[35] S. Basu, R. Mehreja, S. Thiberge, M.-T. Chen, R. Weiss, *Proceedings of the National Academy of Sciences of the United States of America* **101**, 6355 (2004).

[36] L. You, R. S. Cox, R. Weiss, F. H. Arnold, *Nature* **428**, 868 (2004).

[37] Y. Tanouchi, A. Pai, N. E. Buchler, L. You, *Molecular Systems Biology* **8** (2012).

[38] K. Brenner, D. K. Karig, R. Weiss, F. H. Arnold, *Proceedings of the National Academy of Sciences of the United States of America* **104**, 17300 (2007).

[39] W. Weber, M. Daoud-El Baba, M. Fussenegger, *Proceedings of the National Academy of Sciences of the United States of America* **104**, 10435 (2007).

[40] T. Danino, O. Mondragón-Palomino, L. Tsimring, J. Hasty, *Nature* **463**, 326 (2010).

[41] A. Tamsir, J. J. Tabor, C. A. Voigt, *Nature* **469**, 212 (2011).

[42] M. Weitz, *et al.*, *Journal of the American Chemical Society* **136**, 72 (2014).

[43] M. Thattai, A. Van Oudenaarden, *Proceedings of the National Academy of Sciences of the United States of America* **98**, 8614 (2001).

[44] J. M. Vilar, H. Y. Kueh, N. Barkai, S. Leibler, *Proceedings of the National Academy of Sciences of the United States of America* **99**, 5988 (2002).

[45] N. Rosenfeld, J. W. Young, U. Alon, P. S. Swain, M. B. Elowitz, *Science* **307**, 1962 (2005).

[46] J. T. Macdonald, C. Barnes, R. I. Kitney, P. S. Freemont, G.-B. V. Stan, *Integrative Biology* **3**, 97 (2011).

[47] Y. Shimizu, *et al.*, *Nature Biotechnology* **19**, 751 (2001).

[48] V. Noireaux, R. Bar-Ziv, A. Libchaber, *Proceedings of the National Academy of Sciences* **100**, 12672 (2003).

[49] V. Noireaux, A. Libchaber, *Proceedings of the National Academy of Sciences of the United States of America* **101**, 17669 (2004).

[50] J. Shin, V. Noireaux, *Journal of Biological Engineering* **4** (2010).

[51] J. Shin, V. Noireaux, *Journal of Biological Engineering* **4** (2010).

[52] E. Karzbrun, J. Shin, R. Bar-Ziv, V. Noireaux, *Physical Review Letters* **106**, 048104 (2011).

[53] J. Shin, V. Noireaux, *ACS Synthetic Biology* **1**, 29 (2012).

[54] J. Shin, P. Jardine, V. Noireaux, *ACS Synthetic Biology* **1**, 408 (2012).

[55] Z. Z. Sun, E. Yeung, C. A. Hayes, V. Noireaux, R. M. Murray, *ACS Synthetic Biology* (2013).

[56] F. C. Simmel, *ACS nano* **7**, 6 (2013).

[57] D. Y. Zhang, G. Seelig, *Nature Chemistry* **3**, 103 (2011).

[58] G. T. Walker, M. C. Little, J. G. Nadeau, D. D. Shank, *Proceedings of the National Academy of Sciences of the United States of America* **89**, 392 (1992).

[59] T. Fujii, Y. Rondelez, *ACS nano* (2013).

[60] K. Hasatani, *et al.*, *Chemical Communications* **49**, 8090 (2013).

[61] A. Padirac, T. Fujii, A. Estévez-Torres, Y. Rondelez, *Journal of the American Chemical Society* **135**, 14586 (2013).

[62] P. Subsoontorn, J. Kim, E. Winfree, *Arxiv preprint arXiv:1101.0723* (2011).

[63] E. Franco, *et al.*, *Proceedings of the National Academy of Sciences of the United States of America* **108**, E784 (2011).

[64] M. Weitz, *et al.*, *Nature Chemistry* **6**, 295 (2014).

[65] M. Loose, E. Fischer-Friedrich, J. Ries, K. Kruse, P. Schwille, *Science* **320**, 789 (2008).

[66] Y. Saga, H. Takeda, *Nature Reviews Genetics* **2**, 835 (2001).

[67] G. Dong, S. S. Golden, *Current opinion in microbiology* **11**, 541 (2008).

[68] M. Nakajima, *et al.*, *Science* **308**, 414 (2005).

[69] L. Michaelis, M. L. Menten, *Biochem. z* **49**, 352 (1913).

[70] G. E. Briggs, J. B. S. Haldane, *Biochemical journal* **19**, 338 (1925).

[71] D. D. Van Slyke, G. E. Cullen, *Journal of Biological Chemistry* **19**, 141 (1914).

[72] H. Pearson, *Nature* **441**, 398 (2006).

[73] K. H. Nealson, J. W. Hastings, *Microbiological Reviews* **43**, 496 (1979).

[74] G. K. Ackers, A. D. Johnson, M. A. Shea, *Proceedings of the National Academy of Sciences of the United States of America* **79**, 1129 (1982).

[75] U. Gerland, J. D. Moroz, T. Hwa, *Proc. Natl. Acad. Sci. U. S. A.* **99**, 12015 (2001).

[76] P. Erdos, A. Renyi, *Institute of Mathematics University of DeBreceniens, Debrecar, Hungary* (1959).

[77] S. S. Shen-Orr, R. Milo, S. Mangan, U. Alon, *Nature Genetics* **31**, 64 (2002).

[78] S. H. Strogatz, *Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering*, Studies in nonlinearity (Westview Press, 2008).

[79] A. Goldbeter, *Nature* **420**, 238 (2002).

[80] J. Mallet-Paret, H. L. Smith, *Journal of Dynamics and Differential Equaitons* **2**, 367 (1990).

[81] B. Novák, J. J. Tyson, *Nature Reviews Molecular Cell Biology* **9**, 981 (2008).

[82] J. D. Murray, *Mathematical Biology I: An Introduction* (Springer, 2002).

[83] M. B. Elowitz, A. J. Levine, E. D. Siggia, P. S. Swain, *Science* **297**, 1183 (2002).

[84] J. Paulsson, *Physics Of Life Reviews* **2**, 157 (2005).

[85] M. Weitz, F. C. Simmel, *Transcription* **3**, 87 (2012).

[86] J. W. Szostak, D. P. Bartel, P. L. Luisi, *Nature* **409**, 387 (2001).

[87] A. C. Forster, G. M. Church, *Genome Research* **17**, 1 (2006).

[88] P. E. M. Purnick, R. Weiss, *Nature Reviews Molecular Cell Biology* pp. 410–422 (2009).

[89] P. S. Swain, M. B. Elowitz, E. D. Siggia, *Proceedings of the National Academy of Sciences of the United States of America* **99**, 12795 (2002).

[90] A. Hilfinger, J. Paulsson, *PNAS* **108**, 12167 (2011).

[91] D. Huh, J. Paulsson, *Nature Genetics* **43**, 95 (2011).

[92] M. A. Gibson, J. Bruck, *The Journal of Physical Chemistry A* **104**, 1876 (2000).

[93] M. Kærn, T. C. Elston, W. J. Blake, J. J. Collins, *Nature Rev. Genet.* **6**, 451 (2005).

[94] N. Friedman, L. Cai, X. Xie, *Physical Review Letters* **97**, 168302 (2006).

[95] M. He, T. Lu, *Metal-dielectric interfaces in gigascale electronics: thermal and electrical stability*, Springer Series in Materials Science (Springer, 2012).

[96] F. Crick, *Nature* **225**, 420 (1970).

[97] L. L. Latour, K. Svoboda, P. P. Mitra, C. H. Sotak, *Proceedings of the National Academy of Sciences of the United States of America* **91**, 1229 (1994).

[98] P. N. Sen, *Journal of Physics: Condensed Matter* **16**, S5213 (2004).

[99] M. Weiss, M. Elsner, F. Kartberg, T. Nilsson, *Biophysical Journal* **87**, 3518 (2004).

[100] A. Upadhyaya, J. P. Rieu, J. A. Glazier, Y. Sawada, *Physica A* **293**, 549 (2001).

[101] A. M. TURING, *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **237**, 37 (1952).

[102] J. Howard, S. W. Grill, J. S. Bois, *Nature Reviews Molecular Cell Biology* **12**, 400 (2011).

[103] M. L. Dequeant, *et al.*, *Science* **314**, 1595 (2006).

[104] S. Köster, *et al.*, *Lab on a Chip* **8**, 1110 (2008).

[105] J. F. Edd, *et al.*, *Lab on a Chip* **8**, 1262 (2008).

[106] J.-u. Shim, *et al.*, *Journal of the American Chemical Society* **131**, 15251 (2009).

[107] F. Courtois, *et al.*, *ChemBioChem* **9**, 439 (2008).

[108] M. Srisa-Art, A. J. deMello, J. B. Edel, *The Journal of Physical Chemistry B* **114**, 15766 (2010).

[109] J.-C. Baret, *et al.*, *Lab on a Chip* **9**, 1850 (2009).

[110] S. Cho, *et al.*, *Analytical chemistry* **85**, 8866 (2013).

[111] B. Rotman, *Proceedings of the National Academy of Sciences of the United States of America* **47**, 1981 (1961).

[112] M. Nakano, *et al.*, *Journal of Biotechnology* **102**, 117 (2003).

[113] O. J. Miller, *et al.*, *Nature Methods* **3**, 561 (2006).

[114] W.-C. Lu, A. D. Ellington, *Methods* **60**, 75 (2013).

[115] V. Vanag, I. Epstein, *Physical Review Letters* **87**, 228301 (2001).

[116] G. F. Christopher, S. L. Anna, *Journal of Physics D: Applied Physics* **40**, R319 (2007).

[117] M. J. Rosen, J. T. Kunjappu, *Surfactants and Interfacial Phenomena* (John Wiley & Sons, 2004).

[118] T. G. Mason, J. Bibette, *Langmuir* **13**, 4600 (1997).

[119] J.-C. Baret, *Lab on a Chip* **12**, 422 (2012).

[120] J. H. Hildebrand, D. Cochran, *Journal of the American Chemical Society* **71**, 22 (1949).

[121] J. N. Lee, C. Park, G. M. Whitesides, *Analytical chemistry* **75**, 6544 (2003).

[122] C. Holtze, *et al.*, *Lab on a Chip* **8**, 1632 (2008).

[123] G. I. Taylor, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **146**, 501 (1934).

[124] P. Garstecki, *et al.*, *Applied Physics Letters* **85**, 2649 (2004).

[125] T. Thorsen, R. W. Roberts, F. H. Arnold, S. R. Quake, *Physical Review Letters* **86**, 4163 (2001).

[126] P. Garstecki, M. J. Fuerstman, H. A. Stone, G. M. Whitesides, *Lab on a Chip* **6**, 437 (2006).

[127] S. L. Anna, N. Bontoux, H. A. Stone, *Applied Physics Letters* **82**, 364 (2003).

[128] P. Garstecki, H. Stone, G. Whitesides, *Physical Review Letters* **94**, 164501 (2005).

[129] S. L. Anna, H. C. Mayer, *Physics of Fluids* **18**, 121512 (2006).

[130] J.-u. Shim, *et al.*, *ACS nano* **7**, 5955 (2013).

[131] L. Mazutis, A. D. Griffiths, *Lab on a Chip* **12**, 1800 (2012).

[132] A. R. Abate, P. Mary, V. van Steijn, D. A. Weitz, *Lab on a Chip* **12**, 1516 (2012).

[133] P. A. Romero, A. R. Abate, *Lab on a Chip* **12**, 5130 (2012).

[134] M. De Menech, P. Garstecki, F. Jousse, H. A. Stone, *Journal of Fluid Mechanics* **141–161** (2008).

[135] A. S. Utada, *et al.*, *Science* **308**, 537 (2005).

[136] D. Link, S. Anna, D. Weitz, H. Stone, *Physical Review Letters* **92**, 054503 (2004).

[137] N. Bremond, A. Thiam, J. Bibette, *Physical Review Letters* **100**, 024501 (2008).

[138] H. Song, M. R. Bringer, J. D. Tice, C. J. Gerdts, R. F. Ismagilov, *Applied Physics Letters* **83**, 4664 (2003).

[139] K. Ahn, *et al.*, *Applied Physics Letters* **88**, 024104 (2006).

[140] T. Franke, A. R. Abate, D. A. Weitz, A. Wixforth, *Lab on a Chip* **9**, 2625 (2009).

[141] http://en.wikipedia.org/wiki/Quinine, Quinine (2013).

[142] J. R. Lakowicz, *Principles of Fluorescence Spectroscopy* (Springer, 2007).

[143] J. F. W. Herschel, *Philosophical Transactions of the Royal Society of London* pp. 143–145 (1845).

[144] F. G. Prendergast, K. G. Mann, *Biochemistry* **17**, 3448 (1978).

[145] T. Förster, *Annalen der Physik* **6**, 55 (1948).

[146] Olympus corporation, Fluorescence resonance energy transfer (fret) microscopy (2013).

[147] C. R. Cantor, I. Tinoco, Jr, *Journal of Molecular Biology* **13**, 65 (1965).

[148] M. Ronen, R. Rosenberg, B. I. Shraiman, U. Alon, *PNAS* **99**, 10555 (2002).

[149] G. Sezonov, D. Joseleau-Petit, R. D'Ari, *Journal of Bacteriology* **189**, 8746 (2007).

[150] J. Sambrook, D. Russell, *Molecular Cloning: A Laboratory Manual*, no. Bd. 1 (Cold Spring Harbor Laboratory Press, 2001).

[151] S. Kaneda, *et al.*, *Analytical Sciences* **28**, 39 (2012).

[152] J. Kim, J. J. Hopfield, E. Winfree, *Advances in neural information processing systems* **17**, 681 (2004).

[153] J. G. Wetmur, *Annual Review Of Biophysics And Bioengineering* **5**, 337 (1976).

[154] N. Srinivas, *et al.*, *Nucleic Acids Research* (2013).

[155] C. T. Martin, J. E. Coleman, *Biochemistry* **26**, 2690 (1987).

[156] I. Berkower, J. Leis, J. Hurwitz, *Journal of Biological Chemistry* **248**, 5914 (1973).

[157] M. Chamberlin, J. Ring, *Journal of Biological Chemistry* **248**, 2235 (1973).

[158] M. Jiang, M. Rong, C. Martin, W. T. Mcallister, *Journal of Molecular Biology* **310**, 509 (2001).

[159] F. Simmel, B. Yurke, *Physical Review E* **63**, 041913 (2001).

[160] B. Yurke, A. P. Mills, *Genetic Programming and Evolvable Machines* **4**, 111 (2003).

[161] B. Yurke, A. J. Turberfield, A. P. Mills, Jr, F. C. Simmel, J. L. Neumann, *Nature* **406**, 605 (2000).

[162] Y. Krishnan, F. C. Simmel, *Angewandte Chemie International Edition* **50**, 3124 (2011).

[163] K. E. McGinness, G. F. Joyce, *Journal of Biological Chemistry* **277**, 2987 (2002).

[164] N. E. Buchler, F. R. Cross, *Molecular Systems Biology* **5**, 1 (2009).

[165] W. F. Lima, S. T. Crooke, *Journal of Biological Chemistry* **272**, 27513 (1997).

[166] J. Rizzo, L. K. Gifford, X. Zhang, A. M. Gewirtz, P. Lu, *Molecular and Cellular Probes* **16**, 277 (2002).

[167] F. J. Triana-Alonso, M. Dabrowski, J. Wadzack, K. H. Nierhaus, *The Journal of Biological Chemistry* **270**, 6298 (1995).

[168] L. Qian, E. Winfree, *Science* **332**, 1196 (2011).

[169] T. Liedl, F. C. Simmel, *Nano Letters* **5**, 1894 (2005).

[170] S. R. Eddy, *Nature Reviews Genetics* **2**, 919 (2001).

[171] J. R. Babendure, S. R. Adams, R. Y. Tsien, *Journal of the American Chemical Society* **125**, 14716 (2003).

[172] J. S. Paige, K. Y. Wu, S. R. Jaffrey, *Science* **333**, 642 (2011).

[173] M. Schwarz-Schilling, Coordination of molecular processes with a synthetic biochemical oscillator, Master's thesis, Technische Universität München (2013).

[174] H.-X. Zhou, G. Rivas, A. P. Minton, *Annual Review of Biophysics* **37**, 375 (2008).

[175] T. K. De, A. Maitra, *Advances in colloid and interface science* **59**, 95 (1995).

[176] T. Matsuura, K. Hosoda, Y. Kazuta, N. Ichihashi, *ACS Synthetic Biology* (2012).

[177] A. B. Theberge, *et al.*, *Angewandte Chemie International Edition* pp. 5846–5868 (2010).

[178] I. R. Epstein, *et al.*, *Accounts of Chemical Research* **45**, 2160 (2012).

[179] I. Nagypal, I. R. Epstein, *The journal of physical chemistry* **90**, 6285 (1986).

[180] E. Fung, *et al.*, *Nature* **435**, 118 (2005).

[181] M. Tigges, T. T. Marquez-Lago, J. Stelling, M. Fussenegger, *Nature* **457**, 309 (2008).

[182] J. Ackermann, B. Wlotzka, J. S. McCaskill, *Bulletin of Mathematical Biology* **60**, 329 (1998).

[183] D. Gonze, J. Halloy, A. Goldbeter, *Proceedings of the National Academy of Sciences of the United States of America* **99**, 673 (2002).

[184] S. Kar, W. T. Baumann, M. R. Paul, J. J. Tyson, *Proceedings of the National Academy of Sciences of the United States of America* **106**, 6471 (2009).

[185] C. I. Hong, E. D. Conrad, J. J. Tyson, *Proceedings of the National Academy of Sciences of the United States of America* **104**, 1195 (2007).

[186] A. T. Winfree, *The geometry of biological time*, vol. 12 (Springer, 2001).

[187] D. Huh, J. Paulsson, *Proceedings of the National Academy of Sciences of the United States of America* **108**, 15004 (2011).

[188] Y. Liu, S.-Y. Jung, C. P. Collier, *Analytical chemistry* **81**, 4922 (2009).

[189] M. Maslak, C. T. Martin, *Biochemistry* **33**, 6918 (1994).

[190] M. Modesti, *Methods in molecular biology (Clifton, NJ)* **783**, 101 (2011).

[191] R. N. Gutenkunst, *et al.*, *PLoS Computational Biology* **3**, 1871 (2007).

[192] P. Subsoontorn, J. Kim, E. Winfree, *ACS Synthetic Biology* **1**, 299 (2012).

[193] M. Langecker, *et al.*, *Science* **338**, 932 (2012).

[194] B. Cayrol, *et al.*, *Journal of the American Chemical Society* **131**, 17270 (2009).

[195] F. J. Isaacs, D. J. Dwyer, J. J. Collins, *Nature Biotechnology* **24**, 545 (2006).

[196] M. B. Miller, B. L. Bassler, *Annual Reviews in Microbiology* **55**, 165 (2001).

[197] S. Milburn, M. Goldrick, M. Winkler (U. S. Patent 5256555, 1993).

[198] K. Rao, *Toxicology Letters* **81**, 107 (1995).

[199] S. Srivastava, R. Sinha, D. Roy, *Aquatic Toxicology* **66**, 319 (2004).

[200] C. Hansch, A. Leo, D. Hoekman, *Exploring QSAR.: Fundamentals and applications in chemistry and biology*, ACS professional reference book (American Chemical Society, 1995).

[201] F. Chizzolini, M. Forlin, D. Cecchi, S. S. Mansy, *ACS Synthetic Biology* (2013).

[202] W. Song, R. L. Strack, S. R. Jaffrey, *Nature Methods* **10**, 873 (2013).

[203] S. Youssef, S. Gude, J. O. Rädler, *Integrative Biology* **3**, 1095 (2011).

[204] C. Fuqua, E. P. Greenberg, *Nature Reviews Molecular Cell Biology* **3**, 685 (2002).

[205] B. A. Hense, *et al.*, *Nature Reviews Microbiology* **5**, 230 (2007).

[206] C. M. Waters, B. L. Bassler, *Annu. Rev. Cell Dev. Biol.* **21**, 319 (2005).

[207] H. Urbanczyk, J. C. Ast, M. J. Higgins, J. Carson, P. V. Dunlap, *International Journal of Systematic and Evolutionary Microbiology* **57**, 2823 (2007).

[208] T. Sohka, *et al.*, *Proceedings of the National Academy of Sciences of the United States of America* **106**, 10135 (2009).

[209] W. S. Choi, D. Ha, S. Park, T. Kim, *Biomaterials* **32**, 2500 (2011).

[210] R. D. Whitaker, *et al.*, *Journal of Biological Chemistry* **286**, 21623 (2011).

[211] J. Q. Boedicker, M. E. Vincent, R. F. Ismagilov, *Angewandte Chemie International Edition* **48**, 5908 (2009).

[212] J.-u. Shim, *et al.*, *Lab on a Chip* **11**, 1132 (2011).

[213] Y. Bai, *et al.*, *International Journal of Molecular Sciences* **14**, 10570 (2013).

[214] J. J. Tabor, *et al.*, *Cell* **137**, 1272 (2009).

[215] G. E. Dilanji, *et al.*, *Journal of the American Chemical Society* **134**, 5618 (2012).

[216] M. L. Urbanowski, C. P. Lostroh, E. P. Greenberg, *Journal of Bacteriology* **186**, 631 (2004).

[217] B. Canton, A. Labno, D. Endy, *Nature Biotechnology* **26**, 787 (2008).

[218] X. Li, *et al.*, *Journal of Chromatography A* **1134**, 186 (2006).

[219] B. M. Davis, R. Jensen, P. Williams, P. O'Shea, *PloS one* **5**, e13522 (2010).

[220] R. Iizuka, M. Yamagishi-Shirasaki, T. Funatsu, *Analytical biochemistry* **414**, 173 (2011).

[221] B. A. Hense, J. Müller, C. Kuttler, A. Hartmann, *Sensors* **12**, 4156 (2012).

[222] M.-V. Vazquez, A. M. Berezhkovskii, L. Dagdug, *The Journal of chemical physics* **129**, 046101 (2008).

[223] S. Regot, *et al.*, *Nature* **469**, 207 (2011).

[224] R. Lutz, H. Bujard, *Nucleic Acids Research* **25**, 1203 (1997).

[225] J. Quan, J. Tian, *PloS one* **4** (2009).

[226] R. Berg, Characterization of gene regulation functions in *Escherichia coli* (Bachelor's thesis, Technische Universität München, 2013).

Acknowledgement - Danksagung

Vielen Dank...

- allen Mitarbeitern des Lehrstuhl Simmel für die schönen Jahre und die gute Atmosphäre.
- Fritz Simmel, für die Aufnahme am Lehrstuhl und die hervorragende Betreuung.
- Erika, für dein Engagement weit über Deine Pflichten hinaus.
- Helene, für Deine Unterstützung in allen Belangen.
- Jongmin Kim and Elisa Franco, for a successful collaboration and a great time in Pasadena and Riverside.
- Andrea Mückl und Korbinian Kapsner, für die lustige und erfolgreiche Zusammenarbeit.
- dem Elitenetzwerk Bayern und dem IDK-CompInt, NIM und BaCaTec für die finanzielle Unterstützung.
- meiner ganzen Familie.
- meiner Freundin Birgit, für alles was Du für mich getan hast. Es war schön die Promotionszeit gemeinsam mit Dir zu erleben.