Single cell measurements in microfluidic chips to determine the dynamics of transcription during induction

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Einzelzellmessungen in mikrofluidischen Chips zur Bestimmung der Transkriptionsdynamiken während Stimulation

Single cell measurements in microfluidic chips to determine the dynamics of transcription during induction

vom Fachbereich Biologie
der Technischen Universität Darmstadt

zur Erlangung des Grades
Doktor rerum naturalium
(Dr. rer. nat.)

Dissertation
von Jascha Carl Roman Diemer

Erstgutachter: Prof. Dr. Heinz Koeppl
Zweitgutachter: Prof. Dr. Beatrix Süß

Ehrenwörtliche Erklärung

Ich erkläre hiermit ehrenwörtlich, dass ich die vorliegende Arbeit entsprechend den Regeln guter wissenschaftlicher Praxis selbstständig und ohne unzulässige Hilfe Dritter angefertigt habe.

Sämtliche aus fremden Quellen direkt oder indirekt übernommenen Gedanken sowie sämtliche von Anderen direkt oder indirekt übernommenen Daten, Techniken und Materialien sind als solche kenntlich gemacht. Die Arbeit wurde bisher bei keiner anderen Hochschule zu Prüfungszwecken eingereicht.

Ort, Datum

Jascha Carl Roman Diemer

Acknowledgements

I would like to thank my supervisor and mentor Prof. Dr. Heinz Koeppl for the possibility of graduating and his support during the last five years.

Further thanks to Prof. Dr. Beatrix Suess, who helped me with biological questions and problems.

Chris, Leo, Christian, Sebastian, Tim and François receive my gratitude for the time working together.

I like to thank Sarah, Megan, Kathrin and Markus for their helping hand in the laboratory.

To all the members of BCS, thank you all for the time spent together and the interesting discussions. I hope I could expand your knowledge about biology the same way you expanded mine about modeling and math.

To all my friends, especially Philipp, Henrik, Renè and Malte: Thank you so much for your input and the time spent together.

I thank my family for all the love and support they have given me.

Last but not least, Patricia. You are my lobster. Thank you for everything.

Zusammenfassung

Mikrofluidische Chips eignen sich hervorragend für Einzelzellmessungen. Spezielle Designs ermöglichen das Separieren und Einfangen von einzelnen *Saccharomyces cerevisiae*. Das Mikroskop als leistungsstarkes Instrument zur Bestimmung von temporalen und spartialen Abläufen gewährleistet die Aufnahme von relevanten Informationen.

Meine Arbeit wird durch zwei Themen motiviert. Zum einen wird ein mikrofluidischer Chip weiterentwickelt, welcher einzelne Zellen einfängt um langandauernde Experimente unter gleichbleibenden Bedingungen zu erlauben. Zum anderen wird untersucht, wie sich die Transkriptionsdynamik unter Induktion verändert.

Zunächst zum Chip: Die erste Verbesserung betrifft eine wesentliche Verkleinerung des Layouts. Nun finden bis zu fünf Zellobservationskammern auf einem Chip Platz. Zudem wird die Herstellung erleichtert indem zwei von vier Masken zu einer zusammengelegt werden. Drei verschiedene Layouts werden entworfen, um entsprechende Experimente durchführen zu können. Das erste Design bietet Platz für vier unterschiedliche Hefestämme, die in einem Experiment verglichen werden sollen. Das zweite hat vier eng aneinanderliegende Zellkammern in denen unabhängige Messungen in einem Lauf aufgenommen werden können. Das letzte Design erhöht vornehmlich die Wiederverwertbarkeit eines Chips mit fünf Kammern.

Die zweite Optimierung steigert nicht nur die Produktionsrate, sondern testet auch andere Formen der Zellfallen. Die Platznutung wird weiter verbessert und die Chips bieten nun zwischen sieben und zwöf Kammern. Da die Herstellung mit den bisherigen Fallen an ihre Grenze stößt, werden neben bereits publizierten Fallen auch ovale Formen mit erhöhter Fläche entworfen und getestet. Hierbei zeigt sich, dass nur die größten Fallen mit gleichbleibender Qualität hergestellt werden können. Erst mit der Verwendung eines Laser-Direkt-Writers kann die Fabrikation mit allen Formen vollendet werden. Die aus der Literatur bekannten L-förmigen Fallen weisen die beste Fangrate (90%) auf. Kleinere Fallen haben geringere Abstände und kommen trotz 70%iger Fangrate auf die gleiche Anzahl an aufgenommenen Zellen.

Die Kooperation mit Dr. Christopher Schneider (AG Süß) kann auf den Chip mit vier unabhängigen Kammern zurückgreifen. Als genetischer Schalter wird ein Aptamer konstruiert, welcher sowohl auf Neomycin als auch auf Tetrazyklin die Translation inhibiert (yCS). In den vier Kammern können daher zeitgleich die Positiv-Kontrolle, Neomycin, Tetrazyklin und die Kombination beider aufgenommen und ausgewertet werden. Dies gewährt eine Zeitersparnis von rund drei Tagen gegenüber der Verwendung des ursprünglichen Chips. Der genetische Schalter zeigt Wirkung: Neo und Tc reprimieren beide die Produktion von GFP 70 Minuten nach Zugabe.

Der Chip mit vier parallelen Zellkammern kommt in der Kooperation mit Sebstian Höler (AG Thiel) zum Einsatz. Hierbei wird die Genexpression eines Licht-abhängigen Kaliumkanals untersucht (ySH). Mehrere verschiedene Sequenzoptimierungen werden vergleichend aufgenommen. Dabei stellt sich heraus, dass die einfache Sequenz das

beste Signal liefert. Der Kaliumkanal lagert sich in allen Varianten im endoplasmatischen Retikulum ein.

Der zweite Forschungsschwerpunkt meiner Dissertation dreht sich um die Transkriptionsdynamik. Auch hier werden einzelne Zellen unter dem Mikroskop mit dem mikrofluidischen Chip gemessen. Die Visualisierung der Transkription in Echtzeit wird mittels dem Coat-Proteins des Phagen PP7 realisiert. Eine Stammschleife in der mRNA-Sequenz wird erkannt und durch Repititionen kann eine Akkumulation von PP7-GFP herbeigeführt werden, welche als Punkt gemessen wird. 14 solcher Stammschleifen vor dem Gal10-Gen werden mit dem Transkriptionsfaktor GEV unter Zugabe von β-Estradiol synthetisiert (yJD). Da Transkription als stochastischer Prozess großen Schwankungen unterworfen ist (sog. Bursts), ist es von Interesse, wie sich Burstintensität und -frequenz durch unterschiedlich starke Stimulation beeinflussen lassen. Der erste Fund betrifft die Anzahl an *S. cerevisiae* Zellen, welche auf die Stimulation reagieren. Mit der höchsten verwendeten Konzentration von 500 nM β-Estradiol hat man die höchste Fraktion an Respondern mit ca. 45% gefunden. Zudem steigt die Anzahl über die Dauer des Experimentes stetig an, was mit einer Steigerung der Re-Initiationsrate erklärt werden kann. Die hohe Konzentration an vorliegendem GEV im Nukleus führt zu einer höheren Wahrscheinlichkeit aktive Zellen zu beobachten. Entgegen der Erwartung und Literaturerkenntnisse ist dies die einzige Variable, welche mit der Dosis korreliert. Die Transkriptionsdynamik als solche bleibt in diesem Assay unberührt. Exprimiert eine Zelle das Konstrukt, so folgt die produzierte mRNA Menge über alle Konzentrationen der gleichen Verteilung.

In einer weiterführenden Studie soll der Einfluss des GC-Gehalts des DNA-Templates auf die Elongationsgeschwindigkeit ermittelt werden. Die Klonierung eines Plasmides mit 10 Stammschleifen für PP7, 1000 Basenpaaren einer 66% GC-haltigen Sequenz und anschließend 24 Stammschleifen des MS2 Coat-Proteins ermöglichen eine genauere Analyse der Elongationsrate durch die Aufnahme in zwei Farbkanälen. Dabei sollte zuerst PP7-GFP an die mRNA binden und den Begin des GC-haltigen Gen markieren, während MS2-RFP gegen Ende der Transkription sich an die mRNA lagert. Die Integration dieses Konstruktes ist nicht rechtzeitig vollendet worden und keine Aufnahme mit zwei Farben kann ausgewertet werden.

Ich fasse die Kernpunkte meiner Thesis wie folgt zusammen: Die Optimierung des Chiplayouts hat den Herstellungsprozess vereinfacht und zu einer Steigerung der produzierten Chips geführt. Die getesteten Fallenstukturen sind alle funktionsfähig und können Hefen über große Zeiträume hinweg kultivieren. Der genetische Schalter yCS lässt sich mit beiden Liganden reprimieren und der Licht-induzierbare Kaliumkanal ySH zeigt die gleiche Expressionsdynamik für unterschiedliche DNA-Sequenzen. Und die induzierte Transkription am Gal10 Lokus steigert die Rate mit der Zellen aktiv werden.

Résumé

Single cell experiments require a system that is capable of collecting signals on the scale of a cell. Merging a microfluidic system with a microscopic setup opens up the possibility to collect data on the single cell level in a controlled manner. A microfluidic chip can both separate and trap cells with the right design. The so called “cell traps” are perfectly suited to observe the organism *Saccharomyces cerevisiae* over long time periods.

This work contributes to the design of traps and their fabrication process. In order to reduce the complexity, two layers of the layout are combined into one. The overall orientation of the chip is reduced in size to fit multiple copies onto one device. In a first drawing, the total number of cell chambers could be increased from one to five, in a second iteration even further to twelve. Furthermore, different designs are used for different purposes. One design enables the simultaneous measurement of up to four yeast strains with the same environment and conditions. Another design allows to image again four chambers, but with different media compositions. Both designs were planned to reveal variations between the strains and conditions used. In addition, the structure of the cell trap was optimized for the fabrication devices, as those could not be generated with the same quality. A big oval shape seemed to solve the problems as it has no sharp angles and a big surface to be washed out. All other shapes of traps could only be produced with the use of a laser direct writer. Thereby, differences in the catch rate could be observed. The L-formed traps had the highest catch rate with 90%. The drawback is the limited number of traps that can be placed in close proximity. Smaller trap designs could catch higher numbers of cells as they were more densely packed.

In two collaborations both chip designs could be used successfully. The characterization of a genetic logic gate based on two aptamers was performed with the same yeast strain under four different conditions: (I) positive control with inducer of gene expression, (II) exposure of cells with neomycin after 2 h of induction, (III) cells were exposed to tetracycline and (IV) cells had both ligands in the media. The increase in GFP signal generated from cells with at least one repressor was stopped after 70 min.

The second collaboration was about a light-sensitive potassium channel. Here, different strains with alternative DNA sequences for the GFP-tagged channel have been compared. The chip with four chambers for parallel experiment was best suited to keep the media as well as the light conditions identical. Interestingly, the non-codon optimized strain showed the best signal. The temporal development of the relative signal was in all strains the same. In the microscope images the localization of the channel in the ER membrane was visible.

Transcription of mRNA is a stochastic process. So called bursts give rise to the mRNA distribution in the cell. In the case of stimulus dependent transcription, the parameters of the bursting are subject of change. This work investigates the question, which parameters change based on the intensity of activation.

The genetic background of the yeast cell line to perform measurements on the transcription dynamics consists of two parts: a PP7-based system to tag mRNA directly during synthesis and the GEV transcription factor to respond to β-estradiol. The coat protein of PP7 recognizes a stem loop in the mRNA and strongly binds this motif. Multiple stem loops in repetition lead to an accumulation of PP7-GFP on the mRNA, which is visible as a spot in the nucleus. The duration and intensity are directly related to the transcription dynamics. The GEV transcription factor consists of a *Gal4* DNA binding domain, an estrogen receptor and a viral activation domain. GEV relocates into the nucleus during induction and activates gene expression of Gal1 and Gal10, when β-estradiol is present. As expected, the number of cells that respond to the input rose with higher concentrations of the inducer. Interestingly, the number of responders did not reach a plateau, when induced with 500 nM over 4 h, but continuously increased over time. This finding might be explained by an increased re-initiation rate of polymerases. As GEV gets accumulated in the nucleus, it is more likely to observe an actively transcribing cell. The number of mRNA synthesized is not correlated with the induction dose. Once transcription is started, the cell produces a number of mRNA molecules following the same distribution for any β-estradiol concentration.

In addition, a GC-rich sequence should give knowledge about the influence of DNA template properties on the elongation rate. As the GC pair forms a stronger bond, it can slow down polymerases as indicated by literature. During the design of a suitable fragment, the melting temperature was chosen to be constant for a window of 14 base pairs. Although the cloning of such a construct was successful, the integration into the yeast genome was not. Furthermore, the plasmid had repeats of another coat protein, MS2, in the 3' UTR. Such a dual tagging strategy would have simplified the determination of the elongation speed.

In conclusion, the optimizations of the chip design reduced the waste and increased the number of usable chips after one fabrication run. All tested traps were able to catch and keep cells for longer time periods. Different designs had different advantages and disadvantages. The chips could be utilized in three projects. The main results were: the neomycin-tetracycline gate responds to both repressors; the light-inducible potassium channel relocates into the ER membrane and the number of responding cells increases with higher concentration of β-estradiol.

Contents

1 Theoretical basis of transcription **1**

1.1 Transcription: the dogma in biology 1
1.2 RNA polymerases 1
1.3 RNA polymerase II 2
1.4 Control of transcription of pol II 3
1.5 Pre-initiation complex 4
1.5.1 General transcription factors 5
1.5.2 Mediator 5
1.6 Initiation 6
1.7 Elongation 6
1.8 Termination 7
1.9 Bursting 8

2 Metrology **10**

2.1 Microscope 10
2.2 Methods to measure transcription elongation 10
2.3 RNA-binding proteins 11
2.3.1 *Pseudomonas aeruginosa* phage PP7 coat protein 11
2.4 Microfluidic chips 14

3 Aim of this Work **17**

4 Results **18**

4.1 The microfluidic chip as center for projects with *S. cerevisiae* 18
4.1.1 Biological factors on loading and clogging 23
4.1.2 Physical parameters for clogging 24
4.2 Dynamics of transcription in yeast 25
4.3 Dose dependency of transcription dynamics 31
4.4 DNA properties and their influence on transcription elongation 36
4.5 Inference on single cell data 37
4.6 Real time imaging and evaluation of the neo-tc gate 40
4.7 Expression dynamics of light inducible potassium channels 42

5 Bringing the results into context **47**

5.1 Design and fabrication challenges of the chip . . . 47
5.2 Transcription dynamics in response to an external stimulus . . . 49
5.3 Unfinished project about the influence of GC content on transcription dynamics . . . 51
5.4 Conclusion of ROC'n'Ribo . . . 52
5.5 Increasing image quality for correct membrane localization . . . 53
5.6 What should be investigated next . . . 54
5.6.1 Future of the traps . . . 55
5.6.2 Elongation project . . . 55
5.6.3 Aptamer based NOR gate . . . 56

6 Conclusion **57**

7 Material **58**

7.1 Devices . . . 59
7.2 Plasmids . . . 59

8 Methods **60**

8.1 Competent yeast cells . . . 60
8.2 Microscope setup . . . 61
8.3 Chemical competent *E.coli* . . . 62
8.4 DNA Manipulation . . . 62
8.5 Microfluidic chips . . . 64
8.6 Fabrication of master wafer . . . 64
8.7 Casting of microfluidic chips . . . 64
8.8 Experiments for transcription dynamics . . . 64
8.9 Image procession . . . 66

9 Appendix **68**

9.1 Cirriculum Vitae . . . 68
9.2 Supporting Material . . . 68

Abbreviations

Abbreviation	Full meaning
A	adenosine
aa	amino acid
bp	base pair
C	cytosine
CC	closed complex
CDK8	cyclin-dependent kinase 8
CLSM	confocal laser scanning microscopy
CTD	C-Terminal domain
DNA	deoxyribonuclein acid
ER	endoplasmatic reticulum
FRAP	fluorescence recovery after photobleaching
G	guanine
Gal	galactose related gene
GAL	galactose related protein
gal	galactose
GEV	GAL4 DNA binding domain- estrogen receptor - VP16
GFP	green fluorescent protein
GTF	general transcription factors
HYAA	high-throughput yeast aging analysis chip
kb	kilo base pairs
mRNA	messenger RNA
MCP	bacteriophage MS2 coat protein
MSL	MCP stem loop
NLS	nuclear localization sequence
nt	nucleotide
NTP	nucleotide triphosphate
OC	open complex
ORI	origin of replication
P-TEF	positive transcription elongation factor
PCR	polymerase chain reaction
PDMS	polydimethylsiloxane
PIC	pre-initiation complex
pol	RNA polymerase
PCP	Phage Protein 7 coat protein
PPi	pyrophosphate
PSL	PCP stem loop
PWM	pulse width modulation
RNA	ribonucleic acid

Abbreviation	Full meaning
RNAP	RNA polymerase II
RPB	RNA polymerase II subunit B
rRNA	ribosomal RNA
SEC	super elongation complex
SL1	selectivity factor 1
snRNA	small nuclear RNA
STED	stimulated emission depletion
T	tyrosine
TBP	TATA-box binding protein
TF	transcription factor
TFII	transcription factor associated with RNA polymerase II
TFIII	transcription factor associated with RNA polymerase II
TFIIA	transcription factor A for pol II
TFIIIA	transcription factor A for pol III
TL	translation
tRNA	transfer RNA
TX	transcription
UBF	upstream binding factor
yCS	yeast strain C. Schneider
yJD	yeast strain J. Diemer
ySH	yeast strain S. Höler

1 Theoretical basis of transcription

1.1 Transcription: the dogma in biology

Life on earth runs two mechanisms to access stored information. The first synthesizes ribonucleic acids (RNA) from deoxyribonucleic acids (DNA) and is called transcription. Those products are classified by their function: protein relevant information are carried by messenger RNA (mRNA) and are translated by ribosomes into amino acid sequences. This translation of mRNA to amino acids is the second mechanism. The following dissertation will focus on transcription, its biomolecular mechanism and dynamics.

1.2 RNA polymerases

The core enzyme of transcription is the RNA polymerase. DNA in the transcription bubble of the template strand is "read" and an appropriate nucleotide is added to the synthesized RNA. "Reading" in this case, is the addition of the energetically most suitable nucleotide following Watson-Crick base pairing. This pairing, discovered by Watson and Crick in 1953, describes the configuration of the DNA and its complementary strand [100]. Adenine (A) is paired with thymine (T) for DNA or uracil (U) for RNA, cytosine (C) forms three hydrogen bonds to guanine (G). The sugar back bone determines if the information carrier is DNA or RNA. Each RNA polymerase catalyzes the nucleophilic attack of the 3'OH of the chain on the α-phosphate group of the new nucleotide. The phospho ester bond is formed under the release of PPi and the polymerase moves to the next position. In eukaryotes there are 3 different classes of RNA polymerases: I, II and III. The first class is responsible for up to 50% of all RNAs inside the cell. Its transcription machinery consists of the Pol I enzyme, the TBP (TATA-binding protein) -TAF (TBP-associated factor) complex SL1 (selectivity factor 1)/TIF-IB (transcription initiation factor-IB) and the transactivator protein UBF (upstream binding factor). The only product of Pol I is preribosomal RNA (47 S precursor rRNA), which is processed into three mature rRNAs (5.8 S, 18 S and 28 S) and these, together with the 5 S rRNA synthesized by Pol III, comprise the enzymatic and structural scaffold of the ribosome [75]. Pol I shares in its composition of the core protein homologies with the other two classes. Subunit composition is, as expected by the same function, conserved: the two biggest subunits AC40 and AC19 are directly shared with Pol III and find their analogs in the RPB3-RPB11 hetero-dimer from Pol II. The RPB6 subunit of Pol I, shared with Pols II and III, is a structural and functional homologue of the bacterial ω subunit. Of the 14 subunits Pol I consists of, 7 shared directly or with structural consensuses with Pols II and III in most eukaryotes. The activity of Pol I can be directly linked to the growth rate of the organism, as its products

result in the assembly of ribosomes [99]. The regulation of the rDNA promoter is rather simple and synced with the cell cycle as there are approximately 400 copies of the rDNA genes, which share the same promoter. A transcription cycle of Pol I includes the binding of upstream binding factor (UBF) ~100 - 200 bases upstream of the promoter and SL1 at the promoter. Pol Iβ gets recruited and clears the initiation complex after activation through hRRN3. SL1 and UBF stay bond to the promoter to facilitate re-initiation of other polymerases [75].

Pol III is the most complex of the nuclear RNA polymerases. It has a total molecular weight of around 700 kDa and comprises of 17 subunits. 5 of those are unique for Pol III, the others share sufficient homologies to different subunits of Pols I or II. Pol III has similar to Pol I a small set of genes to transcribe [97]. The key role of Pol III is supplying 5S rRNA and tRNAs for protein synthesis [22]. With advancements in sequencing, more transcripts have been found that are linked to Pol III. Regulation of Pol III relies on the initiation on three different promoters [79]. The class I promoter is exclusively for the 5S RNA gene. An intragenic regulation site is recognized by TFIIIA, which in turn recruits TFIIIC followed by TFIIIB, before Pol III enters the stage. Class II promoter occurs mostly at tRNA genes and has two motifs (A and B box) for binding TFIIIC inside the gene. TFIIIB will find TFIIIC and initiation ensues. The class III core promoters were identified originally in mammalian U6 snRNA genes, which encode the U6 snRNA component of the spliceosome. Their discovery came as a surprise because, unlike the then-characterized class I and II promoters, the class III core promoters turned out to be gene-external. They are located in the 5'-flanking region of the gene and consist of a proximal sequence element, which defines the core of RNA polymerase III snRNA promoters, and a TATA box located at a fixed distance downstream [79]. Termination of Pol III is implemented by a short poly(T) site.

1.3 RNA polymerase II

The RNA polymerase II (RNAP) transcribes DNA to RNA of most protein coding sequences. It is a 550 kDa protein complex, consisting of 11 subunits [107]. Among the subunits, one particular is extraordinary unique: its C-terminal domain (CTD). The two other RNA polymerases lack this functional domain, which consists of multiple repeats of Tyr-Ser-Pro-Thr-Ser-Pro-Ser and is subject to protein-protein interaction, phosphorylation (for capping factors of the 5' RNA), proline-isomerization, poly(A)-lynation and glycosylation [25]. The pattern of modification changes during transcription and leads to the recruitment of RNA processing proteins in a specific and time dependent order [42]. The patterns are thought to be a molecular signal code for proteins. The core of the polymerase, the reactive center, has two Mg^{2+} ions complexed for NTP condensation and PPi release [1]. Newly synthesized mRNA leaves the active center through a tunnel on the side, which can be occupied by TFIIB at the start of transcription. DNA-RNA hybrid base pairing offers the stability for RNAP during polymerization and the chance of

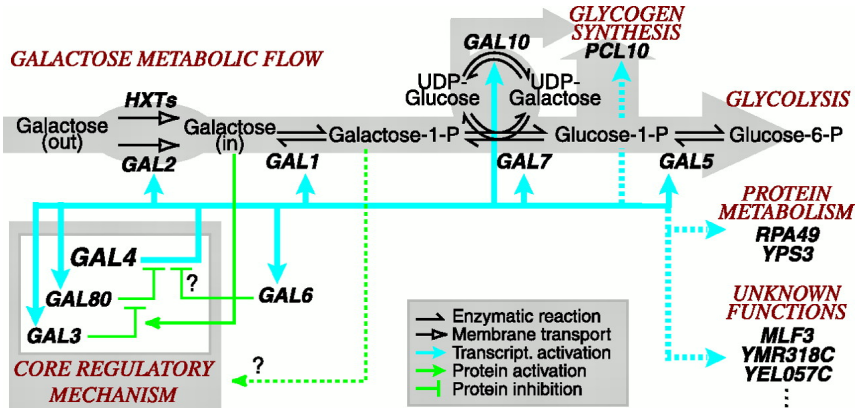

Figure 1.1: Galactose gene network, adapted from [40]. The GAL4 protein activates transcription of *Gal1*, *Gal2*, *Gal3*, *Gal5*, *Gal6*, *Gal7*, *Gal10* and *Gal80* (cyan arrow). The three core regulatory proteins are GAL3, GAL4 and GAL80. They activate and inactivate each other in turn and detect the presence of galactose (green lines). Dotted lines may indicate correlation of interactions. GAL1, 7 and 10 show the highest response on galactose induction.

proofreading. The transcription bubble describes the opening of the DNA helix for ~14 nt [51].

1.4 Control of transcription of pol II

Organisms, from single bacteria to highly complex mammals, are equipped with a repertoire to adapt to the ever changing environment. A classical example is the metabolism of the carbon source. Sugar derivatives are energy providing carbon sources for microorganisms and animals. For their best utilization, different proteins and enzymes are required, which are specialized for a certain sugar. If one available sugar becomes depleted, cells have to adapt to the environment and change their metabolism to use alternative energy sources. As an example, induced transcription ensures the switch from glucose to galactose based metabolism. The genes expressing galactose processing proteins are inhibited by glucose and activated by galactose itself. A combination of repressors and activators ensures the most efficient way to switch the metabolism (see Fig. 1.1) and generate a metabolic memory. Besides sugars, there are a lot of external triggers to stimulate transcription, like osmotic stress, heat, amino acids and many more. A functionality that the induction of all triggers share is a mechanism of binding a specific DNA domain and recruiting the general transcription factors. In the example of the galactose metabolic pathway, GAL4 is binding several promoters of galactose associated genes, like *Gal1* and *Gal10*.

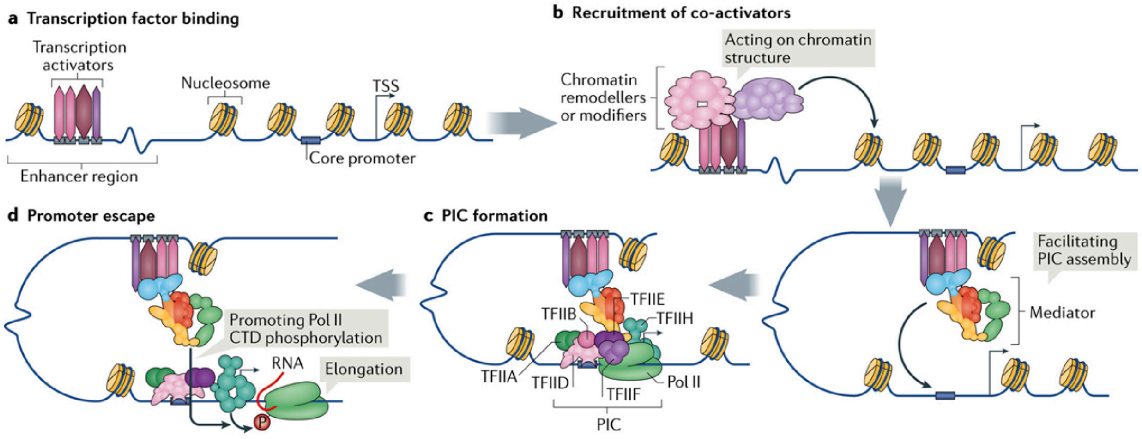

Figure 1.2: Pol II assembly after Soutourina *et al.*, 2018 [87]. **A** Transcriptional activators bind on enhancer regions. **B** Chromatin remodelers take action and make the core promoter and transcription start site (TSS) accessible. **C** Mediator bridges the enhancer to the core promoter and recruits the GTFs. **D** RNAPII leaves the promoter into elongation.

1.5 Pre-initiation complex

For a single transcription event, multiple molecules have to assemble on a genes operon. This assembly is highly regulated and is a result of either stimuli or exists constantly. In cells, we find constitutively expressed house keeping genes, cell cycle dependent, metabolite, stress or heat induced genes. A key step in initiation is the formation of the pre-initiation complex (PIC). It consists of TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH, Pol II and Mediator in both the open and closed form of the PIC (Fig. 1.2). The term opening refers to the transcription bubble, which is the partial melting of the double stranded DNA template. Before the melting of the DNA can transit into elongation, the packaging units of DNA, namely nucleosomes, have to be rearranged. A nucleosome consists of 2x4 subunits (histones) and are encircled by 147 base pairs. Depending on the location of the promoter, nucleosomes have to be remodeled to gain access for the PIC (Fig. 1.2, b) [46]. The packaging of DNA is a physical repression of transcription [59]. In human cells, one of the two chromosome copies is highly packed into so called heterochromatin, which shows no transcription at all. In yeast, most chromosomes are accessible and regulation is shifted towards other protein-DNA interactions [6]. Another topic which is highly related to transcription control in higher developed eukaryotes is the methylation of CpG sites and histones. Those two mechanisms are linked to silencing genes. The modification of bases or histones and its influence on gene expression and phenotypes is called epigenetics [81]. Additional literature can be found in [92] and the references therein.

1.5.1 General transcription factors

The term general transcription factor (GTF) includes all necessary proteins that have to bind on the promoter region to recruit Pol II: TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH [69]. In all eukaryotes, these proteins play a pivotal role in transcription and share great homology between species. GTFs are involved in transcription, but gene specific TFs can bind and recruit GTFs as well (in the galactose network GAL4, Fig. 1.1). A general promoter motif, like the TATA-box is independent of specific TFs. Some highlights of each TF are summed up in the following:

TFIIA can be cleaved into two subunits alpha and beta. It is not essential for transcription *in vitro*, but enhances it. TFIIA has regulatory function in development [38].

TFIIB aids the DNA opening with the B linker; the template strand is then placed into the RNA polymerase II cleft and the bubble is stabilized by the B reader (open complex formation). RNA polymerase II and B reader scan the DNA in order to position the transcription start site. Then, the first phosphodiester bond is formed. Production of short abortive transcripts happens due to clashes with the B reader loop. The growth of nascent RNA chain to 12-13 bases leads to ejection of TFIIB to prevent further clashes [52, 71].

TFIID is first to bind the promoter region with its subunit TATA box binding protein (TBP) [36]. Binding of TBP introduces a 90° curve in the DNA [89].

TFIIE recruits TFIIH to the initiation complex and stimulates the RNA polymerase II C-terminal domain kinase and DNA-dependent ATPase activities of TFIIH. Both TFIIH and TFIIE are required for promoter clearance by RNA polymerase [67].

TFIIF strongly stabilizes the binding of polymerase to TFIID and TFIIB at the promoter. It is required for entry of TFIIE and TFIIH into the preinitiation complex, for subsequent open complex formation catalyzed by the TFIIH DNA helicase, and for synthesis of the first phosphodiester bond of nascent transcripts [105].

TFIIH is a multifunctional RNA polymerase II transcription factor that possesses DNA-dependent ATPase, DNA helicase, and protein kinase activities. TFIIH enters the preinitiation complex and fulfills a critical role in initiation by catalyzing ATP-dependent formation of the open complex prior to synthesis of the first phosphodiester bond of nascent transcripts [24].

Mediator is not essential for transcription *in vitro* per se, but is a lethal knock-out *in vivo* [73]. It binds to/is a part of the pre-initiation complex (PIC) [33] and interacts with pol II and multiple TFs.

The temporal order of action and binding to the DNA/PIC is as follows: TFIID/TFIIA, TFIIB, RNA polymerase II/TFIIF, TFIIE, TFIIH [86].

1.5.2 Mediator

Mediator is a central protein complex in the initiation of RNAP and is conserved over mammalian species. The term Mediator refers in yeast to a 21 subunit complex and

bridges signals between TFs and RNAP in a central role. Its 3D structure is segmented into 4 domains: head, middle, tail and CDK8. During assembly, subunits can be removed or added to proceed from the PIC to elongation. This finding suggests an adaptation of Mediator to different necessary gene expression patterns for the development of cells. The list of interactions between TFs and Mediator is long and not the topic of this thesis. To give a brief introduction to a pivotal enzyme in transcription regulation, some highlights are presented:

1. In the case of a TATA-box containing promoter, TFIID binds with its functional group TBP and is recognized by Mediator [92].
2. In the case of gene specific transcription factors, Mediator is also the second enzyme which binds.
3. The tail domain of Mediator is capable of binding enhancer regions, several thousand nucleotides upstream or downstream of the promoter core region [87].
4. The CTD of RNAP is modified by Mediator during PIC assembly [74].

1.6 Initiation

To start transcription, the RNA polymerase II has to get access to its template, the DNA. The helicase activity of TFIIH opens the double helix and RNAP can start with transcription. During the synthesis of the first 8 nt, RNAP is stabilized by the PIC. Once the interaction of the RNA-DNA hybrid is robust enough, the bubble collapses at the back of RNAP again into the normal DNA helix. Binding of TFIIB is lost and RNAP leaves the transcripiton start site. A phenomenon observed in human cells is promoter proximal pausing [45]. After 30-60 nt transcribed, RNAP comes to a stop as an additional regulatory step and has to be released from its paused state.

1.7 Elongation

Once the polymerase escapes the promoter it loses contact to the PIC [3]. If RNAP is in a transcribing state together with the proteins P-TEFb, AFF4, and ELL, or their species specific analogons, this assembly is called super elongation complex (SEC) [61]. P-TEFb (positive transcription elongation factor b) is a cyclin dependent kinase, which is involved in the release of RNAP from the promoter proximal pause site. It phosphorylates its counterplayer NELF (negative elongation factor) and the CTD of RNAP to keep the polymerase in an transcribing state. As the polymerase moves through the gene body, nucleosomes have to be bypassed. Data available on the dynamic histone compilation of the nucleosome supports multiple mechanisms on different genes [104]. Four modes of operation have been found: (I) Sliding the nucleosome along the DNA to make the gene accessible, (II) partially unwrap the DNA from the nucleosome, (III) complete eviction of all histones and (IV) partially removing or exchange of histones [104]. During elongation NTPs are consumed and incorporated under the release of PPi. The reaction is a

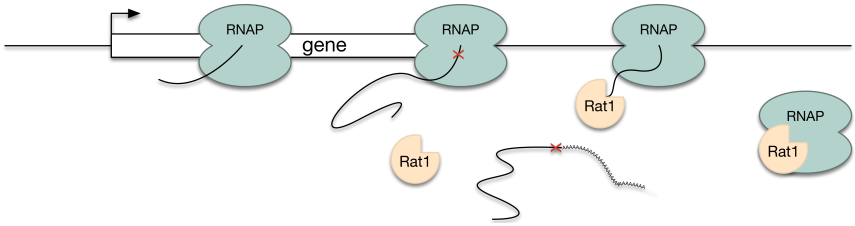

Figure 1.3: Scheme of the torpedo termination mechanism. Upon reaching the polyA site (red cross), Rat1 cuts the RNA and starts to digest the newly synthesized RNA. The cleaved RNA gets its polyA tail and moves on to translation, while Rat1 collides with RNAP and stops its transcriptional activity.

nucleophilic attack of the 3' hydroxyl group at the α-phosphate of the incoming ribonucleoside triphosphate with Mg^{2+} as a catalyst and stabilizer. The different nucleotides have to enter and leave RNAP in order to find the correct base for the DNA template. The elongation rate is found in a range of 15 to 60 nt/s [98, 14]. Veloso *et al.* have found in 2014 several features of the DNA itself or genomic context that influence the elongation rate [98]. In their sequencing data of human cell lines negative correlation is found for exon density, GC content, long terminal repeats and CpG methylations. Higher elongation rates are more frequent for genes with low complexity, long genes and distance to other genes. All findings are not yet explained in detail. It is unclear how a long gene can facilitate a higher transcription rate.

1.8 Termination

The termination of transcription has different mechanisms in prokaryotes and eukaryotes. For example, the formation of a terminator structure in the pre-mRNA in *E. coli* leads to the dissociation of RNAP from the DNA. Recent developments take advantage of this simple and efficient mechanism and use competitors as orthogonal regulators of transcription [11].

The most acknowledged hypothesis of termination in eukaryotes was formulated by Tollervey in 2004 [93]. His "torpedo" mechanism acts as follows (Fig. 1.3): RNAPII passes the poly(A) site and introduces a cutting sequence for RAT1. After a successful cut, the pre-mRNA gets cleaved and freed from RNAPII. The next step of pre-mRNA modification is the addition of its poly(A) tail and the export to the endoplasmatic reticulum for translation. Meanwhile, RNAP keeps transcribing and Rat1/Xrn2 digests the newly synthesized "non-sense" RNA. Termination occurs when the RNase catches up with RNAP and the two proteins collide [49]. Teixeira *et al.* [91] found a co-transcriptional cleavage site at some genes, which result in the formation of a ribozyme with self-cutting functionality. Rat1 or Xrn2 is recruited and starts digesting the RNA directly from RNAP. The physical mechanism of the collusion on the DNA remains illusive.

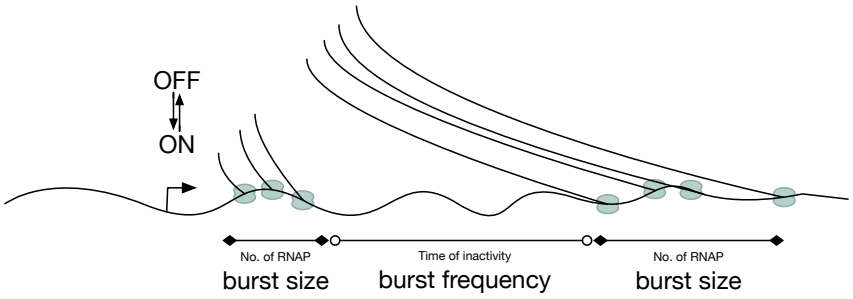

Figure 1.4: Scheme of bursting. The number of RNAPs dictates the burst size, while the time between two bursts is considered as burst frequency. This stochastic behavior results in transcriptional noise and is amplified by translation.

1.9 Bursting

Based on the randomness of the binding process of single molecules to start transcription, a term for this stochasticity was introduced together with mathematical models: the phenomenon of bursting (Fig. 1.4). A burst is a time interval in which a promoter is in an active state and yields RNA. Duration and frequency change drastically from gene to gene. In 2005, Paulsson [72] reviews this observation by multiple laboratories, often measured as huge noise in single cell experiments [26]. The telegraph model [29], consisting of two states for the gene (on and off) is used to explain the observed noise levels and it captures the stochastic binding of TFs to the gene. In this model, the gene switches at exponentially distributed times between the states and the number of mRNA produced in one burst is random, but follows a Poisson distribution [31]. The total amount of RNA depends on two parameters: burst size and burst frequency. Burst size describes the number of polymerases that start transcription with small pauses (Fig. 1.4). This parameter can be influenced by the activation domain of the transcription factor, or enhancer regions [83]. The burst frequency, the time between two bursts, can be linked to the strength of the DNA binding domain of the transcription factor. House keeping genes, where the expression only rely on the GTFs to bind at a TATA box, share the same properties and therefore also the same noise. Slight variations in the TATA box effect the binding of TFIID on the DNA, tuning the RNA level of the encoded gene. In the study by Dar *et al.* in 2012, the authors analyzed the expression level of randomly inserted GFP throughout the whole genome in a human cell line [16]. Based on single cell traces of GFP signal, fluctuations are analyzed, which can be decomposed to burst size and burst frequency. Highly expressed loci modulate predominantly the size of the burst, while low expressed loci modulate more often the frequency. Besides the finding that in all 8000 loci examined bursting is the predominant mode of operation, the authors conclude, that a threshold frequency exists, which can not be increased, but instead the burst size is tuned [17]. Further, tumor necrosis factor α has an effect on the on-state of a gene, and increases the burst size.

Advancing experimental setups have uncovered more and more details on the single cell level of transcription. With new data, models that explain the data had to be expanded constantly to provide insightful meaning. Several studies have proposed multiple states for a gene to explain the observations. In the case of induced transcription, either as a stress response or based on metabolic changes, two states are not enough to capture all dynamics. At least one extra off-state has to be added to the telegraph model to distinguish the long time inactivity from pauses in between bursts [90]. In 2016, Corrigan *et al.* used a direct read-out for the amount of mRNA produced in a single cell (see Section 2.3.1). The distribution of mRNA measured is best described by a continuum state model, where the gene, once active, can yield any number of products [13]. This might be the solution for heavily expressed genes, like the actin5 locus.

2 Metrology

What is needed to conduct measurements? When we look at the mechanism of transcription, which is a dynamic binding and dissociation of molecules, time resolved devices produce the most accurate data. Through sample drawing at different time points, the dynamics can be reconstructed at the population level. But when it comes to information obtaining from single cells, these techniques are inadequate. As transcription happens inside the cell, and in eukaryotes inside the nucleus, the device should utilize a method, which can penetrate the membrane without damaging the cell during the measuring time. Light is the most promising candidate with these properties. It is easily detected and interacts only with specific molecules. Microscopy is best suited for these kinds of light based single molecule measurements.

2.1 Microscope

The microscope has been developed for hundreds of years. In biological laboratories it is most often used in combination with fluorescent samples. The signal-to-noise ratio obtained through the Stokes Shift is far superior compared to standard absorbency or phase contrast with white light. Since the discovery of the green fluorescent protein (GFP) by Shimomura [85] the wide field microscope has become a major player for single cell studies. By specifically targeting proteins that are involved in the basic mechanisms of the cell, the microscope gives insights in the localization, movement and temporal abundance.

2.2 Methods to measure transcription elongation

Visualizing the transcription on a gene is a challenging task. As the polymerase moves along the DNA first attempts had tagged Pol II to watch it move. Apparently the density of pol II in the nucleus is too high to track single polymerases over time. Early studies of transcription have been performed therefore *in vitro* to track single polymerases or used FRAP techniques. Fluorescence recovery after photobleaching (FRAP) [65] uses strong point-wise illumination, until all fluorescent proteins in that region are bleached. The recovery is directly linked to the movement and diffusion of functional fluorophores fused to target proteins, replacing the damaged ones. In the study from Yao *et al.* from 2007 [106], the authors tagged the RNA pol II subunit p33 with EGFP to follow its recruitment in the cell. Upon heat treatment, pol II increasingly locates on the Hsp70 loci in drosophila flies. First estimations of elongation speed are at ~25 bp/s.

2.3 RNA-binding proteins

Directly targeting pol II has more drawbacks *in vivo* than advantages. A better approach is to visualize newly synthesized RNA. Best results can be archived with specific proteins, that recognize a RNA sequence. Their natural functions range from processing pre-mRNA, like splicing or poly(A)-ylation over translation to degradation. Those are common task inside a cell, so most of RNA binding proteins have low specificity and target most of the RNAs. But some proteins are found that have explicit high affinity and specificity for their target. The bacteriophages MS2, Qβ and PP7 have coat proteins that bind tightly specific RNA sequences and are utilized to tag RNA [2, 5]. Binding proteins from bacteriophages have the advantage of being orthogonal; they can be used with little to no cross-talk in eukaryotes. The natural function of MCP (MS2 coat protein) and PCP (PP7 coat protein) besides the coating of the virus genome, is the translational repression of viral replicase, through binding of an RNA hairpin that encompasses the replicase start codon. Despite their high affinity to RNA and analog function, the two proteins share a minimum homogeneity at the sequence level and the target structures have diverged evolutionarily. The coat proteins bind the RNA on the surface of a large β-sheet [58]. The visualization of transcription in living cells with coat proteins has revealed bursting as the mode of operation for various genes [55, 57, 66].

2.3.1 *Pseudomonas aeruginosa* phage PP7 coat protein

One key player for tagging RNAs is the protein PP7 from a *Pseudomonas aeruginosa* phage [10]. Compared with analogs like MCP it has lower probability to form inclusion bodies or create artifacts through not releasing RNA [94]. PCP is a 127 amino acid long protein and can be C-terminally fused to other proteins [68]. Its natural target RNA has the sequence `UAAGGAGUUUAUAUGGAAACCCUUA` (Fig. 2.1) [58].

The protein forms a dimer at the target structure. For the imaging of transcription, a fusion protein with one or two GFP variants is often used [55, 57]. In Larson *et al.* [55], the authors describe a method for the real-time observation of transcription initiation and elongation on an endogenous yeast gene. A constitutively expressed PCP-NLS-GFP fusion protein targets the pre-mRNA of the modified *GLT1* locus. The native *GLT1* promoter is exchanged with a cassette of HIS3 auxtrophobic selection marker, the cell-cycle regulated POL1 promoter and 24 repeats of the PCP stem loop sequence (PSL). Upon transcription of the construct, a diffraction limited spot is recorded with the wide-field microscope inside the nucleus as result of the accumulation of PCP on the transcription site (TS).

Following the intensity of the spot over time reveals the binding of multiple RNAPs. An autocorrelation function $G(\tau)$ is developed for the analysis of the TS intensity trajectories, which is a discrete atutocorelation over all the transcription probabilities for RNAP. The

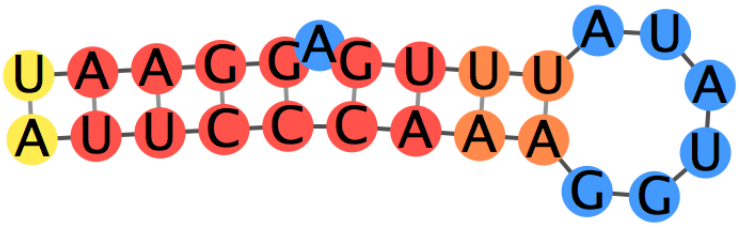

Figure 2.1: Structure of the PP7 stem loop. The color indicates the strength of the binding from weak (blue) to red (strong).

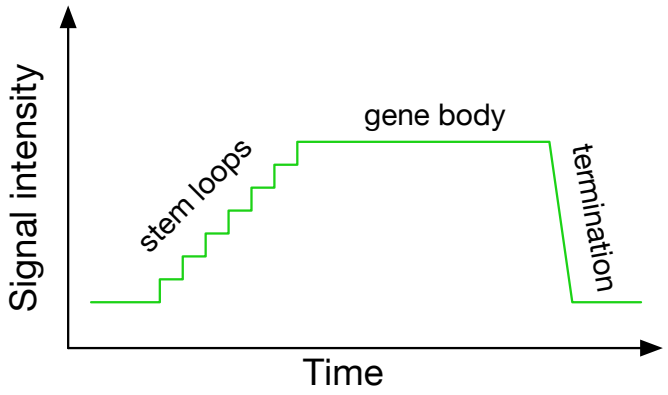

Figure 2.2: Single transcription event. After promoter release, RNAP processes the stem loop region, resulting in an step-wise increase in fluorescence after binding of PCP-GFP. A plateau is reached, when RNAP transcribes the gene body. After release of the pre-mRNA from the TS, the signal is lost. Multiple polymerases generate a superimposition of this signal.

function depends on the total dwell time of a transcript (T), which includes elongation as well as termination, and the initiation rate (c).

$$G(\tau) = \frac{T-\tau}{cT^2} H(T-\tau)$$

The autocorrelation decay is depicted as τ, H denotes the Heaviside step function, which is unity when $\tau < T$ and zero when $\tau > T$. The authors could estimate an average total dwell time of a transcript at the TS with 290 ± 30s. The calculated autocorrelation vanishes after the dwell time, which can be interpreted as non-existence of a transcriptional memory. With a second construct, the authors sought to decompose the signal into the process of initiation, elongation and termination. The target for modification is the housekeeping gene MDN1, with a length of 15kb. A cassette of 24 PSL is inserted in the 5'UTR after the promoter or respectively a cassette in the 3'UTR at the end of the gene. To compute the elongation velocity, the dwell times of the different constructs can be used, by the equation $v = \frac{L_{3'}-L_{5'}}{T_{3'}-T_{5'}}$. Although the data are based on single cell experiments, the autocorrelation calculates the average velocity for *MDN1* at 20 ± 8 bases per second. The authors point out, that cells in S/G_2 phase show a faster mean initiation rate, which would result in a higher estimated velocity of 46 ± 6.2 bases per second. This variable dwell time of single RNAPs is obscured by averaging the single molecule data and could be misinterpreted as RNAP pausing. For the gene targets in those studies, no bursting was observed as the single events of transcription appear to be not correlated or more specifically: they follow a stochastic model where transcription activation is determined by recruitment of a rate limiting factor. Based on those experiments, the laboratory of Robert Singer conducted a dual-color experiment, where the start and the end of transcripts are monitored with MCP-GFP and PCP-RFP [37]. The first finding in the work of Hocine *et al.* is the independence of transcript production from two alleles. Through mating of a *MDN1*-24xMSL and a *MDN1*-24xPSL yeast strain, the pre-mRNA of the two alleles can be followed over time and their abundance appear to be stochastic, intrinsically noisy and not correlated. This approach comes to similar conclusions as in Elowitz *et al.* [26]. Whereas the two alleles are indistinguishable by the environment and regulatory mechanism, their transcription initiation is no more correlated than any other genes and can be explained by a diffusive transcription factor binding model with low abundance. In addition to this, the authors tagged one gene with PSL and MSL in the 5' respectively the 3' UTR to directly measure the transcription dynamic on a single molecule and used the GAL1 promoter for induction. The found elongation rates range from 14 bases per second up to 61 bases per second, cleaned from cell cycle dependencies. The average velocity confirms prior work of the group with 25 ± 2 bases per second. The authors missed the opportunity to study the effects of different induction levels on the transcription dynamic and bursting parameter. In Lenstra *et al.* [57] the focus is on the galactose sensitive gene network. The biological setup includes MSL in the 5'UTR and PSL in the 5'UTR antisense direction of native *Gal10*. The aim of their work is to understand the abundance of non-coding RNA (ncRNA) within the *Gal7-Gal10-Gal1* locus. More interestingly for this

thesis is the evaluation on the burst parameters for this inducible system. Autocorrelation again is used to determine the dwell time in live traces and single molecule fluorescence *in situ* hybridization (smFISH) is used to estimate bursting parameters. These experiments are the first to show bursting in living yeast cells with an average burst size is 1.9 RNAs per burst and a frequency of 4.2 bursts per dwell time. Furthermore, the results of ncRNA and mRNA traces indicates an anti-correlation of the two species. If the traces of the sense strand are aligned to the begin of transcription, there is a 100 second time window where the production of the anti-sense ncRNA is completely absent. This can be an indication that the remodeling of the chromatin template around the promoter allows only one direction of operation. Interestingly, together with the transcription of the mRNA, ncRNA is found. An explanation might be that the passing of the polymerase in anti-sense direction is immediately followed by an initiation of the sense direction. A direct observation of the temporal order of those events is with this system not possible.

2.4 Microfluidic chips

For single cell measurements, a variety of designs for microfluidic chips have been proposed [15, 20, 43] (Fig. 2.3). A first separating characteristic is the generation of droplets or using continuous flow. On the one hand, the usage of droplets yields isolation of single cells and it is easy to get control and manipulation of the droplets [9]. Sorting devices are widely established and offer the possibility to enrich or screen populations [44]. The major drawback of the droplets is the inaccessibility of the cells during an experiment. Once the droplets are formed, the contents can hardly be altered. Time dependencies are restricted to an initial pulse in treatment. Some alternatives have been proposed [63], but this comes with a down scaling of throughput. On the other hand, continuous flow experiments have full control over the environment of the cell. Adding or removing chemicals is performed with ease. Adherent growing cells are placed inside the chip, can grow and undergo development and can be directly measured. The flow of new media provides nutrients and removes excreted cellular waste. The material of the chip (polydimethylsiloxane (PDMS)) is gas permeable: oxygen levels inside the chips are constant and will match externally applied gas composition [53]. For suspension cells like *Saccharomyces cerevisiae* it is a challenge to keep cells in one place and as monolayer. Simple devices immobilize cells between the glass coverslip and the ceiling [12, 56]. This generates a monolayer of cells, but lacks the possibility to wash away daughter cells. As a result, those chips cannot be imaged over longer time periods, as the field of view becomes clogged and cell segmentation becomes more and more challenging. A solution to this problem is the usage of traps, which are limited in size. In Ryley *et al.* [76], the authors developed small cages for cells. Mother cells have to be loaded during the assembly of the device, reducing the reliability of reproduction. The cage for cells consists of 3 rectangular pillars arranged in a equilateral triangular pattern, spanning an area of approx. 90 μm^2 (Fig. 2.3, A); gaps between the pillars ensure the removal of daughter cells in a

laminar flow. Reversal of the flow direction increases the separation from daughter cells from their mother. Overall, this device is capable of trapping mother cells and maintain them in a suitable environment. But the traps are big in structure and limit the number of cells, which can be recorded at each time point. In Crane *et al.* [15], the authors describe a trapping device, which can immobilize yeast cells in a continuous flow. "The device has three inlet ports upstream of a flow chamber in which the cells are observed. The central port is used for introducing cells; the other two ports are connected to programmable syringe pumps, which drive media flow through the device. Altering the relative flow rates of the pumps allows us to switch the medium in the flow chamber. This chamber contains an array of more than 1,500 individual cell traps, each of which consists of two PDMS pillars that extend from the ceiling of the device to the glass coverslip forming its floor (Fig. 2.3, C). The media flows through each trap from top to bottom with the pillars oriented such that the gap between the pillars is wide upstream, at the trap entrance, and narrows downstream. Refinements of this basic design have produced a simple device in which daughters are efficiently removed from the mothers with a low rate of failure. Once a cell is caught, the hydrodynamic resistance of the trap increases: diverting the fluid flow and creating a pressure differential across the cell from the top to the bottom of the trap. This low energy pocket, combined with viscous forces from fluid flow, prevents a cell from being pulled out of the trap. Diversion of the streamlines around trapped cells also creates a positive feedback loop during cell loading: cells bypass occupied traps in preference for empty ones, rapidly filling the array. The same process aids removal of daughter cells from the device during operation. Daughter cells are removed by the flow in two ways depending on the polarity of the trapped mother cell. If the bud is oriented towards the top of the trap, the diverted flow moves the bud to the side and later removes the resulting daughter cell. If the bud forms downstream of the mother, it is extruded between the pillars and separates rapidly from the mother on completion of cytokinesis. Mother cells in this position usually continue to bud through the gap because haploid cells form successive buds in close proximity." [15]

The authors claim to have highly loaded chambers over the interval of 20 hours. Clogging of the device is reduced by increasing the spacing between individual traps, leaving more space for cells to be washed away. The influence on the cells by the device itself is estimated by the viability, or more precisely the replication rate of the trapped yeast. In the time of 62 hours several hundred cells show no significant change in cell cycle. Although an increase in time between two divisions is observed, the cells remain vital in the physiological regime and this finding reflects the effect of aging. Several stress related transcription factors like Msn2p have been studied during the experiments in the device, but no unusual dynamics are reported. In conclusion, the Alcatras chip shows no influence on the behavior of trapped yeast cells, is capable of removing daughters and switch media on a small time scale. In 2015, Jo *et al.* published their results for the high-throughput yeast aging analysis (HYAA) chip [43]. The design has the shape of two Ls facing each other. Compared to the Alcatras ship, HYAA has better retention rates for trapped cells

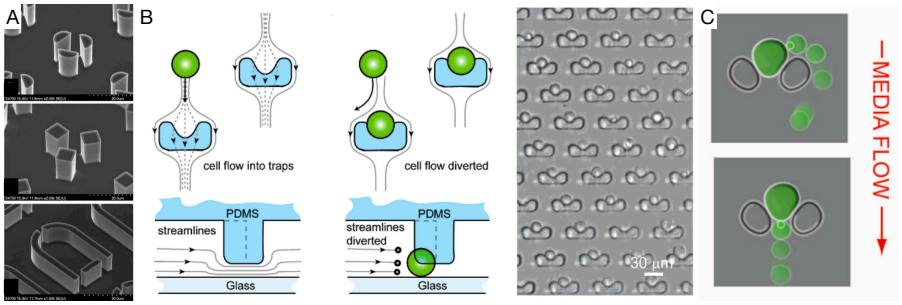

Figure 2.3: Collection of microfluidic traps. **A** Design taken from [76]. Yeast cells are trapped during assembly of the chip. **B** Picture taken from [20]. A two layer design traps the cells between the cover glass and the ceiling. **C** Design from [64]. Alcatras trap design helps to prevent clogging of the device as daughter cells are washed out.

and keeps 90% of the cells for days. The authors tested different row and column spacing in terms of clogging. Narrow spacing at 6 μm in each dimension comes with a great catch rate, but lowest probability of clogging is shown with 18 μm in each dimension.

3 Aim of this Work

My work is inspired by two ideas: The establishment of a cell trap chip in the laboratory to perform single cell experiment with yeast and the characterization of transcription dynamics during induction.

The first goal is setting up a microfluidic device and increasing the production efficiency. The basic design is received from Michael Unger [95] and includes a trap design, similar to the Alcatras chip [15]. One wafer of this design holds 6 chips, where only a subset is usable after the whole fabrication process of soft-lithography. To improve the fabrication, the size of the traps will be increased and the required layer with different heights will be reduced. In addition, the design will be reduced in size to fit multiple cell trap chambers onto one chip. The chip should serve as platform for collaborations with other working groups using *Saccharomyces cerevisiae*, who are interested in the temporal behavior of genetic switches and gates. Three different types of chips will be produced. The first is a copy of the old design with a size reduction. The second will place 4 cell chambers in close proximity for parallel measurements of independent experiments. The last design will have again 4 chambers, but all share the same media inlet. This way, 4 different yeast strains can be observed and will be exposed to the same media composition. The motivation is to increase the data output for overnight experiments. The images from the experiments have to be processed in an image procession pipeline. I will contribute to early stages of the pipeline and provide a suitable interface to apply improved, robust and fast algorithms on the raw single cell data.

The second goal focuses on the transcription dynamic in living cells. To achieve this, several repeats of PSL will be integrated genomically into yeast, while PCP-GFP will be constitutively expressed from a low copy plasmid. Upstream of an inducible gene, the mRNA tagging will produce transcription trajectories upon exposure of inducer. Those will be analyzed and interpreted for the responsiveness to β-estradiol as inducer with the GEV transcription factor. Further, the influence of properties of the transcribed DNA on the transcription dynamics will be investigated, like the deoxyguanosine and deoxycytosine content (GC content). Different genetic constructs with an homogeneous GC content ranging from 33 to 66% will be evaluated. Here again, the images of the experiments need to be processed and handed over to more complex data analysis. A tracker should follow the intensity of the spots on the transcription site during time in 3D and extract the intensity of a voxel. Those traces are subject to inference. The number of polymerases and their position should be estimated during the observation.

4 Results

4.1 The microfluidic chip as center for projects with *S. cerevisiae*

The microfluidic chip to trap yeast cells is a suitable device to begin cooperation within the department of biology and other working groups. The chip has been redesigned to increase the usability per chip and parallel experiments. The original chip as used by M. Unger [95] had one cell chamber on one physical chip. Media inlet channels were made with AZ4562, including a post bake step to round the channel with a height of 10 μm. The outlet channel had a height of 25 μm and cell loading channel were at 20 μm. The trap region was slightly higher than an average yeast cell with a height of 7 μm. In the new design, the heights of the cell loading and outlet channels are both set to 20 μm to decrease the fabrication steps and times. Three different layouts of the overall chip are tested. The first contains 5 old chip designs re-scaled to fit on one chip. This allows the same chip to be used up to 5 times for independent experiments, which is a huge increase in production speed and a reduction in PDMS waste. The mixing channel after the valves is adapted from the former design to guarantee the same mixing behavior for the *pulse width modulation* (PWM). The second design has four independent cell chambers, all lined up together in the center of the chip. Each trap has its own inlets and outlets to conduct separate experiments with e.g. different media concentrations, but on the same strain. The close proximity of the chambers allows the stage to move to the next trap within the shortest possible time. The third design also has four chambers positioned closely together, but the input channels are shared. This chip enables the direct measurement of up to four different yeast strains with the same experimental conditions. Each trap chamber has a separate cell loading channel and outlet to limit cross-contamination. The cell loading channels are relocated from the sides to the front for the four parallel and four independent chip. These three designs are tested with trap areas optimized for the 60x and 100x objective field of view (FOV) (Fig. 4.1B). The fields with traps for the 100x objective are arranged in an 5x5 pattern, with the same geometrical structures for separation as the 60x. In total, one wafer holds the negatives for 6 chips, three times 60x variants and three times 100x, and 26 cell chambers for experiments. The layout of a single trap itself is basically kept the same among the chip designs. The only change targets the back of the trap and elongates the body, resulting in a bigger shape for better fabrication (Fig. 4.2 4 (old) and 5 (new)). Besides the changes in the design, the fabrication of the wafer with photo-lithographic methods in the facilities of the institute “Elektromechanische Konstruktionen” (EMK) has come with problems. The size critical features of the chip, the traps, could not be fabricated repetitively with the same quality (Fig. 4.1). The traps, which require a chrome mask for the exposure to UV light, are not sufficiently washed out from the unexposed photoresist SU-8. This results in hanging traps, which

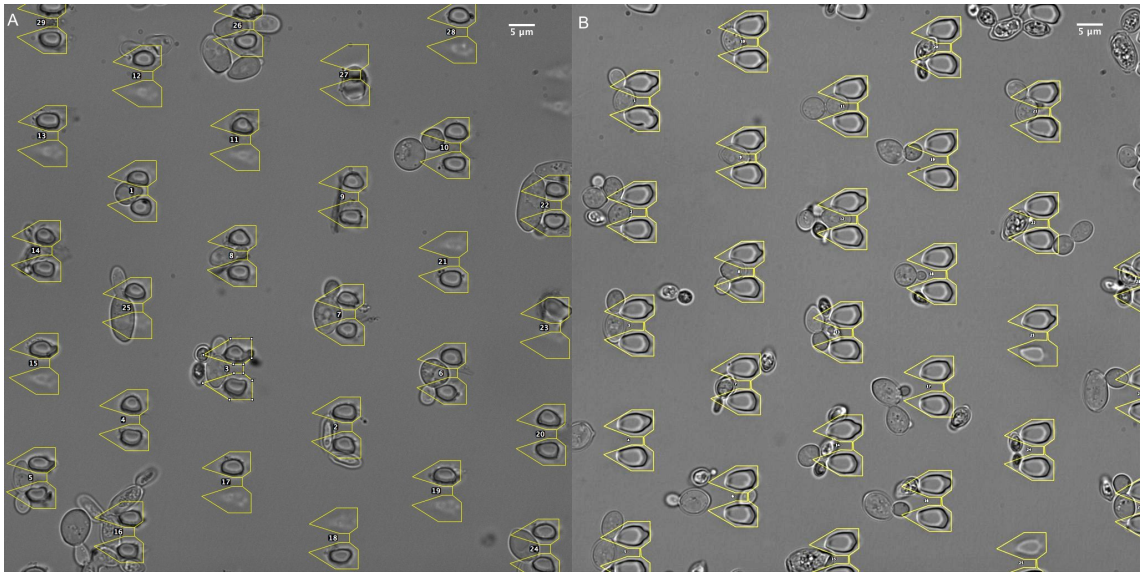

Figure 4.1: Examples of fabrication errors. The desired trap design is highlighted in yellow for two different rounds of wafer producing under the same settings.

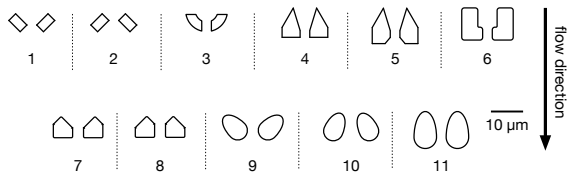

Figure 4.2: The second design iteration. Traps 1,2,6 are directly adapted from [15, 43]. Flow direction is from bottom to top. Scale bar equals 6 μm.

are not able to catch yeast cells at an acceptable rate. Only two wafers in initial trials are produced despite great efforts in the exploration of all fabrication parameters and especially for exposure time and development duration. Since one wafer is robust enough to fabricate hundreds of chips, further investigations were postponed until the next round of design.

With the second iteration of design, the shape of the traps is changed to bypass the fabrication problems. Several different designs are tested on the wafer (Fig. 4.2). Some designs are taken directly from the literature [15, 43], despite their dimensions being smaller than the limits I encountered. The area of the traps ranged from 15 μm^2 to 46 μm^2. The inter-trap spacing is adjusted to the new dimensions, lowering the total traps in a FOV to 12 for the biggest structure with the 100x objective and 30 for 60x. Compared to the older designs (Fig. 4.2, 4+5), the oval traps (Fig. 4.2, 9-11) are designed to bypass the production limits as the total size of the trap is increased and it has no sharp angles. The L-shaped traps followed the same dogma: increased area, decreased sharp angles [43]. The half arcs and the design of Alcatras have significantly smaller areas, but also consist only

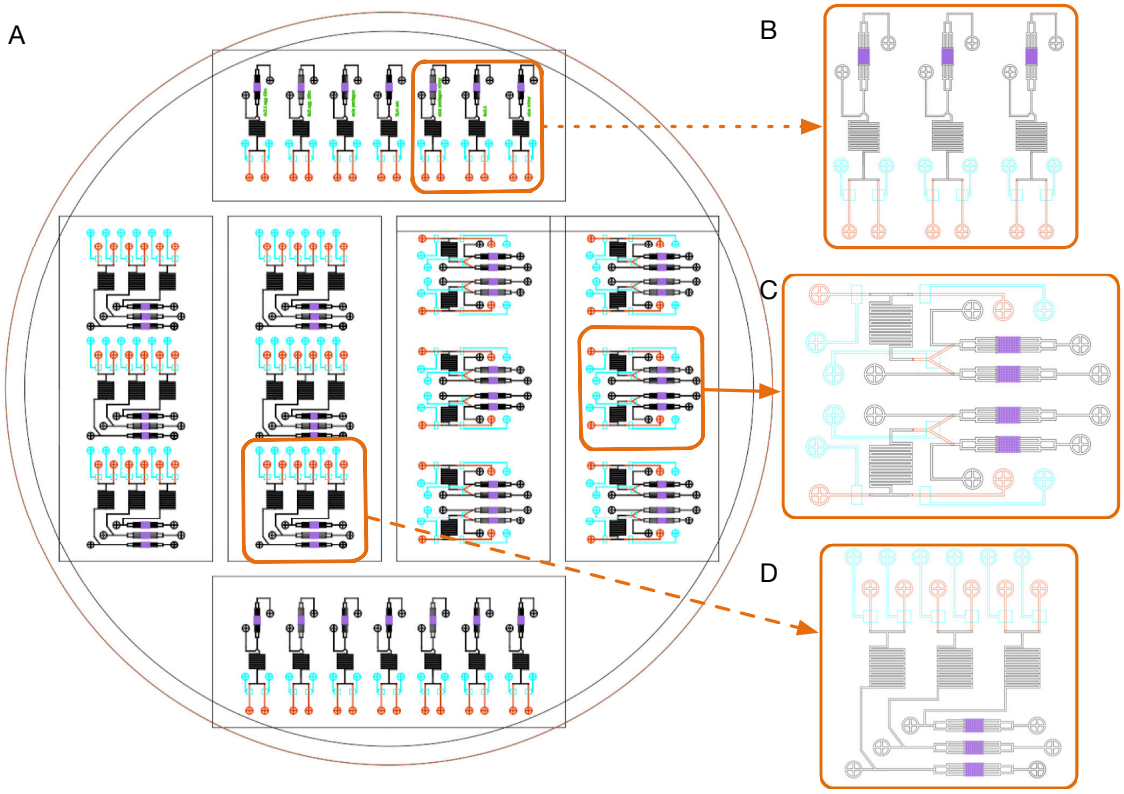

Figure 4.3: Sketch of the wafer (100 mm diameter) for the second design iteration. **A** The colors correspond to one layer of fabrication: cyan is the control layer, red is the AZ layer, black is the SU8 2025 and purple the cell traps with SU8 2005. 3 different chip layouts are tested: **B** 7 independent chambers **C** 6 x 2 dependent parallel chambers. The 6 x 2 chambers have in addition a valve to separate the two strains during the loading process. **D** The last design 3x3 independent parallel cell chambers

of 90° angles. Two versions of pentagons are based on the Alcatras, and I increased the overall body area. The first version is pointing outwards with a greater distance between the tips; the second is basically a mirrored version with inwards pointing tips. Two parallel standing oval traps have achieved the best fabrication results (Fig. 4.4F).

Each of the three chip variants are further optimized in terms of usable cell chamber per fabrication (Fig. 4.3). The 5 chips are shrunk even further to fit 7 independent experiments with a cell loading channel coming from the front in a Y-shaped junction (Fig. 4.3, B). The four independent chamber layouts are moved and reorganized to fit 3 times 3 chambers together (Fig. 4.3, D). In respect to the demands of the experiments, the four parallel chamber chip are redesigned to have 2 chambers connected to one inlet (Fig. 4.3, C). A Y-junction with an overlaid valve ensures no cross-contamination. Such a layout is mirrored to produce an area with four chambers in close proximity. Further space optimization leads to a total of 12 chambers on one physical chip. In all new designed chambers, the

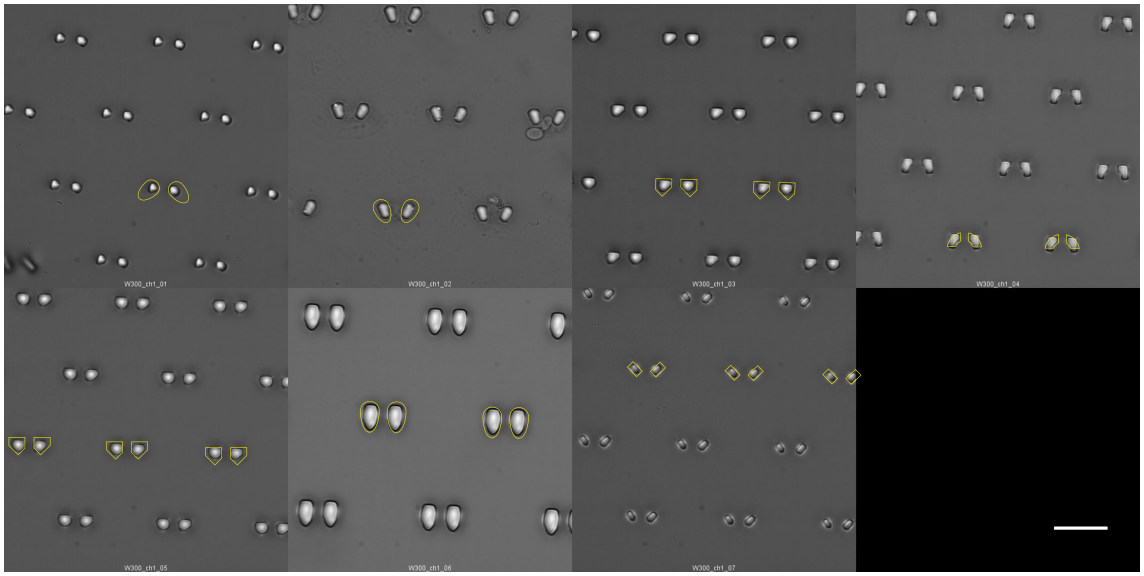

Figure 4.4: Montage of traps made with direct laser writing. Drawn in yellow are the specified designs, images taken at 100x . Resolution and section are matched. Scale bar equals 20 μm.

orientation marks and separation structures are removed to make the chip usable with the 60x and 100x all the same. The distribution of yeast cells throughout the chamber revealed, that the separation in different field of views (FOV) does not yield advantages, besides stitching of all captured images. Although the structures in between should not catch cells, bigger aggregates got stuck here. The fabrication process is pushed to its limits with the small traps and could not be successfully completed with all designs. As expected, the bigger oval structures show the best results with the used mask aligner, although only a fraction of the chips produced are usable. The sources of errors remain hidden, but changes in the device itself, like UV lamp exchange, repair of positioning table and contact mode could have been influential. A solution was found in a direct laser writer (Fig. 4.4), omitting the step of generating masks. The drawing is directly uploaded to the writer and the reproducibility is enhanced by eliminating the need to correctly place lamp, mask and wafer. Although the structures are not perfectly written, the chips produced from this wafer are usable for experiments.

A typical experiment for the transcription elongation project has 70-85 % traps filled with yeast (Fig. 4.5 and Fig. 4.6). I do not distinguish between traps that are not loaded with cells and broken traps caused by fabrication errors. Both situations lead to a smaller number of cells that can be captured and limited the experimental output. As the loading efficiency is similar in most of the experiments, little knowledge about the trap design on the loading efficiency can be gained to improve the design process. The rate of catching and maintaining cells is comparable to the Alcatras chip. Cells are caught most commonly with a bud (33% total, or 42% of trapped), if the bud of half the size of its mother is

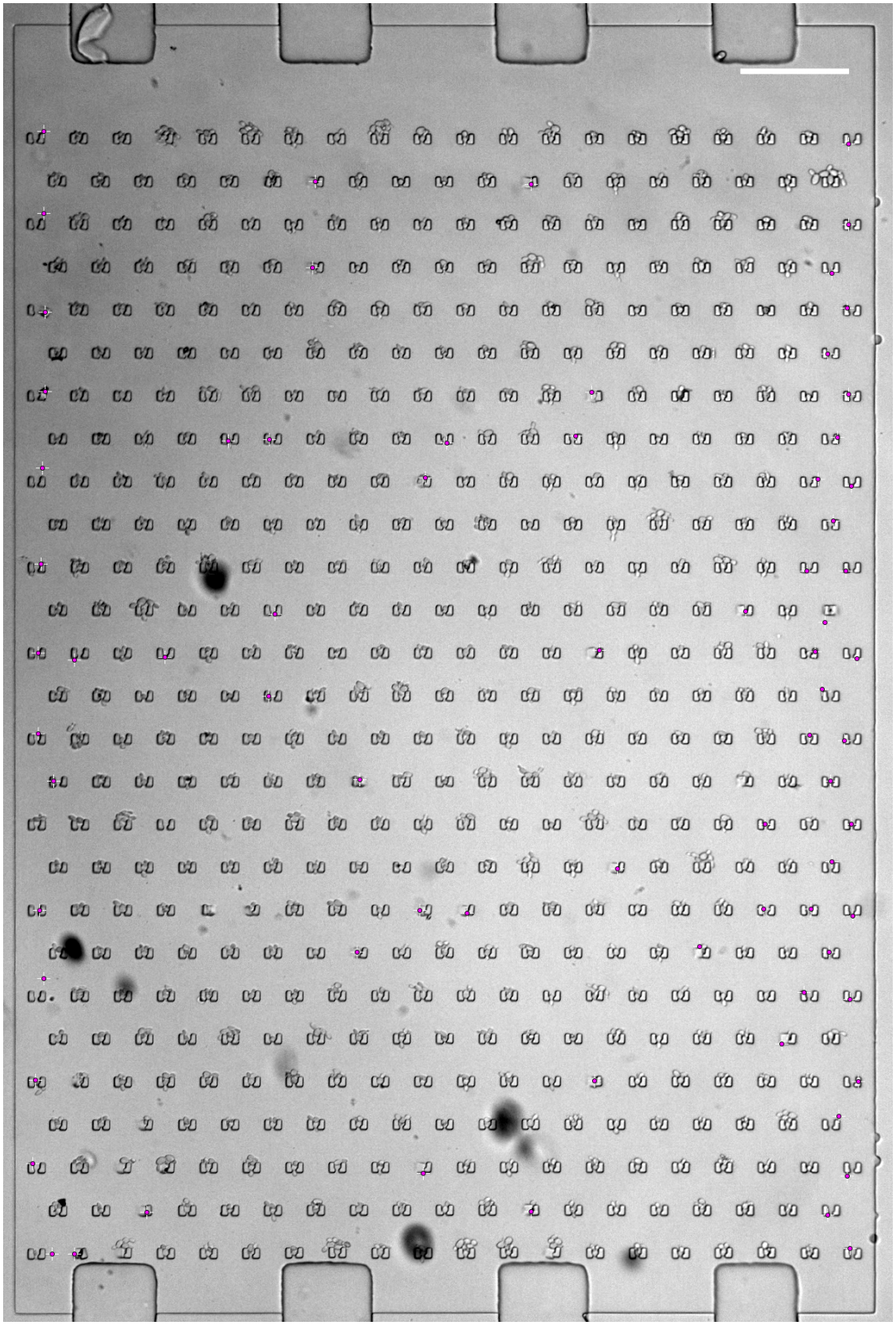

Figure 4.5: Example for a loaded chip before an experiment. The L-shaped traps show excellent trapping efficiency. Of the 527 possible traps, 22 have fabrication errors and 53 on the sides of the chip are not filled with yeast (purple dots). Scale bar equals 100 μm.

counted as cell. In 18% of the traps (25% of filled traps), I found the desired single cell. Bigger cluster sizes (3 and bigger) are found in 25% of the traps with fewer counts for bigger aggregates. The overall loading efficiency depends on multiple factors: the yeast strain itself, cell density and growth conditions as biological factors; trap design, flow rate and chip height represented some mechanical factors.

4.1.1 Biological factors on loading and clogging

For the demonstrated device, the yeast strain should ideally be composed of mostly single cells. The budding of yeast leads to a grouping of mother and daughter cells, which can become problematic for the chip in longer experiments. The yeast strain BY4741 and its descendants have a deletion in the gene FLO1 that is known to be involved in flocculation [101]. In the transcription speed project (yJD), the BY4741 strain is the genetic background, and single cells should be found widely, if grown in suspension. In my experiments, most of the cells come in a cluster of 2-4 cells (Fig. 4.5). The daughters still stick to their mothers and further processing is required to receive single cells in the majority of cases. However, this can often be skipped, because the length of the recording in the transcription project rarely extends 4 hours, during which a clogging of the device is not observed. The yeast strain used in cooperation with C. Schneider has the genotype MATa/α ade2/ade2 his3/his3 trpl/trpl leu2/leu2 ura3/ura3 (RS453, with neo-tc gate: yCS) and no additional flocculation reducing deletions. After a time course of 5-6 h, the chip is at its limit of traceability and operability (Fig. 4.25) caused by clogging. Separation procedures like enzymatic cell wall digestion or ultra-sonication are omitted in order to reduce the stress on the cells before the experiment. In the potassium channel project, leading scientist S. Höler modified yeast in the background of MATa/α his3/his3 leu2/leu2 ura3/ura3 (ySH), also with no additional deletions of flocculation genes. But this strain yields the highest percentage of single cells.

For all experiments, the best cell density (measured in optical density at the wavelength 600 nm (OD_{600})) for loading is found between 0.05 and 0.1. At an OD_{600} of 0.05 cluster forming cells (yCS) have still small clusters (2-6 cells), which fill the traps without clogging the device directly during the loading process. A higher cell density can be used for cells that form no or small clusters (yJD, ySH). Here, the time needed to fill the traps is relatively small, so a pre-chip dilution is suitable. In general lower OD_{600} have a higher fraction of single cells or doublets, but also a higher population of dead cells, when not cultured long enough. Best results are obtained, when the cells are diluted from an overnight culture to 0.01 and re-diluted at an OD_{600} of 0.1 to 0.03. Diluting the cells at this point does not open clusters and generate more single cells, but a bigger volume is

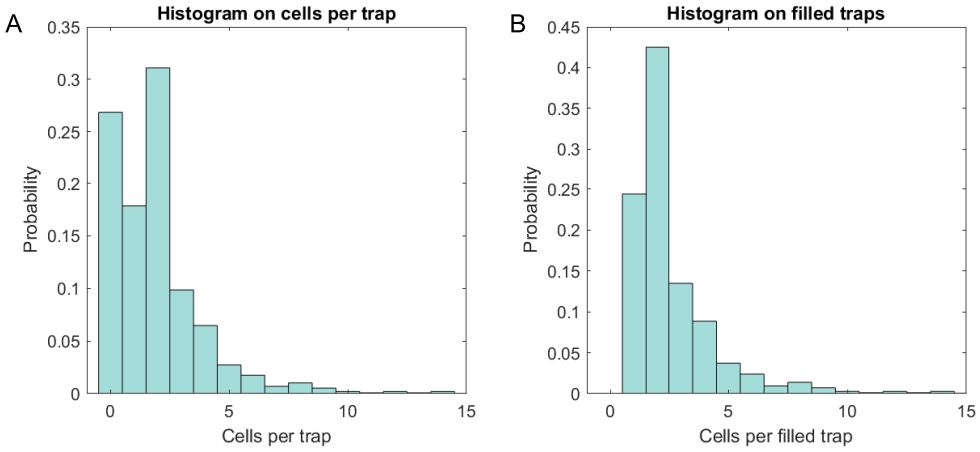

Figure 4.6: Histogram of the loading efficiency for the transcription experiments. In total, 5000 cells are observed in 100 FOVs. **A** 74 % of all traps are filled with cells. for the strain yJD, the most common caught cluster size is 2 cells (33%) followed by single cells (18%) **B** Statistics without empty traps. 25% of traps with cells are occupied by one single cell, 43% have caught a doublet. The remaining 32% of captured cell aggregates consists of 3 cell or more.

necessary to bring the same amount of cells to the chip. The extra time spent in loading the chip grants more control over the loading process and the experimentator can stop at a suitable loading ratio. However, this procedure is not practicable, when the cells are light sensitive as in the potassium channel project and the timing of the cells in the light influences the desired experiment. A reverse strategy is used: the cells are diluted to OD_{600} of 0.01, grown to 0.05 and collected in the centrifuge to artificially increase the density to 0.1. This way, the loading process can be shortened and a light exposure of the cells minimized.

As clogging is the result of too fast growing cells in the chip, the growth conditions and properties of the used strains are key points in operating the chip for longer time periods. Yeast has slight variations in its doubling time based on the carbon source in the medium. For the different projects, various sugars as carbon sources are used in the media and result in diverse growth rates. For the three strains the doubling times ranges from yCS = 80 min, yJD = 90 min to 110 min for ySH. Cells of yJD grow in media containing raffinose, which required additional proteins to be digested into its three monosaccharides galactose, glucose and fructose. The media for ySH had to contain additional K^+-ions to compensate the deletion of native potassium channels.

4.1.2 Physical parameters for clogging

In addition to the growth rate and likeliness of flocculation, the chip itself had some parameters to be tuned. A quantitatively study of the different trap designs could not be

finished as the fabrication is not completed at the time of this thesis. In theory, the first change of design, extending the trap body, reduced the gap between two rows of traps, leaving less space for cells to be washed out. In practice, this is negligible in the available data from the three different strains. Nevertheless, as clogging is observed in parts of the experiments (more often with yCS), the spacing is increased in the second round of design to counterbalance the trap size.

The height of the chip should lay in the range of one diameter of a yeast cell (~6 μm). When the chip is higher than 8 μm, multiple yeast cells have been found on top of each other, increasing background intensity and the chance to clog the chip. Besides the fabricated height, the pressure inside the chip can influence the cell chamber ceiling. Since PDMS is elastic, the pressure within the chip can expand the channels and chamber, especially during the loading process, when an external syringe is flushing in liquid in addition to the flow of new media from the reservoirs. On the one hand, the flow rate determines the chances of washing away daughter cells. The best flow rate for experiments is found, when the reservoir had an applied air pressure of 1 bar to bring liquid to the chip. The estimated flow rate at 1 bar equals 4 ± 0.2 μL/min and was enough to keep the chip running for several hours. On the other hand, a high flow rate can flush out already trapped cells by chance or during budding, when a daughter cell is exposed to the flow and does not break the connecting with its mother.

In conclusion, the different trap designs are all suitable for trapping and maintaining yeast cells. Based on the data presented here, no advantages or disadvantages for the experiment is found. The smaller designs have influence on the fabrication quality and bigger structures are repeatedly produces with better resolution. The flow rate at 4 μL/min has the best trade off between flushing out cells unintentionally and removing daughter cells to prevent overgrowing of the device.

4.2 Dynamics of transcription in yeast

The biological importance of this work is the characterization of the transcription dynamics derived from microscopy images. A minimal-system for this purpose requires a way to display the amount of mRNA that is produced. Additional components, like the transcription factor or the translated protein help to gain more insights about the dynamics. The first experimental design includes a *Gal1* promoter, several repeats of the PP7 stem loop, followed by a coding sequence for a fluorescent protein and the usage of the synthetic transcription factor, consisting of GAL4 binding domain, estrogen receptor and VP16 (GEV) [64] for activation. The ideal case in which all three quantities can be imaged is withdrawn rather quickly for practical reasons. The system is reduced to monitor the mRNA production and the localization of the transcription factor. The option to display the translation and therefore the coding sequence of the fluorescent protein after the stem loops is skipped because of the three following reasons: (I) in order to work with PCP-GFP and GEV-mCherry, (II) since the 5' UTR stem loop inhibit translation of the

following gene and (III) no adequate fluorescent protein could be found to match with the spectra of GFP and mCherry while being bright enough to be constantly monitored. Blue fluorescent proteins have poor quantum yield and their excitation with violet light is harmful for living cells. Cyan alternatives overlap strongly with GFP, making their usage inefficient. Spectral relevant are far-red fluorophores, but again, the quantum yield and brightness is problematic for those proteins. Furthermore, the installed light source does not have the correct excitation wavelength and filters. Another option is the switch from the GFP-mCherry pair to a CFP-YFP pair. The spectra of CFP and YFP have overlapping emission from CFP and excitation from YFP, generating a suitable FRET combination. This narrow band width is a disadvantage for the measurement. The suitable filters for CFP and YFP leave ~30 nm to measure CFP and ~40 nm to measure YFP, allow space for an additional third fluorophore in the red spectrum. In comparison: the GFP-mCherry pair has combined a spectral range of 100 nm usable for detection. The expected signal of this single molecule experiment is at the limit of the sensitivity of the camera; for this reason the GFP-mCherry pair is preferred and used (Fig. 4.7).

The construction of the gene cassette was more challenging than anticipated. As the cloning and insertion of the repetitive stem loops was unsuccessful with multiple primer pairs, different polymerases, reaction buffer compilations, DMSO and Mg^{2+} concentrations or classical restriction site cloning (for details see 8.4 and Fig. 4.8 and 4.9), I received a strain from Tineke Lenstra from the Larson Lab, NIH Bethesda, Maryland, USA, with the genotype: MATa/α his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 met15Δ0/MET15 ura3Δ0/ura3Δ0 14xPSL-gal10. Those cells were transformed with a GEV-mCherry containing plasmid to yield β-estradiol inducible cells. GEV with its Gal4 binding domain exceptionally activates Gal1 and Gal10 [64], making the engineering of the genome of this strain redundant. Together with the transformation of GEV-mCherry, PCP-GFP-GFP is introduced into the cells. This fusion protein includes a nuclear localization sequence (NLS) to directly label newly synthesized mRNA with the PCP domain (Fig. 2.1 and Fig. 4.7). The intra-cellular levels of PCP are kept low with a low copy plasmid (ARS-CEN) and a weak constitutive promoter (MET25) to reduce the detection limit for the transcription sites.

A typical theoretical trace (Fig. 2.2) shows a rapid increase in intensity, followed by a short plateau before it diminishes into the background. The RNAP moves with approx. 20 $\pm$ 8 nt/s along the DNA harboring the gene [55]; given the length of the construct (840 bp for PSL and 2100 for Gal10) one would expect a signal length for a single polymerase for 100-200s (10-20 frames) reduced by a minimum of PCP-GFP which has to accumulate on the mRNA to become detectable (Section 5.2). In a naive approach, one can guess where and when transcription has started by the trace alone. To estimate the number of polymerases at the gene at each time frame is a more complex problem. Compared to a trace from a measurement (Fig. 4.11) the plateau is hidden in noise and might covered in overlaying signals from multiple polymerases. The first results obtained with the system show heavy bleaching although the GFP signal is rather at the end of sensitivity for the

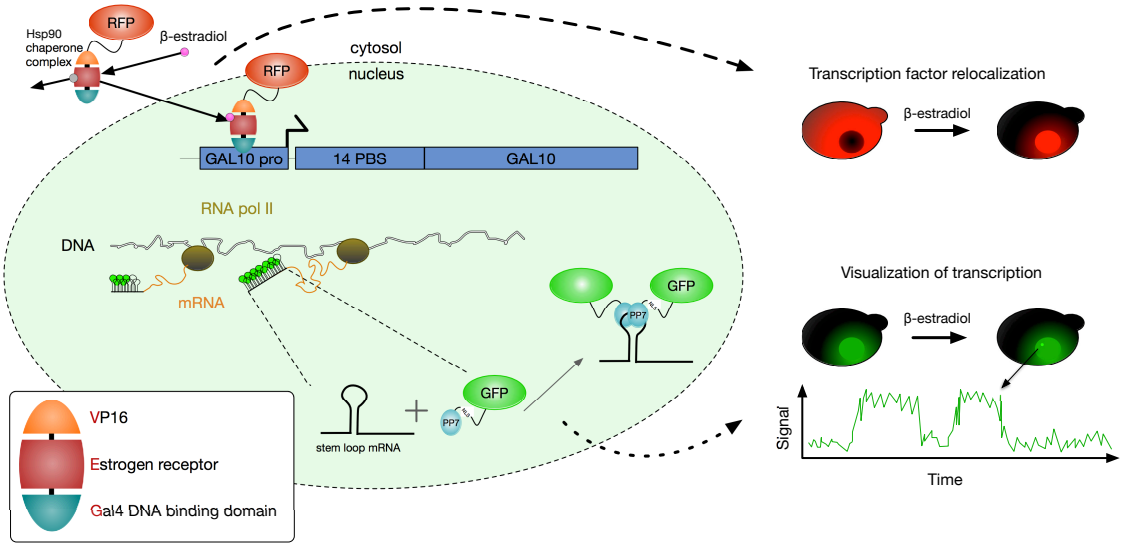

Figure 4.7: Scheme of experimental system. PCP-GFP and GEV-mCherry are constantly expressed and present in the nucleus and the cytosol respectively. Upon exposure with β-estradiol, GEV is relocated into the nucleus to start transcription on the GAL10 locus. This re-localization can be observed in the red channel (right side, top). Once transcription is activated, the stem loops are synthesized and can be bound by PCP. This event can be displayed as spot appearing inside the nucleus (right side, bottom). Following the intensity of the spot over time describes the dynamic of transcription.

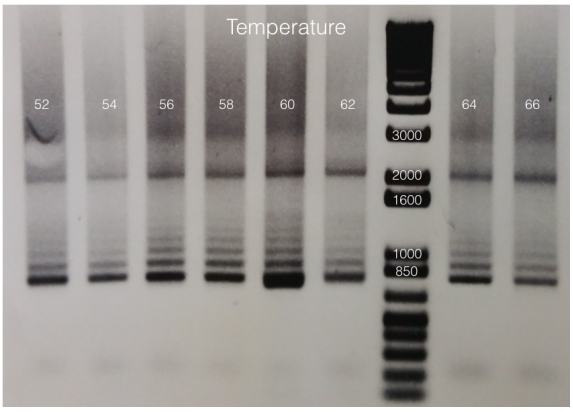

Figure 4.8: Example of agarose gel with PSL product. The used primer pair binds 200 bp upstream and 300 bp downstream the PSL region to minimize side products. Annealing temperatures tested range from 52°C (right) to 66°C (left). The product is expected to run at 2100 bp.

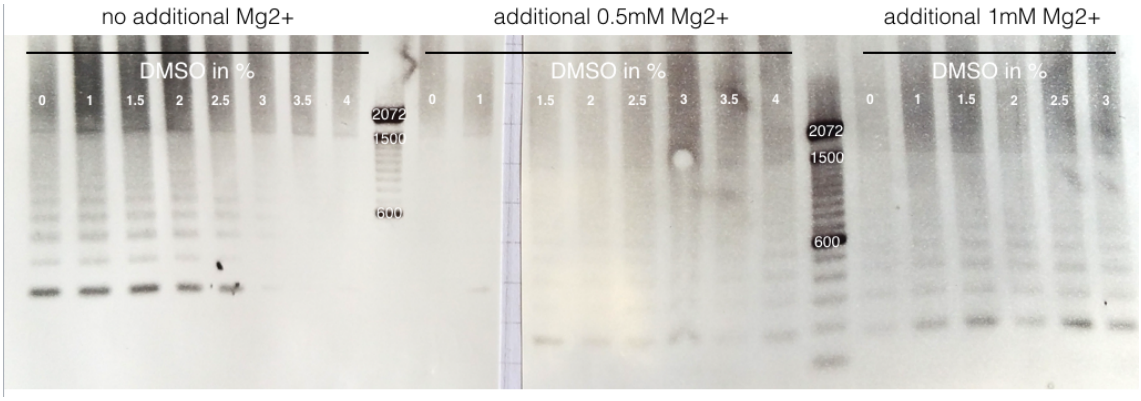

Figure 4.9: Example of agarose gel with PSL product. The product is expected to run at 1600 bp. DMSO is tested ranging from 0% to 4% under 3 different Mg^{2+}-concentrations (0, 0.5 and 1 mM additional).

camera to start with (Fig. 4.11). Only two positions at once can be recorded with the interval of 10 s. Due to the loss of signal over time, the recording duration of 1 h is not feasible with this imaging setting. Resulting traces hold barely any information in the last third of the experiment (Fig. 4.12**A**, between 40 and 60 and 100 and 120 min). As an alternative, multiple positions are recorded sequentially for 20 min each (Fig. 4.12). The pattern is optimized to minimize cross-bleaching of a position to the next (Fig. 8.1). When one position is recorded, the direct neighboring traps are visited at earliest three segments later, giving the GFP time to recover from possible bleaching (Section 8.8). The information density increases every time the position is switched to the next (Fig. 4.12**B**). The bleaching is visible in all positions and the total level of intensity is not biased by the position. At time $t_0 = 0$ s, β-estradiol is introduced to the cells, which show directly an increase in activity.

In most cases, the real data (Fig. 4.11) differs in definiteness of signal from the theory (Fig. 4.7). The PCP-GFP expression level inside the cells is rather uniform (Fig. 4.16) and is not necessarily higher in the nucleus. The weak promoter on the low copy plasmid provides enough protein to target transcription sites and has a low background. The correlation between the PCP level and the trace intensity is not significant (Fig. 4.16) and one can assume that PCP is not limiting the detection of the spot. The re-localization of GEV takes longer than the actual activation of transcription (Fig. 4.10). This observation can be explained by the fact that a single relocated transcription factor can already start transcription, while the nucleus has to hold more mCherry molecules than the cytosol for a noticeable relocation. The cell shows an active transcription site just 50 s after induction with β-estradiol, while most of the transcription factor is still in the cytosol. Although the TF is already present in the nucleus, transcription is not active all the time.

The dynamics of transcription for one induction level are in agreement with literature in terms of elongation speed and burstiness. A spot, once it exists longer than 200 s has

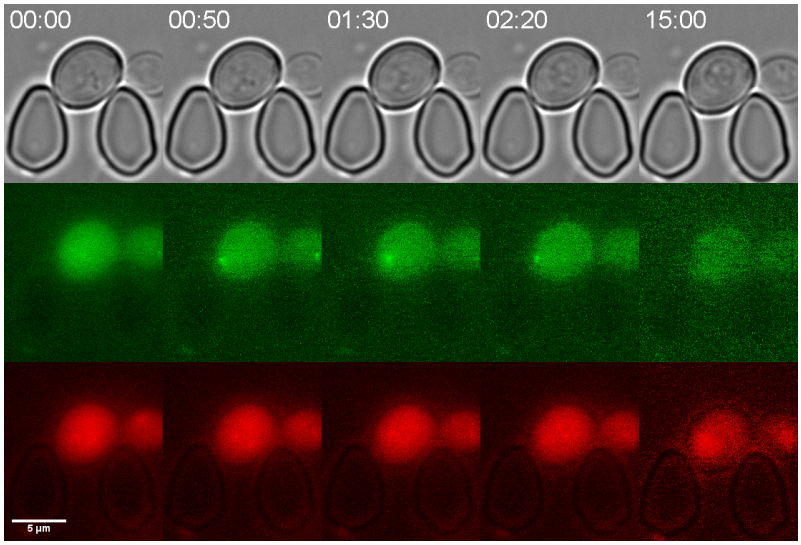

Figure 4.10: Montage of a cell in the channels bright-field, GFP and RFP from top to bottom. The appearance of a spot can be spotted in the GFP channel on the lower cell. The re-localization of the transcription factor is only visible at the end of the recording (after 20 min).

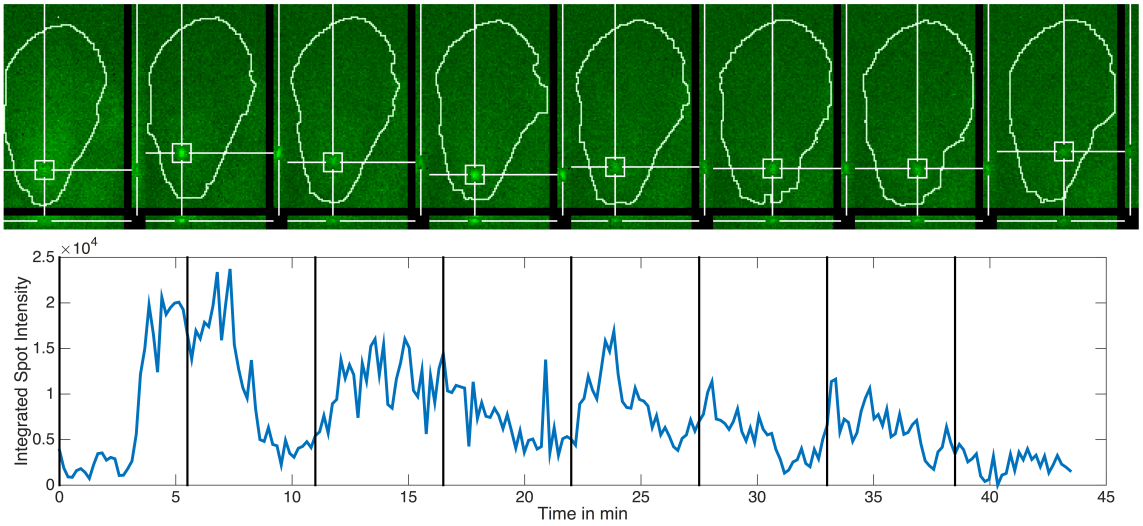

Figure 4.11: A typical measurement. Top row shows the cell in a maximum projection, cell membrane highlighted in white. The corresponding intensity values are beneath each frame and marked with a vertical line. Several transcription event can be observed in this cell.

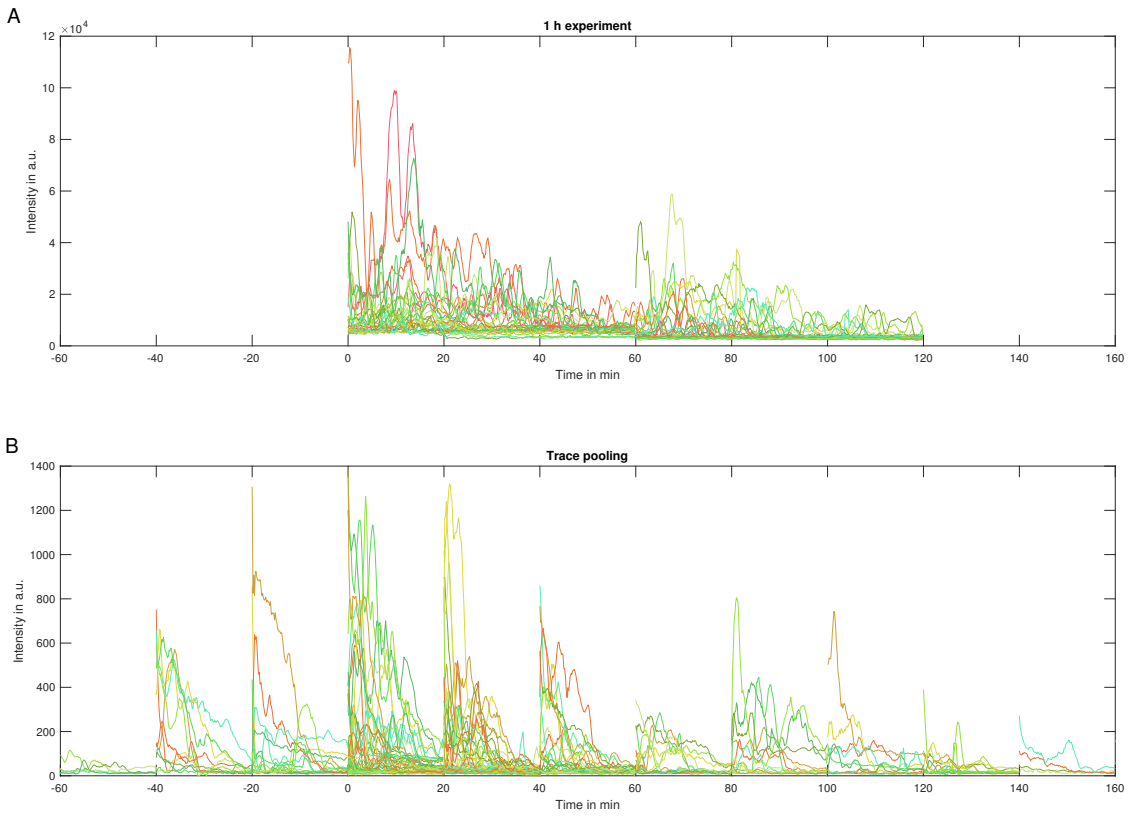

Figure 4.12: Raw data wit different recording times. **A** Single cell traces for experiments with 10 s interval over the period of 1 h. Strong bleaching is visible and the last third of the trace hold little useful information. Two sequential time points are plotted**B** Pooled single cell traces for sequential recordings of 20 min. The total information captured in 1 h is increased with shorter movies at different locations. Each line represents a trace from a single cell and the different colors are for a better visualization. The fast time decay of the signal is caused by photobleaching.

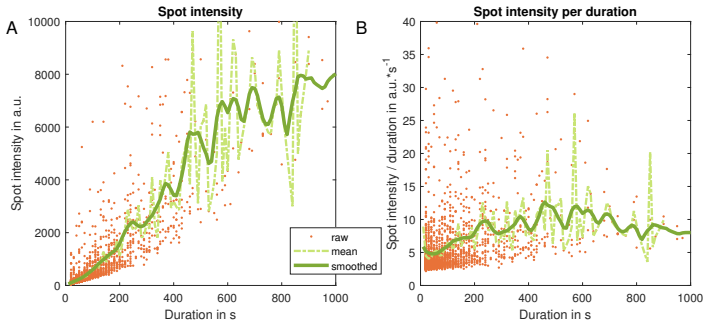

Figure 4.13: Spot intensity by its duration with 50 nM of β-estradiol. **A** The accumulated spot intensity is increasing with the duration of the spot. **B** Even when normalized by its duration, a correlation at the beginning is observable. Shorter spots can include abortive events and those have a lower spot intensity as less stem loops are present.

no significant increase in its average intensity (Fig. 4.13). Shorter lasting spots might be abortive transcription events, which have not transcribed all stem loops and accumulate less PCP-GFP. The average duration for spots in a physiological regime is 198 s $\pm$ 70 s. The more interesting case is the comparison of different levels of induction with varying concentration of β-estradiol.

4.3 Dose dependency of transcription dynamics

The properties of bursting dynamics under non-stationary conditions are not well characterized. Therefore, performing multiple experiments with different level of induction with β-estradiol are insightful for the transition from non- to stationary expression levels. I choose a variety of inducing conditions starting from 1 nM to 1 μM to observe changes in the transcription dynamic. The expectation is a rise in number of cells, which respond to the stimulus, with increasing concentration of stimulus. In addition, a single cell should be transcriptionally more active at higher concentrations. Despite the fact that some publications have reported which parameters change, a quantitative study has yet to be performed. In Larson *et al.* [54] the authors used four concentration of ponasterone A to induce mRNA production in cells. During their evaluation of spot intensity and on/off times of transcription, a positive correlation between induction and on-times (spot appearance) and total spot intensity (summed up over time) was not found.

The first observation, which can be quantified with increasing induction is the fraction of cells that show spots (Fig. 4.14). The single cell traces from single experiments do not support more statements. Although the experiment with low β-estradiol concentration could have less intensity, also the total number is cells is not comparable (Fig. 4.14**A**). If all experiments are pooled together in a cumulative fashion a better estimate of the underlying truth can be made (Fig. 4.15). Below 10 nM β-estradiol, the number of cells seems barely to be influenced by the stimulus. Although the negative control showed less

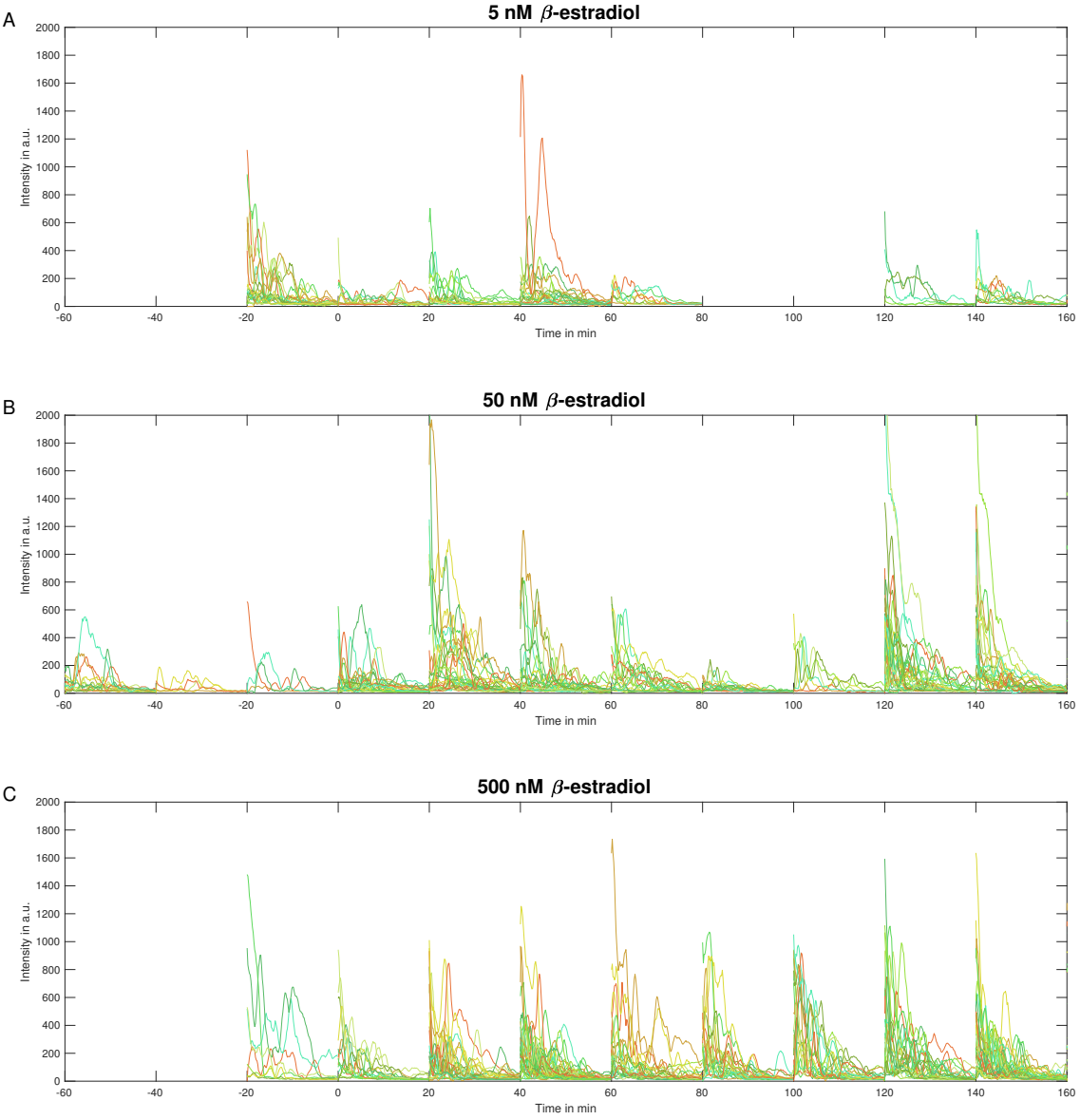

Figure 4.14: Raw data from experiments with different β-estradiol concentrations. **A** 5 nM, **B** 50 nM, **C** 500 nM β-estradiol. The increasing number of responders is visible in the higher number of traces generated by one experiment (109 to 271 to 301). Each line represents a trace from a single cell and the different colors are for a better visualization. The fast time decay of the signal is caused by photobleaching.

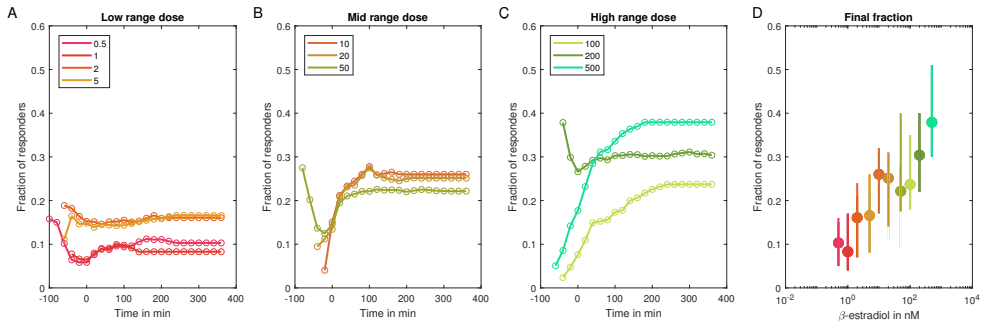

Figure 4.15: Time-resolved dose-dependent transcriptional activity. The cumulative fraction of cells showing a spot per total cell count for the time since β-estradiol (time 0 min). The curve converges to the average for each dose, as all cells are taken into account that have shown activity from former positions. **A** Lower concentrations (0.5 to 5 nM β-estradiol, pink to orange). The curves show little influence of the inducer on the cells. The fraction of responders does not change significant around time point 0. Even before β-estradiol is introduced to the chip, cells show activity. **B** For 10 to 50 nM an increase at time 0 is visible. **C** The highest concentrations result in the highest number of induced cells (500 nM). Each point is averaged over at least 3 independent experiments. **D** Average fraction of active cells per dosage. A positive correlation is visible.

activity, the observed spots might have their origin in the leaky regulation. This can be observed at time points before the actual induction. In some experiments the cells have already been active (up to 40% in one experiment with 200 nM, Fig. 4.15**C**). The plot is extremely influenced by single experiments at the beginning as not all do have the same length. At a concentration of 500 nM, the highest number of active cells is observed. A big difference between intermediate (Fig. 4.15**B**) and high concentrations (100-500 nM) is a cumulative activation of cells. While cells stimulated with e.g. 20 nM show a saturation at ~25 % active cells after ~80 min, the fraction of active cells at 500 nM does not reach a plateau and increases up to ~38 % after 160 min. A possible explanation for this finding is an increase of the on-time of the gene with longer induction. More β-estradiol leads to faster relocation of GEV, and higher levels of TF in the nucleus favor more transcription. Cells might be caught in a state where re-initiation is increased and the chances of observing active cells within the 20 minute window increase over the course of the experiment. The average responder rate follows the concentration of β-estradiol (Fig. 4.15**D**).

In contrast to the number of responders, it is surprising to see no clear trend in the accumulated intensity with increasing inducer concentration (Fig. 4.17**B** and Fig. 4.18). To directly see, if the maximum of gene activation is already reached at the lowest concentration of 1 nM, more experiments with lower dosage are performed, although the original paper showed no significant activity below this concentration. Another possible explanation for this finding might be the low expression of GEV inside the cells, which reacts

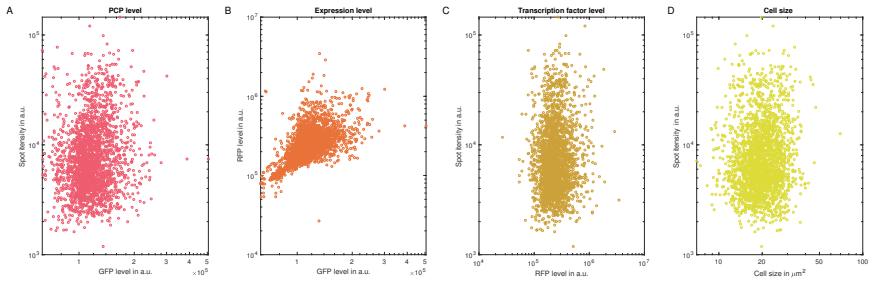

Figure 4.16: Dotplot of the integrated trace intensity in relation to other parameters. **A** The calculated GFP value for each cell is size corrected and rather uniform (one log difference). **B** Despite a size correction, a strong correlation between GEV and PCP expression level is observable for low expression levels. The amount of GEV in the cell (**C**) and its size (**D**) do not have an influence on the total spot intensity. These data are based on 3500 cells.

to the stimulus. The used plasmid has a weak *STE5* promoter and a low copy *ARS/CEN* ORI. All free GEV could be already be bond to β-estradiol at 1 nM, so an increase would not result in higher transcription; at 1 nM each cell has approximately 70 molecules of β-estradiol without any accumulation. To test this hypothesis, I measured the total mCherry concentration and plotted it against the accumulated trace intensity (Fig. 4.16C). The accumulated spot intensity is not correlated with the expression level of GEV-mCherry. Cells that express GEV at high levels do not show higher accumulated trace intensities. Therefore, GEV is not limiting the transcriptional activity. The background activity (Fig. 4.15, before time 0) seems highly variable and the fraction of responders increases over the time of the experiment. Control experiments with a strain lacking GEV showed little activity when exposed to β-estradiol, and never reached a comparable fraction of responders (Fig. 4.15). Also, control experiments without any β-estradiol show insignificant activity (less than 3% of cells show a spot).

The statistics on the spot duration and its corresponding sum of intensity add to the findings. The variations in spot duration or gaps between spots are not obviously linked to the stimulation. A slight effect on the integrated intensity could be found for the experiments with more than 100 nM (Fig. 4.18). The total GFP level in these three groups (<10 nM (red), <100 nM (orange, dotted) and >100 nM (green, dashed)) has minor changes that might be based on the natural heterogeneity. The integrated spot intensity, however, seems to differ for the last group. The mean of the high induced cells is shifted upwards, in the histogram visible as reduced abundance of low spot intensities. The bulge at $log_{10}(spot\ intensity) = 4.5$ increases with more β-estradiol. Besides the fraction of responders this parameter can be dose dependent. Although the presented data are not conclusive enough to provide a quantitative analysis of changing transcription parameters during induction, the data are based on statistically significant numbers. The reduced responsiveness of the system can have multiple origins that are not linked directly to the

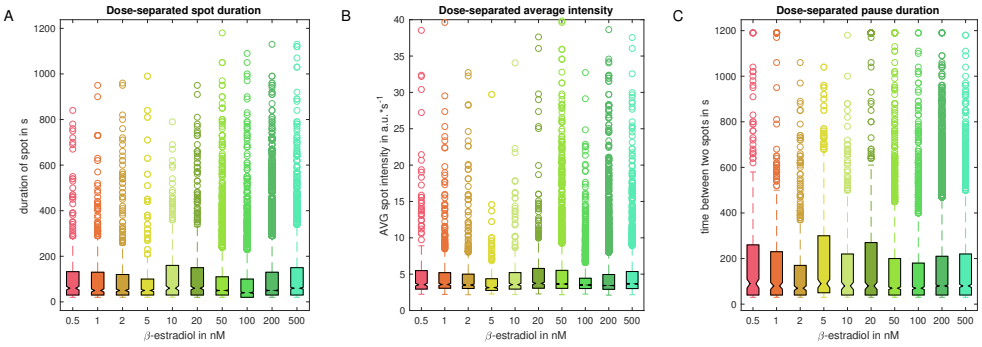

Figure 4.17: Spot duration and intensity separated by dosage. **A** The calculated spot duration does not change with the concentration of β-estradiol. Cells that show a spot over 90% of imaging time are rare events (0.9%). **B** Average spot intensity shown for each dose. No correlation between the dosage and the spot intensity can be found. **C** Average pause duration between two spots per induction level is at ...± s. The number of events for each dose varies: 369, 487, 547, 470, 402, 607, 2335, 1627, 5164 and 2187 spot appearances have been observed from 6000 cells.

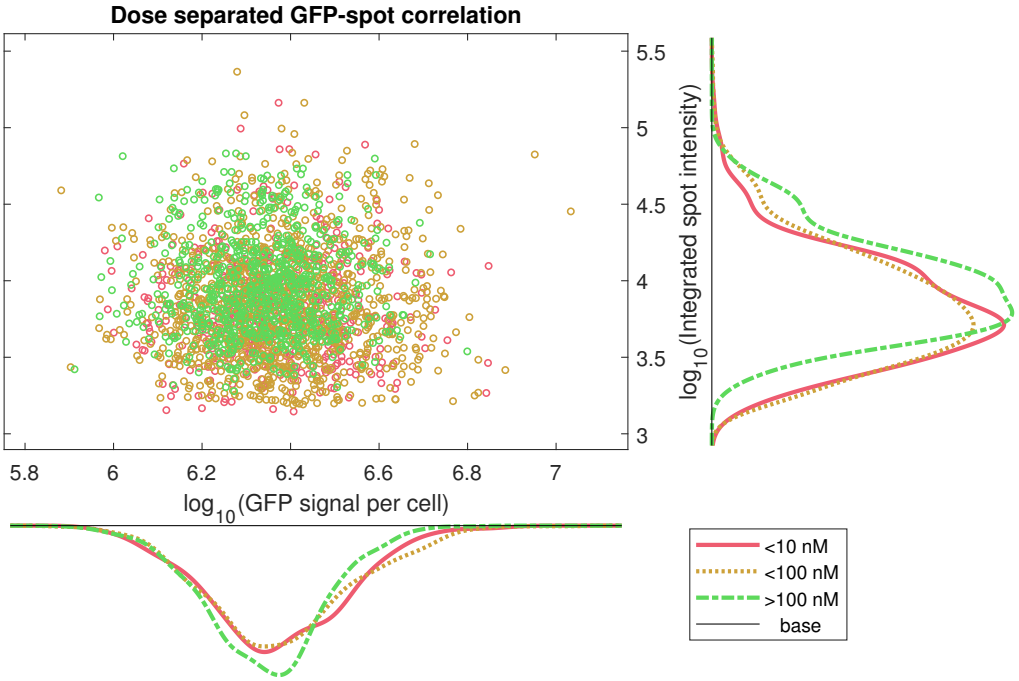

Figure 4.18: Dotplot of the integrated trace intensity in relation to other parameters. **A** The calculated GFP value for each cell is size corrected and rather uniform (one log difference). A slight correlation between the GFP level and trance intensity is observable for cells induced with more than 100 nM. Each set contains at least 800 cells.

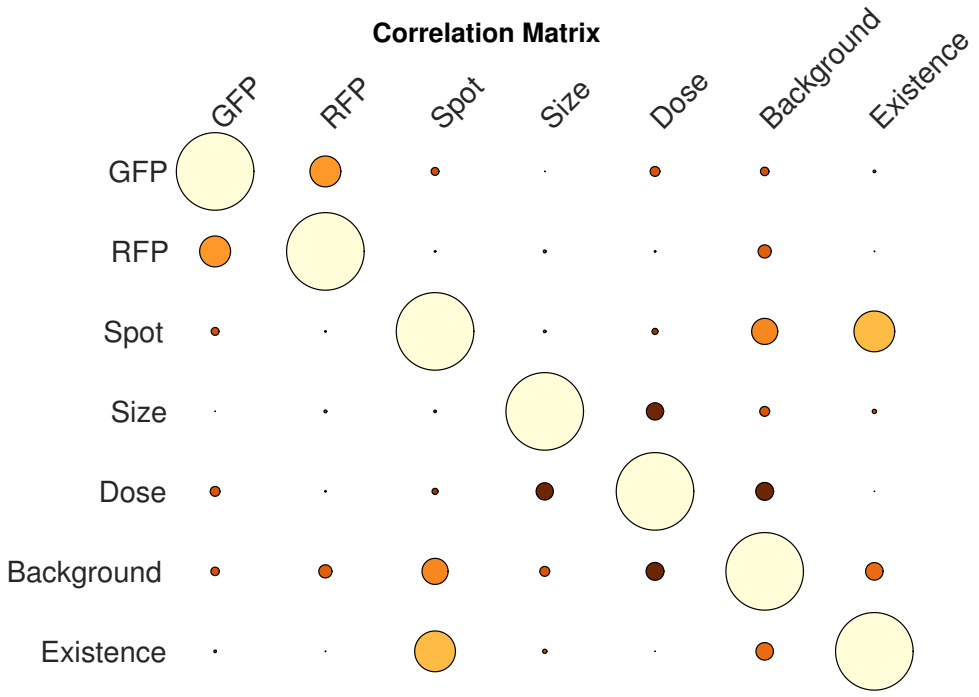

Figure 4.19: Correlation matrix from variables. The GFP and RFP values are corrected for size, as a dependency is here expected. The correlation between the dose and RFP is based on the re-location of the transcription factor. As the maximum projection is used to estimate the RFP signal in the cell, the brighter nucleus does bias the data.

transcription factor and its abundance. One of them might be encrypted in the DNA and the nucleosomes.

The result of the correlation matrix between the parameters was to be expected (Fig. 4.19). The positive correlation between the existence of the spot and its intensity is a logical consequence. Also, the correlation between the background signal and spot intensity seems trivial. Cells with brighter nuclei have as consequence a higher background and also higher intensity of spots. Besides this, the spot does not correlate with other physical quantities like GFP or RFP values.

4.4 DNA properties and their influence on transcription elongation

Although the yeast genome is not tightly packed as in humans, the speed of elongation within a gene can still be influenced. A classical example is the packaging protein complex of a nucleosome. The DNA is wrapped around the 8 histone subunits of a single nucleosome and interacts with its surface. Some mechanisms have been proposed on how

the RNAP can get access to the DNA template [104]. Veloso *et al.* estimated in 2014 the influence of some template properties on the elongation rate with genome-wide RNA sequenzing [98]. Among others, the GC content of the template DNA in general influences the speed of RNAP negatively, but Veloso was not able to provide a quantification of this finding. To fill this gap of knowledge, I decided to directly observe the influence on transcriptional speed. In section 8.4 the construction of a DNA part is described that has overlapping regions for integration into yeast. It has several repeats of PSL in the 5' UTR followed by a defined GC content sequence, with several repeats of MSL adjacent (Fig. 4.21). The combination of both stem loop sequences on one gene offers the possibility to operate a dual color construct (PCP-2xGFP and MCP-mScarlet) to directly measure the elongation speed, which is directly linked to the time delay between the emerging signals in green and red. A Matlab script is used to generate an homogeneous GC content over 900 bases in length (Fig. 4.20A). The corresponding melting temperatures of the moving transcription bubble are shown in Fig. 4.20B for different contents; the native Gal10 gene (orange), the GC66 construct (green) and a AT66 control sequence (lilac). While the synthetic sequences GC66 and AT66 have minor fluctuations, the native gene body of Gal10 highly fluctuates in the GC content and in the estimated melting temperature for the transcription bubble. The last part of the integration cassette is the HIS3 auxotrophy marker for selection purposes. The overlapping regions are selected based on gRNA finder (www.e-crisp.org) [34] for the use of CRISPR/Cas9. The appropriate gRNA and Cas9 plasmids are constructed with Gibson assembly. In total 3 different gRNA are constructed and tested (Fig. 4.22). Despite the wide range of applications in which Cas9 is successfully used for integration [41], only one transformation with one gRNA shows colonies after four days of incubation. Sequencing results show the existence of the insert DNA but no colony PCR from the genome into the insert yielded positive results. A microscopic examination of the strain after introducing PCP-GFP, MCP-mScarlet and GEV shows the expression of GFP and mScarlet. Despite induction with various β-estradiol concentrations (1, 10 and 100 nM) no spot in green nor in red is detected.

4.5 Inference on single cell data

The methods to estimate the elongation speed in literature for data like the presented are mostly utilizing auto-correlation (see Section 2.3.1; [55]). Although used widely, some problems arise: auto-correlation generates optimal results, when applied to long time series in an equilibrium state. The recordings presented here are not suitable to be analyzed by auto-correlation with their 20 min duration under inducing conditions. Nevertheless, computational approaches can be used to extract meaningful knowledge out of the data. Bayesian inference on a continuous-time Markov chain is able to estimate

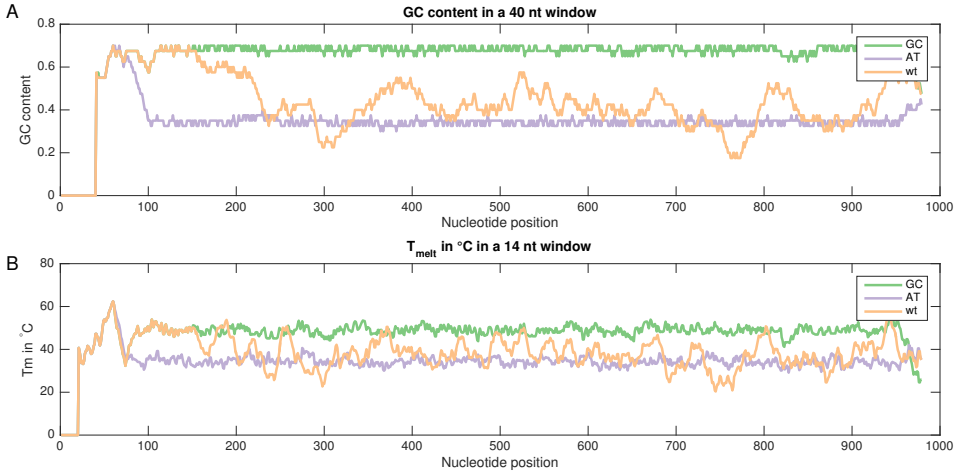

Figure 4.20: Properties of synthetic DNA. **A** Relative GC content in a 40 nt window. The calculated GC66 sequence (green) has a consistent GC content at 66% on average. The native Gal10 gene (orange) shows fluctuations, while the AT66 control sequence (lilac) is at 33% GC. **B** Estimated melting temperature with after SantaLucia (1998) [77] in a 14 nt window representing the transcription bubble.

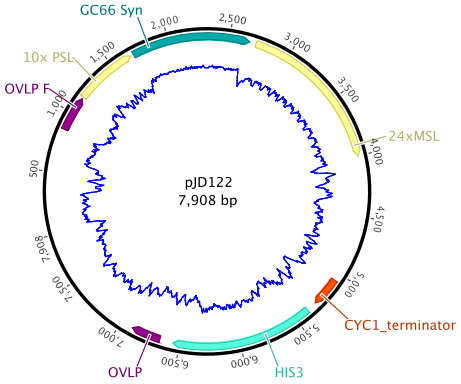

Figure 4.21: Plasmid map of one of the final construct. The overlapping regions correspond to GAL10.


```
818 CCTTCATTTTCATTGTAGG 800
1275 GCGCCTATATAAGCACTATCAGG 1252
2097 GGAAAATCTGTAGACAATCTTGG 2075
```

Figure 4.22: Genome map with annotated gRNAs. The already modified Gal10 locus with 14 PSL has 3 available gRNA targets (at position 800, 1252 and 2097) in its sequence. gRNA is searched with E-CRISP [34]. Below are the sequences of the 3 targets in 5' to 3' orientation. Decreasing number indicate the complementary DNA strand.

the number of polymerases visible in the data at each time point. The inference is based on the theorem:

$$P(Position\ of\ polymerases|Data) = \frac{P(Data|PosP) * P(PosP)}{P(Data)}$$

The probability of the data itself is often unknown and additional methods have to be utilized to handle this issue. The elongation model assumes some simplifications: the locus 14x PSL-Gal10 is segmented into 48 sites, each site as long as one stem loop for PCP and the number of polymerases allowed at the same time is limited to four. This reduced the possible state space about several orders of magnitude. The movement of a polymerase follows a continuous-time Markov chain (CTMC); meaning the jumping time to the next site is exponentially distributed and captured in the rate c_E. The same applies to the initiation (c_I) and termination (c_T). A mathematical description of the process can be found in table 4.1:

Elongation model	
$X(t)$	position vector for occupied sites
c_I	initiation rate
c_E	elongation rate
c_T	termination rate
Observation model	
$N(t) = \alpha^T X(t)$	produced stem loops
$I_{spot}(t) = \gamma * N(t)$	intensity calibration
$I_{obs} = b_0 + e^{-\lambda t}(b_1 + I_{spot})$	spot intensity with ground noise and photobleaching
$Y(t) = I_{obs} * e^{\sigma\epsilon}$	observed intensity with ϵ following Gauss(0,1)

Table 4.1: Observation model of transcription at this locus

The inference algorithm, developed by Leo Bronstein and Christian Wildner, generates possible trajectories from the smoothing distribution. By simulating several thousand of

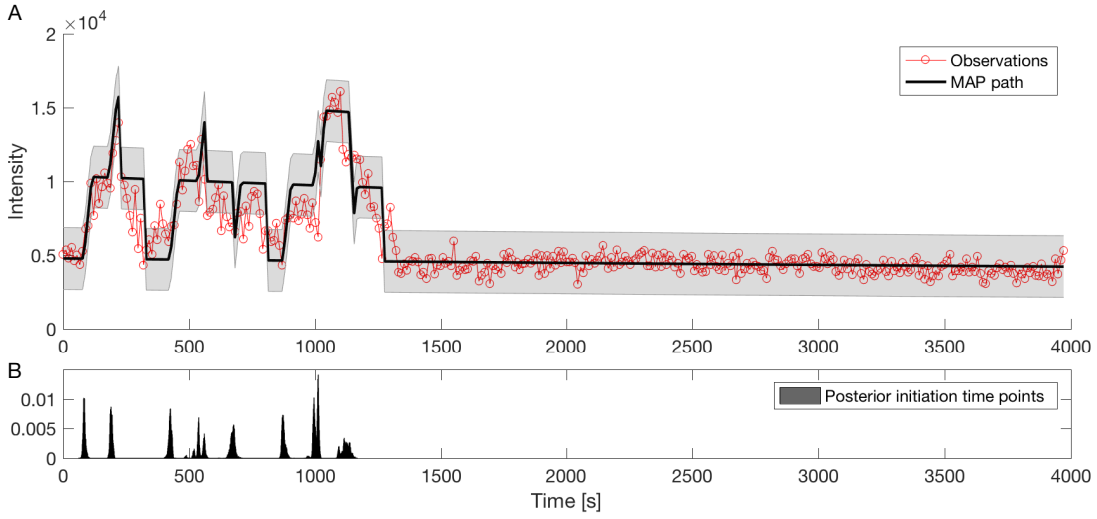

Figure 4.23: Exemplary inference result. **A** Based on the observations (red) a likely path of produced stem loops (black). **B** Starting times for each of the estimated 8 polymerases

those trajectories, the most likely can be estimated together with the variance. Several parameters have to be guessed and estimated roughly beforehand, like the bleaching rate, scaling factor and background signal. The model excludes a take over of polymerases and introduces a minimum distance. The intensities from the tracking algorithm ($Y(t)$) are calibrated to a corresponding number of GFP-tagged stem loops ($N(t)$), which is then translated into a number of polymerases ($X(t)$). The start times for each polymerase are given (Fig. 4.23**B**) and correlate strongly with the first increase in signal. The dwell time of the spot can be departed into two major processes: elongation and termination. The termination can take some time after the actual gene of interest is already transcribed (Fig. 1.3). One output of the model is the distribution of likely elongation rates. For this example, an elongation rate with 30 ± 7 nt/s explains best the observed data (Fig. 4.24**A**). The termination time is estimated with ~35 s (Fig. 4.24**B**). Combining these numbers, one polymerase spends about 135 s on the gene, which is in agreement with the results concerning the length of spot existence (Fig. 4.13 and 4.17).

4.6 Real time imaging and evaluation of the neo-tc gate

In cooperation with Christopher Schneider from AG Süss, yeast with a logic gate are measured with the microfluidic chip [78]. As the system is inducible as well as repressable, several experiments with different combinations of inducer and repressor are performed with the design for four independent measurements. Although the usage of the four independent chip requires more preparation time and cells reside until all chambers are loaded and the experiment can start, the advantage of this chip become evident. A sin-

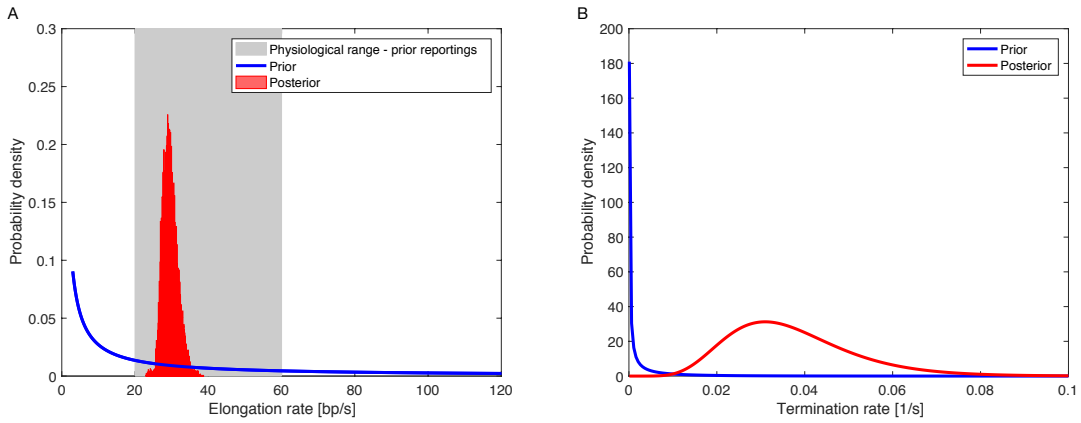

Figure 4.24: Posterior distribution of selected parameters. **A** Estimated elongation rate. The blue curve shows the prior probability density, the red the posterior. The physiological regime is highlighted in gray. The data are best explained by a elongation rate of 30 nt/s and a termination rate of 0.033 events per second **B** Probability density for the termination rate

gle experiment requires approx. 1-2 h of preparation and is run overnight. This chip, in combination with an appropriate imaging interval, parallels the recording, allowing multiple sets of experiments to be finished within a week. The logic gate consists of the GEV-mCherry transcription factor as inducer, a Gal1 promoter, neomycin aptamer followed by the tetracycline aptamer, both upstream of a yeast optimized GFP variant. Upon induction with β-estradiol, cells transcribe the RNA, giving access to the translational regulatory aptamers. Both aptamers form strong secondary structures after binding their ligand and block the ribosome from accessing the GFP sequence. The effective dosage for the inhibition varies between both elements, requiring a higher concentration of tetracycline to achieve the same switching factor in bulk measurements. In all experiments with this strain, 10 nM β-estradiol is used to induce transcription. For the repression of the logic gate, 100 μM neomycin (neo) or 250 μM tetracycline (tc) is used and added after 2 h after induction. The last set of experiments is conducted with a combination of the two repressors. In Fig. 4.25**A** flow cytometry data are shown over the time course of 7 h. The three populations (induced (green), repressed (gray) and uninduced (black)) are clearly distinguishable by their signal intensity. This holds true for the microfluidic experiment. After a run-time of 6 hours, the average intensity per cell separates significantly from fully induced to repressed. At the chosen concentrations, no significant difference can be found between neo and tc based repression. Both molecules alone repress the translation sufficiently. The cells remain vital during the time curse (Fig. 4.25**B**) and clogging of the device is observed after 7 h (data not shown). Interestingly, the peak of fluorescence for the repressed experiments is reached after ~190 min after induction and 70 min later then repression; this trend is found in the cytometry data, too. Another trend in the data

suggest a decay of the EGFP after 200 min since induction. This is most prominent with both inducers, but can be found with either of the two other experiments.

4.7 Expression dynamics of light inducible potassium channels

In another example, the chip is used to follow light induced gene expression over time. I received a yeast strain from Sebastian Höler from AG Thiel with a light inducible potassium ion channel N-terminal tagged with EGFP [39]. The main questions which can be addressed by microscopy are: What are the dynamics of the gene expression on the single cell level and where does the channel localize inside the cell? (Fig. 4.26). The transcription factor relocates at the exposure with light of 495 nm into the nucleus and starts transcription [48, 39]. In the first round of experiments, the same strain is imaged with different light conditions (Fig. 4.27). While imaging with an interval of 15 min in the GFP channel is not enough to trigger transcription, the interval of 5 min triggers the gene expression in some cells. For continuous illumination a LED lamp is used, which is switched on for 3 min between two frames in the first 30 min of the experiment. A direct comparison of the illumination shows no differences between the discontinuous and continuous case for the expression dynamic (Fig. 4.27D). After 50 min the exposure during the recording seems to be not sufficient to activate more transcription factor. Even for the continuous illumination during the first 30 min, a decrease in signal is observed after 120 min. The experiments are not performed on the same day under the same conditions, as a partially illumination of the chip is not feasible.

In the next experiment series, the four parallel chip design could be used. Three different yeast strains are measured under the same conditions with continuous illumination for the whole duration of the experiment (Fig. 4.28). To compare the outcome of the experiment, the previously mentioned strain is again measured. The second strain has a the same potassium channel, but a codon optimized variant of GFP for yeast. The genotype of the last strain consists of a codon optimized version of the potassium channel tagged with the codon optimized GFP. Although the GFP has not changed in amino acid sequence, there are differences in the expression levels visible after 120 min. Surprisingly, the optimized versions of the protein show less signal than the old one. Since both variants of the channel exhibit the same dynamic, the signal difference must come from the changes in the GFP. Addressing the localization of the channel, the hypothesis is an incorporation into the membrane of the endoplasmatic reticulum (ER) [80]. This can be confirmed, based on comparison with literature images [18]. But, a quantification of the tagged channels inside the membrane is not practicable in an automated fashion, because of the lack of a secure method to segment the ER. A selection by hand is only feasible for a small set of cells, which will not satisfy scientific acceptable statistics to formulate a statement. The ER membrane can only be imaged as an blurred and fussy area (Fig. 4.26), which is several times broader then its actual dimension (250nm compared to 7nm). For a quantification of the localization, the use of a less background corrupted technique like confocal

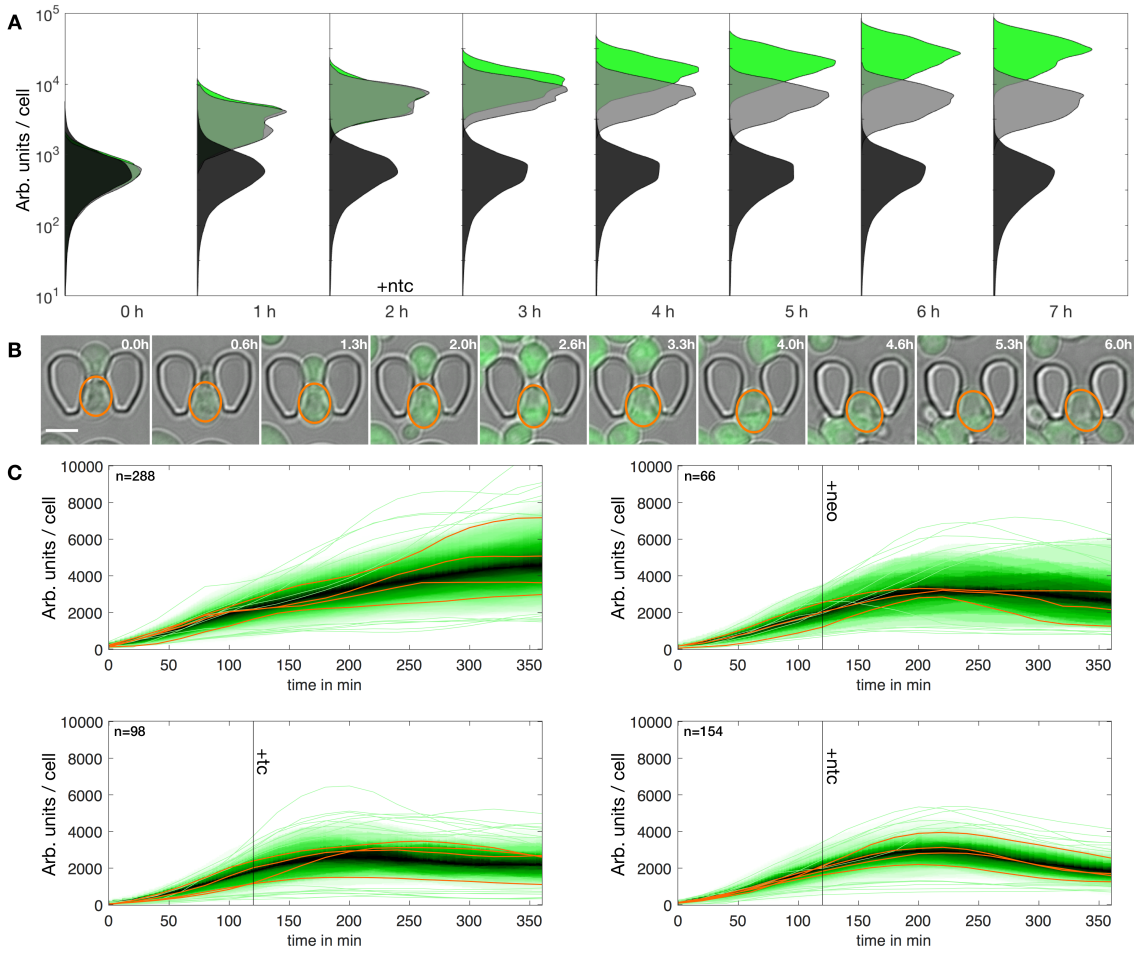

Figure 4.25: Qualitative population dynamics of the NeoTc gate. **A** Histogram of time-lapse flow cytometry data from the +2 h experiment. Uninduced cells (black) show a broader distribution than cells induced at 0 h (green) and cells induced at 0 h and repressed at +2 h (gray). **B** The picture series shows a representative cell trapped on the microfluidic chip and recorded over time by fluorescence microscopy. Scale bar 5 µm. **C** Density plots of segmented single-cell traces. Random examples for individual traces are highlighted in orange. Induction with 10 nM β-estradiol, repression with 100 μM neomycin and/or 250 μM tetracycline. Color becomes darker for more cell traces approaching the median.

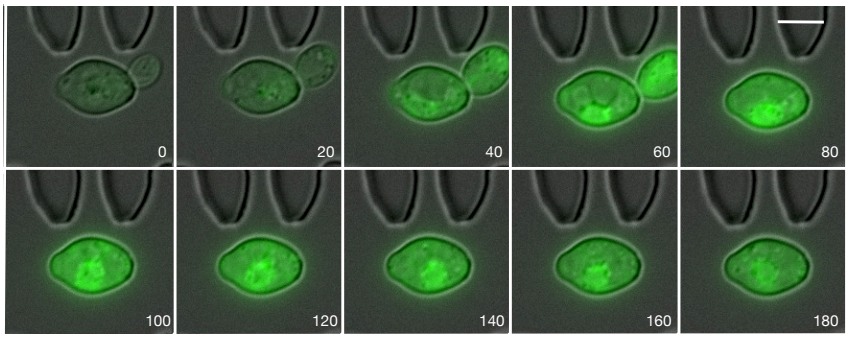

Figure 4.26: Example images for the light induced gene expression of a potassium channel. The channel gets integrated into the endoplasmatic reticulum membrane around the nucleus. Scale bar equals 5 μm

laser scanning or spinning disk microscopy is desirable in combination with a membrane marker with a different wavelength. Such a dye could be found in FM4-64. It belongs to the lipophilic styrol dyes and gains emission, when located into a membrane. Cells are incubated for 30 min with a 20 μM solution of FM 4-64, before loading into the chip. The media contains 4 μM to counterbalance the depletion of dye from the membrane. Based on the observations from several hundred cells, FM4-64 is mostly incorporated into the membrane of the vacuole (Fig. 4.30, [28]). The overlay with the GFP signal from the ion channel eliminates FM4-64 as suitable stain for the ER membrane. Alternatives in the form of genetically encoded marker, like TOK1-mCherry, have to be reconsidered for further experiments.

In conclusion, the dynamic of light induced expression of KCV-EGFP is for the different constructs comparable, but the total level of protein reaches its maximum with the non codon optimized sequence. The four chamber chip is most suitable to receive comparing data sets with different strains, although the light sensitivity of the cells limits the time that can be spend in the preparation.

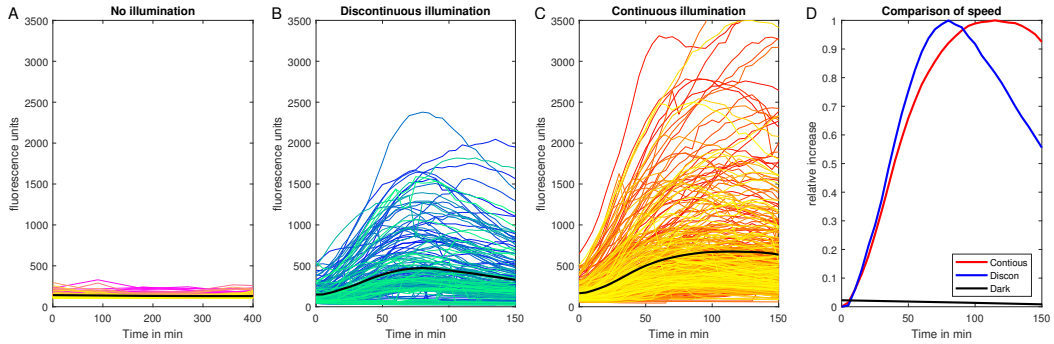

Figure 4.27: Experiment with strain PLY240. Display of single cell traces from 3 different experiments. Left: Control experiment with no additional light excitation, images taken every 15 min. n=460 **B** Images in the GFP channel are taken every 5 min. This is enough light to trigger gene expression. n=389 **C** Illumination of cells for 3 minutes between microscopic image recordings for the first 30 min of the experiment. n=512 **D** Comparison of the dynamic of the GFP expression. For better visualization the mean of the single cell traces is plotted

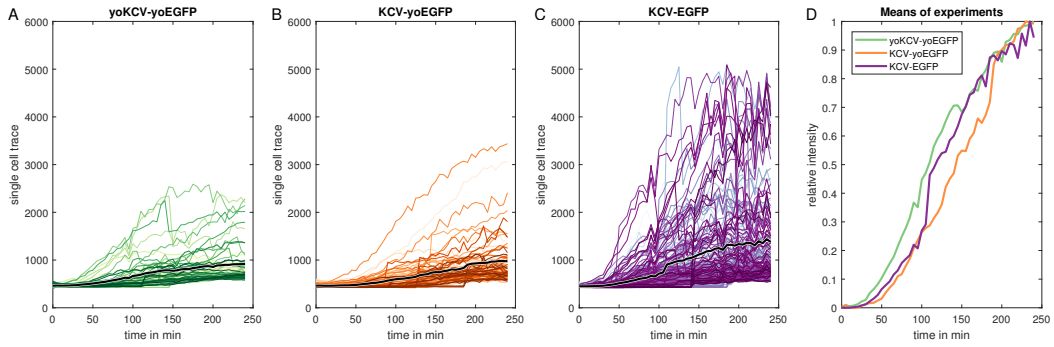

Figure 4.28: Result of parallel experiment with 3 different strains. The light induction of gene expression is continuously on. Images are taken in a 10 min interval. **A** **y**east**o**ptimized KCV-yoEGFP, n=176 **B** KCV-yoEGFP, n=219 **C** KCV-EGFP as reference, n=481 **D** Comparison of relative intensity increase of the average cell per experiment.

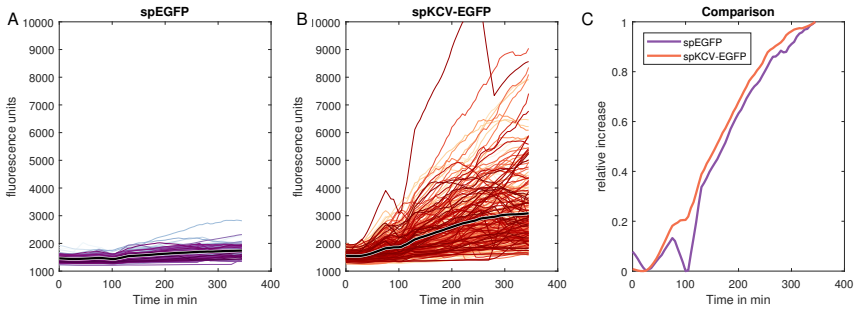

Figure 4.29: Result of parallel experiment with 2 strains. **A** The strain with singleplasmid EGFP, contains the light sensitive TF and an EGFP as target gene. n=161 **B** spKCV-EGFP expressed after induction with light. n=383 **C** The individual dynamic of the plasmid is comparable, the drop of intensity results from out of focus time points. What is surprising though, is the total intensity after induction. Cells with KCV-EGFP express more functional EGFP than cells without the channel.

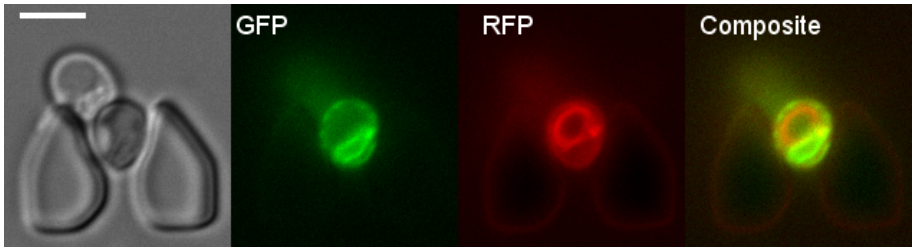

Figure 4.30: Experiment with stain FM4-64. Example cell with KCV-EGFP after 1h of light induction. The image montage shows the localization of the ion channel in the ER membrane and the dye FM4-64 in the membrane of the vacuole. The composite images shows no co-localization of the signals. Scale bar equals 5 μm.

5 Bringing the results into context

In the following chapter, I will discuss the results and compare them to published data on similar topics. Possible sources for the encountered errors will be exemplified and solutions are proposed. In addition, future steps are highlighted that could be investigated and directs of long term studies are given.

5.1 Design and fabrication challenges of the chip

The fabrication of two wafers for microfluidic chips was successful with the first redesign. The usable chips produced from those wafers drastically decreased PDMS waste and lowered the time needed to produce a suitable cell chamber. As the design provides multiple chambers on each chip, the total production costs could be decreased by a factor of three (three cell chambers were the lowest number observed of usable cell traps on a chip with the five rescaled former designs). Nevertheless, some problems of the fabrication process could not be resolved so far. For instance, a poor structure was found, when the trap had either not the correct shape or was not bond to the glass at the end of PDMS chip fabrication. Unfortunately, incorrect exposure as well as resist development could be the source for both issues. On the one hand, an ineffective wash-out of the traps during development can leave a layer of resists at the bottom of the trap, reducing its depth and preventing the binding onto the glass cover slip. In case the developer was not removed efficiently from the wafer in time, the shape of the structures were affected. The wafer was placed on an orbital shaker to simulate a flow of PGMA developer and increase development efficiency. As an alternative, ultra-sonic bath with PGMA was tested, but no change in quality was found.

On the other hand, increased gap size during UV exposure decreased the resolution and the structures of the traps were not fully protected from UV light. This could have led to a cross-linked layer of resist at the surface of the wafer from the overlapping light paths. The manuals of the resists recommended a developing time of 2-3 min for the desired height of 7 μm for the cell chamber and the best results were obtained after 45-60 s. To bypass the fabrication issue, a direct laser writer at another facility was tested with the trap design. Instead of a complete exposure of the coated wafer, only the actual structures were illuminated. This technique offered the possibility to form structures in 3D directly within a substrate [82]. With advancing laser techniques, like stimulated emission depletion (STED) [35], the writing resolution could be increased up to the range of few hundreds nm [27]. The drawback of this approach was the time needed for a full wafer to be written. Depending on the aspect ratio and area of the desired structures, a direct writer consumed too much time in exposing and cross-linking. For this reason, all bigger structures (> 30 μm) were fabricated with the standard photolithographic methods.

Only the biggest ovals tested were fabricated with sufficient quality. But the spacing between traps lowered the amount of data yielded by this chip. Some differences between the layouts could still be found. The L-shaped traps had the best catching rate from all designs. The cell is buried deeply in the PDMS structure and an unwanted washout was barely possible. But this applied also to cell aggregates. If one cell of an aggregate was caught, the flow rate did not suffice to remove the rest of the cluster from the chip. Depending on the planned length of the experiment, different designs were used. Smaller traps, like the Alcatras design were densely placed and allow the recording of a lot of cells for a short period of time. As those traps lost cells more easily, it was not the best option for overnight experiments. The duration of the majority of the experiments presented here (up to 8 h) did not favor any of the designs.

All tested trap designs were capable of capturing and holding cells for several hours, with 20 h as the longest conducted experiment. Chips from the second generation and with yCS cells were clogged after 7 hours in the shown experiments (Fig. 4.25). Compared to the Alcatras [15] or HYAA [60] chip, where the chips could be maintained up to 40 h repetitively, the chip presented here was limited in its use. Nevertheless, in the 7 hours in which yCS was recorded, I obtained enough data to confirm and support the results from the flow cytometry experiments. The limitations in chip operation could be linked to the presence of flocculation genes and fast growth rate for this strain. Introducing deletions on the flocculation relevant genes could improve the conditions for the loading process, as more single cells or doublets would have been flushed in. Alternatives could be found in methods that open cell clusters. The yeast strain yJD was based on a FLO1 mutated strain and frequently showed smaller aggregates of cells, but missed to produce single cells as majority, too. Trapping of yJD and maintaining experiments on the order of 4-5 h was successfully done in a large number of experiments. The flow rate of 4 μL/min was in agreement with literature [15, 43] and had no negative effect on the growth rate for the used strains. The strains from the light-inducible potassium channel project had the highest ratio of single cells. Based on my observations, I could conclude that the culture conditions and growth rate of *S. cerevisiae* had more impact on the loading efficiency and clogging than the actual cell trap design.

New features of the chip presented here were the parallel designs and the combination of cell traps together with on-chip-valves for pulse width modulation. No other device to my knowledge combined the fast switching of media with a local arrangement of cell chambers, for either 4 different strains and the same media conditions or 4 different media conditions with up to 4 strains. The number of caught and maintained cells was with 500-700 per chamber in mid-range (Table 1 in [15]) of established devices. The novelty of this chip was best demonstrated in the projects with the logic gate and the light inducible system [78]. No other device from literate could have been used to generate data at this time resolution.

5.2 Transcription dynamics in response to an external stimulus

The labeling of mRNA based on coat proteins is an established method to visualize transcription in living cells [8]. Besides MCP and PCP, other binding proteins can be used to tackle the challenge [96]. The PP7 coat protein is mainly favored for its binding specificity and properties. The detection of RNA with the PCP system has drawbacks, the high back ground fluorescence being a major one. In order to bind newly synthesized stem loops, PCP-GFP has to be present in the nucleus. The expression level of PCP-GFP dictates the detection limit of the RNA. In case of a strong promoter for PCP, the nucleus will be bright and the mRNA may be detected only after the completion of the stem loop region or never overcome the background intensity. In the opposite extreme, where one has negligible amounts of PCP, the protein could become depleted and newly RNAs cannot be tagged with functional PCP-GFP. A partial solution for this problem is offered in [96, 102]. In both studies, the authors utilized split GFP fused to two different RNA binding proteins. The corresponding stem loops are in close proximity and allow the split GFP to fold and gain function. The drawbacks of this solution lays in the temporal resolution as two protein have to bind to each other in the correct ratio to result in a fluorescent read out. When a new stem loop is transcribed by RNAP it should directly be bond and therefore be detectable. Split GFP however needs time to completely become measurable. Even with the fastest split GFP variant the order of minutes is never undercut [102]. Those studies solve the issue of a high background, but are unacceptable for real time transcription dynamics. Another method is introduced by the Singer lab [103]. Here, the authors uses sequentially PSL and MSL to co-locate PCP-CFP and MCP-YFP for FRET measurements. The time resolution with this technique does not differ from the single PCP method. The total signal is lower, as for a single coat protein less stem loops are present, but for the FRET acceptor, the background intensity is minimized and the signal-to-noise ratio is improved.

An alternative to RNA binding proteins are fluorophore binding RNAs. Instead of bringing the signal to the mRNA, the mRNA is generating the signal itself. The main advantage of this method is a low background, the disadvantage however is the accessibility of the ligand. The variety of ligand binding RNA (aptamers) with fluorescent properties increased over the last years [70, 23, 88], raging from blue over green to red. Specifically targeting the nucleus inside a yeast cell with such a ligand is up to now not possible. In such an experiment the ligand would be also present in the cytosol and label already transcribed mRNA. An additional segmentation of the nucleus might be required to exclude all RNA and only take spots into account with changing intensities. As the signal of the aptamer emerges directly on the binding, the signal-to-noise of each spot would be higher compared to the PCP system. From all those methods, measurements based on the NLS-PCP-GFP provide the fastest reaction time for mainly new mRNAs in the nucleus. The signal could be improved with a FRET acceptor to lower the background. Split variants

have less time resolution and aptamers lack the specificity for newly transcribed RNA. That is why the PCP system is preferred in this study.

The construction of an inducible system to visualize transcription was successful. The cells showed a reaction to the external stimulus: at higher induction levels, more cells showed a spot. Following the observations from McIsaac *et al.*, my construct of Gal10p-14xPSL-Gal10 + GEV + PCP-GFP-GFP met the expectation and recapture bulk measurements of total GFP expression driven by this system. Interestingly, all cells were supposed to show active transcription (Fig. 4**A** in [64]). From other experiments [78] it is known, that the whole cell population shows increased activity. This leads to the insight that every single cell produces more protein with higher amount of β-estradiol. One possible explanation could be the low level of accessible GEV inside the cell. The transcriptional activity could be limited by GEV, as the fusion protein is already bond at minor β-estradiol concentrations.

Since there was no direct correlation between the level of GEV-mCherry expression and the total transcriptional activity (Fig. 4.19), one could conclude that the transcription of this locus was not modulated by the amount of β-estradiol the cells receive, but rather transcription switched to the same on-state, when β-estradiol was present. This finding was partially in contrast to the results of McIsaac [64], who built and characterized the GEV system. Therein, the authors describe how the system could be tuned to reach a certain protein level using β-estradiol. The switching range from background activity to fully induced never exceeded 10 fold for the population mean. The differences in total amount of protein after 12 h could be explained by the accumulation of GFP proteins over time. The time window to measure GFP is exponentially bigger than the time period of active transcription. In my experiments only a sub-population showed transcriptional activity during the 20 minute interval at each position. It could be possible, that the non-responders were actually responding, but not during the recording interval. They might have been active at some point in a 4 h movie and therefore also contribute to the measured protein level, but failed to be recognized in my assay. The imaging window could be to short to recapture the results from McIsaac. Due to the observed bleaching, longer movies were not feasible. Lower imaging frequencies could have been used as countermeasure, but would have lost the precision on the RNA synthesis dynamic.

A model, which fits better to my observation was introduced in [90]. Here, the authors investigate the mRNA dynamic with a HIV1 promoter driven RNA production in human cells. Using the MS2 coat protein and stem loops to visualize newly synthesized RNA produced by a convoy of polymerases on the gene. Mediator bound to enhancer and TATA-box in their wild-type strain enabled a fast re-initiation of RNAP on the gene, before the connection was lost. Mutations in the TATA-box led to longer time intervals of no activity, supporting the idea of a additional inactive state for the gene. Those mutations altered the binding affinity of Mediator to the DNA, resulting in less binding events, and therefore less initiation. The mutated case is better comparable to the system I used,

since it is not a constitutively transcribed gene. The promoter of a housekeeping gene (POLR2A) is analyzed too and shows similar results.

5.3 Unfinished project about the influence of GC content on transcription dynamics

The cloning of a plasmid with MCP stem loops, PCP stem loops and a homologous GC rich sequence was successful. The designed plasmids with the overlapping region for the three gRNA targets were confirmed via sequencing multiple times during cloning. But, the integration of the construct into two different yeast strains was unsuccessful with three different protocols. Possible explanations for the failed transformations and integrations include mutations or mismatches in the genome, no expression of the gRNA or Cas9, not sufficient amount of DNA, inefficient transformation or the lack of enough competent cells.

Mismatches or/and mutations can be checked with the sequencing of the genomic DNA from the competent cells. Two approaches are used to control the DNA: (I) PCR on the genome to produce enough DNA fragments for the sequencing run and (II) extract genomic DNA from cells. The PCR on the genome is done (primers in 8.4), but failed several times. The extraction of the genomic DNA was successful. The sequencing trials have been too short to gain any insights. Only 100 bp are covered in the runs. Therefore, the genome itself might be a source of error, which could not be excluded. To check the correct expression of the required protein Cas9 an SDS-PAGE with a reference strain could have been made. As the plasmid was successfully used in another project, the correct function of Cas9 was proven. The same holds true for the plasmid containing the gRNA. Although the individual gRNA sequence was altered, the plasmid with promoter and scaffold RNA seem to work [78]. A validation could have been performed with an RNA purification and reverse transcription and amplification via PCR (RT-PCR). The next possible source of error could have been the amount of DNA which was used in the transformation. The authors of CasEMBLR [41] used for their integration 4 picomoles of each of the five parts with 50 bp overlap. My design had just one fragment that was generated by a digest with EcoRI and EcoRV of pJD122. For all the transformations 4 picomoles of plasmid (~20 μg) are digested prior use. The digest was purified and concentrated with kits from NEB to remove the enzymes and minimize the volume of the DNA solution. The purification on a column could have led to loss of DNA, especially when the capacity of the column is reached. The manufacturer of the columns had estimated a limit of 25 μg DNA. In addition, the digest could be incomplete and lower the available DNA for the integration. Test gels of the digests are made and no incomplete fragments are found after 1 h of digest (not shown). To check whether the used cells were competent enough, different transformation protocols were tested. Hence, the transformation with all necessary plasmids was separated into multiple steps, to lower the burden on the cells at each transformation. Cells that already contain the PCP-GFP-GFP, MCP-mScarlet and

GEV were used for the genome editing. This could lower the efficiency of the transformations as the cells have already unused proteins expressed. The fact that both, PCP and MCP, were transcribed from a low copy plasmid with a weak promoter was limiting the negative effects of the plasmids being present. The conduction of the genome editing as first step might increase the efficiency, because the additional auxotrophic marker can reduce the growth rate of yeast [47]. For this reason, a different strain (yRS453) without PCP and GEV was simultaneously used for the transformations. Both strategies showed no functional clones. Those error sources have been thought of and could partly be excluded. Of course, different explanations were possible and could be further investigated. At this point, it would be promising to conceptualize other cloning strategies, since the one presented here has turned out to be unsuccessful.

In a gedankenexperiment, the possible outcome can be analyzed. The usage of two colors would allow a better estimation of the elongation rate. The difference between the stagnation of PCP signal and increase of MCP signal is the time the polymerase needed to pass through the gene [37]. In terms of biophysics, we find arguments for and against a slower movement of RNAP in a GC related manner. After the initial helicase activity of TFIIF, the energy to move the polymerase could be generated by the synthesis reaction itself [62]. The energy released during the addition could be enough to move the SEC and might not be sensitive to the stronger bond of G and C [4]. It is also possible, that the SEC consumes ATP for its helicase activity. So far, no experiment have been performed, that excluded ATP during transcription to gain insights about the total NTP consumption of the synthesis of RNA. In eukaryotes, many enzymes are necessary to start and maintain transcription that rely on ATP and no experiment could be designed yet that measures only the movement.

An argument for a slow down in high a GC template is based on the idea, that the polymerase might follow a passive helicase mechanism to open the DNA on the next nucleotide position. The difference between the native and GC rich transcription speed is then based on the additional hydrogen bridge between G and C (3 bridges) compared to 2 bridges between A and T. The elongation speed might then correlate with the melting temperature of the DNA template.

5.4 Conclusion of ROC'n'Ribo

The result from sections 4.6 is a good example for the usability of the chip to generate real time single cell data. The traces (Fig. 4.25) confirm the findings of the flow cytometry data [78]. Experiments with the strain yCS in the microfluidic chip are used as control experiments for the flow cytometry data. The calibration of the single cell model in the paper is based on the cytometry data, because the microfluidic experiment could not capture 20,000 cells. The experiments yield between 100 and 400 traces for each ligand; not enough compared to the available flow cytometry data. But, the experiments with the cell trap chip are performed to gain additional knowledge to snap shot techniques such as

flow cytometry. The localization and distribution of proteins inside the cell is not accessible with those techniques. For the yeast strain yCS, the time course of localization of GEV into the nucleus is in agreement with the published data [15]. The produced GFP is cytosolic distributed and does not form any clusters. After size normalization, there is no correlation between the level of GEV and expressed GFP and the reduction of GEV variability in the population had no effect on the variability of the expressed GFP ([78], Fig. 14). The transcription factor is integrated into the genome, to reduce the copy number from the CENS/ARS plasmid with one to three copies to just one copy in all cells. In conclusion, GEV has to be present to react on the stimulus, the amount of GEV that is needed to yield positive induction is subordinate in those experiments. This result is supported by the transcription dynamic study, where I found the same synthesis rate for all used concentrations. To a certain extent, the usage of GEV-mCherry is redundant. The successful induction is primarily visualized by the increase and accumulation of GFP in the cell. The system could be extended to display the cytosolic levels of one of the ligands with an appropriate sensor or the abundance of the mRNA. Up to now, there is no available sensor to directly estimate the ligand abundance of neomycin or tetracycline. One possibility would be another gate which goes into an active state upon the inhibition of the other. Whereas this is again an indirect way to estimate the mRNA levels inside the cells, the single cell model described in the paper would be able to incorporate the data to improve its accuracy. Up to now, there is no measurement technique capable to produce a direct readout of the binding dynamics of neo or tc inside the cell ([78], Fig. 12). In addition to a modified biological system, the image processing has to be improved for a program with minimal human input. On the other hand, the biological results of the gate performance are impressive. The collected data show a fast acting logic gate with 3 inputs to operate. In most studies of new riboswitch-based logic gates the authors describe the performance just by endpoint measurements, omitting dynamic experiments to capture the time scales of interaction [7, 11, 32, 50, 84]. The lack of knowledge about the dynamics of a gate limits its usability in biotechnological and medical applications.

5.5 Increasing image quality for correct membrane localization

The experiments with the light inducible gene expression of a potassium channel fused to GFP have given spatial and temporal insights about the gene expression. The three different strains with codon optimization for the most used codons in yeast yield the same temporal expression of the respective construct. However, the non-optimized form of the gene has the best signal to noise ratio and produces either a potentially brighter GFP or more functional GFP. As the protein sequence is not changed, one would expect no change at all. Nevertheless, this finding suggests that the optimization was not successful and transcriptional or translational problems were encountered. Adding to this, the positive control with just the EGFP results in less intensity than the fusion protein of KCV-EGFP (Fig. 4.29). An additional experiment to quantify the amount of protein would clarify

the speculations about the origin. Possible experiments like SDS-PAGE or a western plot are best suited to obtain the protein level. Besides the gene expression dynamics, one can imagine experiments to test the functionality of the channel in the microfluidic chip. Up to now, the experiments are performed in media with high potassium concentrations, to limit the stress for the cells. A switch to low potassium concentrations after 2 hours of light induction could display the fitness of the cells and the functionality of the channel. If available, a potassium concentration dependent dye inside the cell could monitor the function directly.

While the image quality of a wide field microscope is not optimal suited to quantify the fraction of membrane incorporated channels, the images demonstrate nevertheless the localization over time (Fig. 4.26). Generally speaking, there are three ways to improve the quality of images: (I) the already recorded images can be deconvolved, (II) new experiments are recorded at a background "free" microscope, like a spinning disk microscope or (III) the biological system is altered to enhance image quality. The first option of deconvolution, a mathematical calculation on the image, which can reduce background signals, needs the point spread function (PSF) for the optical system. To derive the PSF, it is necessary to record small beads (<50 μm in diameter) on many Z slices. Those experiments have been performed and the deconvolution could be done on a computer cluster. The expected computational time for all recorded images so far exceeds several weeks. The second option of generation new data on another microscope is limited to the hardware, which is needed for the microfluidics. The system relies on pressured air, which is not accessible near every microscope and MatLab to control the valves. In addition to these obvious difficulties, other laboratories on the campus have contamination concerns, as the microscopes are mainly used for human cell lines. The resistant and sporulating fungi yeast can easily contaminate the setup and cause a downtime for cleaning. The last option of changing the biological setup to improve the imaging quality involves potentially the most work. For the quantification of the membrane the incorporation of an additional red fluorescent marker is to favor. This can either be realized with a genetically encoded tagged membrane protein or with a specific dye, which diffuses into the membrane. The marked membrane can be taken as mask for the GFP channel to correlate the two signals and estimate the fraction in the membrane.

5.6 What should be investigated next

The multiple projects presented here have different directions to advance the development and gain insight into biological processes. The microfluidic chip as platform for collaborations can be improved for trapping efficiency and longer operations. The elongation project could make use of a different genetic switch and getting a GC construct to run. The two collaborations with the yeast strains yCS and ySH can be expanded and more advanced questions can be investigated. This section will give possible topics for the next step in research.

5.6.1 Future of the traps

The fabrication of a master wafer can be improved with the laser direct writer. Already the first run resulted in usable structures. The mismatches are linked to the software conversion from the drawing tool AutoCAD® to the writer software. For instance, the intra-trap gap was expanded during the conversion from original 3 μm in the design to 4.5 μm on the wafer. Such errors can be omitted, when the design is correctly transferred to the other file format. The resolution of the machine itself is sufficient to print all trap designs in a suitable quality. After solving the exposure issue, one thing that is worth investigating is the loading procedure. Better loading could be achieved with (I) a homogeneous cell front, that all sides of the chip have in average the same number of cells run through or (II) control of the cell flow directionality. Both variants can be combined and require new design elements at the front of the cell chamber. Geometries that increase the distribution of the cells like a triangle with several pillars can be combined with a layer of valves on-top of the entry to tune the flow. In general, the size of the cell chamber can be increased to capture more cells for an experiment. Another option would be the expansion to multiple inlet channels. A chip with 4 different media inlets would have been practicable for measurements with the strain yCS as gene expression can be influenced with 3 chemicals. Besides the biological parameters of clogging, a separation technique could be directly installed on the chip to filter out or break cell clusters and introduce a majority of single cells to the traps. Even without these improvements, the chip is a great tool to study single cells.

5.6.2 Elongation project

Without any doubt, the project about the GC content and its influence on the transcription speed should be prioritized. The cloning strategy should be shifted to different backbones for cloning and integration. In addition, the integration side should be selected anew. The calculated GC fragment could be compared to high GC rich genes which occur in nature; an example is *Streptomyces coelicolor*. In addition, the influence of nucleosomes on the elongation could be studied *in vitro* on these templates. The cell free system should be prepared out of yeast extract in order to study the same polymerase under the controllable environment. Different concentrations of nucleosomes could be introduced to pack the DNA. The investigation of single transcription events *in vitro* requires a new concept of reporting the synthesis. The lack of a nucleus and the removal of mRNA from the transcription site limit the usage of PCP-GFP in an cell free system. First experiments have been undertaken with beads as anchor for the DNA to reduce diffusion. In the bachelor thesis of Katharina Decker, a construct with several repeats of PSL and the GC-rich region was tethered via biotin to Streptavidin-coated beads [19]. In the cell free mix of *E. coli* with already expressed PCP-GFP an increase around the beads over time is found. This setup could be expanded with nucleomoes and a technique with less background.

With such a system, questions could be addressed like: What are the differences between elongation *in vivo* and *in vitro*? How do nucleosomes influence transcription at this DNA template? Do the elongation rates for yeast and *E. coli* RNA polymerase differ under these conditions?

The inducible genetic construct did not show the expected increase in transcriptional activity in combination with higher concentration of inducer. There are two options to proceed: Change the inducible transcription factor or change the genomic environment. For a better comparison to other publications with the GEV system, the chromosomal environment should be changed to another locus. Although all publications on this topic show a direct effect of GEV induction and *Gal10* activity, the promoter of *Gal10* might not respond as expected.

5.6.3 Aptamer based NOR gate

The presented logic gate is characterized to best of our knowledge and without changing the genetic construct no additional experiments are needed. The model in the paper can produce any type of data and give insights about the intra-cellular neomycin concentration. The one thing that might need improvements is the image processing. The script would need a complete autonomous algorithm to segment the cells, track them over time and extract wanted features like size and fluorescence intensity in multiple channels. Best practice would be, when all genetic switches are characterized following the procedure described in Schneider *et al.* [78]. The measure of accuracy and speed for a possible genetic switch describes its dynamic performance. Other genetic constructs, like a neo-neo gate for example, should be characterized following the same methods. An ideal solution would be a data bank with all kinds of different genetic switches and gates with temporal characterization to provide a good comparison.

6 Conclusion

In both topics that inspired my work, progress was made. In the project about the microfluidic cell trap chip, the parallel chambers are most beneficial for quantitative and comparative experiments. Unlike any other published design on this topic, the chip offers the possibility to remove daughter cells to run long time experiments, while imaging different strains and media conditions in parallel. The fast switching of the on-chip valves tops the application field off. All the different trap designs are capable of holding cells for long time periods and prevent clogging of the device. During the loading process, the designs become distinguishable. The traps I designed, from the pointing pentagons to the ovals (Fig. 4.1, 5 and 9-11) are comparable with the Alcatraz chip [15]. The catch rate of those designs are repetitively at 70%. The HYAA design with its L-shaped traps could attain rates up to 90%. The unfilled traps in Fig. 4.5 are mostly at the sides of the chip, where a reduced flow of the media is expected and therefore less cells can be caught. Only 2% of suitable traps in the center of the chamber remain empty. The biological question of interest in this study resulted in a genetic background that is responding to an externally applied β-estradiol concentration as an inducer. The trapped cells generate bright spots at the transcription site that can be tracked and an intensity estimate can be extracted. The transcriptional output of the chosen promoter and transcription factor is not linked to the amount of inducer in the medium. More β-estradiol increases the fraction of cells that enter active transcription, but every transcription event results in the same response intensity drawn from the same distribution (Fig. 4.17). As GEV and PCP could be excluded as limiting factors (Fig. 4.16), the locus itself might be restricting transcription. Nucleosomes and the modification patterns of histones are known to influence the transcriptional output of genes. Another explanation might be the high traffic on the locus. The genes Gal1 and Gal7 are in close proximity and transcription of both of these genes can influence the output of Gal10 [57].

The chip is a perfect platform to work on collaborations. I have shown in two examples, that some scientific questions can be best addressed by catching cells in a device in order to be imaged with the microscope. The results from yCS reflect a high throughput approach for time resolved single cell traces. Different conditions can be measured simultaneously in one experiment with the correct chip. This reduces manual work and improves the comparability of experiments. The performance of the inducible NOR gate can compete with similar published aptamers [84]. The project with ySH shows similar results. At the one hand, one can follow the expression dynamics of single cells for the light induced GFP expression. On the other hand, the localization of the potassium channel can be estimated from the images. Another chip could be used in these experiments. As multiple different genetic backgrounds have been produced, chips with connected and parallel cell chambers are best suited to get the most out of one experiment.

7 Material

Substance	Supplier	Used for
PDMS	Corning	Microfluidic chips
Methyltrichlorosilane	SigmaAldrich	Coating of wafers
SU8 2005		Wafer fabrication
SU8 developer		Wafer fabrication
SU8 2025		Wafer fabrication
AZ4562		Wafer fabrication
Q5 polymerase	NEB	Cloning
Phusion polymerase	NEB	Cloning
Herculase II		Cloning
Synthesized DNA	Thermofisher	Cloning
BSA	SigmaAldrich	Chip preparation
PCR CleanUp kit	NEB	Cloning
Gel Ex kit	NEB	Cloning
MiniPrep kit	Qiagen	Cloning
dNTPS		Cloning
Ammonium sulfate	SigmaAldrich	Yeast media
Yeast Nitroxigen Base	SigmaAldrich	Yeast media
Glucose	SigmaAldrich	Yeast media
Galactose	SigmaAldrich	Yeast media
Raffinose	SigmaAldrich	Yeast media
Amino acid drop out	SigmaAldrich	Yeast media
Ethidiumbromide	SigmaAldrich	Gel staining
Tygon Tubing 0.51 mm ID	Darwin Microfluidics	Micofluidic
Metal fittings 23G	Darwin Microfluidics	Micofluidic
Cover glas 50 x 24 mm	Carl Roth	Micofluidic
Ethidiumbromide	SigmaAldrich	Gel staining

Table 7.1: Material used

7.1 Devices

Type of device	Manufacture
Inverted microscope	Nikon
Centrifuge	CellStar
Spin Coater	Laurell Technologies
Ozon UV lamp	Novascan
Plasma oven	Diener electronics
Soleniod valves	The Lee Company
Gel iX20 Imager	INTAS
SimplyAmp PCR	LifeTechnologies

7.2 Plasmids

Identifier	Features	Origin
pJD103	24x MSL	[37] (addgene #45162)
pJD102	24x PSL	[37] (addgene #45163)
pJD101	PCP-GFP-GFP	[55] (addgene #35194)
pJD100	MCP-GFP-GFP	[37] (addgene #45929)
pJD130	MCP-mScarlet	This study
pJD129	MCP-yomRuby2	This study
pJD146	TEV-GEV-mCherry	[108]
pJD151	STE5-GEV-mCherry	[78]
pJD121	10xPSL - GC66 - 24 PSL for Gal10@2097	This study
pJD122	10xPSL - GC66 - 24 MSL for Gal10@1252	This study
pJD123	10xPSL - AT66 - 24 MSL for Gal10@0818	This study
pJD164	Cas9	[21] (addgene #43802)
pJD165	gRNA tempalte	[21] (addgene #43803)
pJD166	gRNA - Gal10 @ 2097r	This study
pJD168	gRNA - Gal10 @ 1252r	This study
pJD167	gRNA - Gal10 @ 0818r	This study

8 Methods

8.1 Competent yeast cells

Competent yeast cells are produced following two different protocols. The first is part of the Frozen-EZ yeast transformation kit II form Zymo Research. The manual contains a 4 step protocol, growing, washing, preparing, transform the cells. The second protocol is known as lithium-acetate method. The steps required for competent cells is as follows:

	Solutions needed
TE 10x	100mM Tris pH8 10mM EDTA pH8
Li-Ac 10x	1 M Li-Ac filter sterilized (FS)
50% PEG_{4000}	prepare fresh every 1-2 month because evaporation may lead to higher concentration of Peg in the transformation and this may lead to lower efficiency
TE/Li-Ac	(to prepare freshly): 1 vol 10xTE + 1 vol 10xLi-Ac + 8 vol H2O
TE/Li-Ac/PEG	(to prepare freshly): 1 vol 10xTE + 1 vol 10xLi-Ac + 8 vol 50% PEG4000
	working instruction
1.	Dilute cells from an overnight culture to an OD_{600} of 0.35 and should be collected at an OD_{600} of 1.4 after appr. 6 h
2.	Collect cells from at least 10 ml in the centrifuge for 3 min at 3000 rpm
3.	Remove supernatant and redissolve the pellet in 1 mL water and transfer into 1.5 mL tube.
4.	Spin the cells once more at 6000 rpm for 3 min.
5.	Resuspend the cell in 1 mL TE/Li-Ac solution.
6.	Centrifuge the cells 1 min 6000 rpm.
7.	Resuspend the cells in 100 µl of sterile TE/Li-Ac. Incubate at 30 °C in a waterbath until DNA is ready.
8.	Prepare in 1.5 ml sterile Eppendorf tube the DNA to transform: 10 µl of Carrier DNA at 10 mg/ml (denature approx. 5 min. at 95 °C before use) 1-3 µg cleaned PCR product
9.	Add 100 µl of cells in TE/Li-Ac to the DNA
10.	Add 700 µl TE/Li-Ac/PEG. Mix with pipet (try to avoid air bubbles) Incubate 2-2.5 hours at 30 °C in a waterbath. Invert the tubes from time to time.

11.	Add DMSO till final concentration is 10% and vortex 5 sec
12.	Incubate at 42 °C for 15-20 min
13.	Centrifuge the cells 1 min 10000 rpm.
14 .	When using NatMX (Nourseothricin resistance) or KanMX (G418; Geneticin resistance) as selection marker, the transformation efficiency can be increased by taking the cells up in 1 ml YPD, and allowing the cells to recover for 1 hour at 30 °C. In case of other selection markers, directly proceed to step 15
15.	Centrifuge the cells 1 min 6000 rpm.
16.	Rinse pellet in 1 ml sterile water.
17.	Centrifuge the cells 1 min 6000 rpm.
18.	Resuspend cells in 100 µl sterile water
19.	plate 50% on a selective plate. add the other 50% to 1 ml YPD and incubate ON under agitation at 30 °C
20.	**Day 3** Centrifuge the 1 ml culture 3 min 4000 rpm.(or just plate out 100 ul of the o/n culture)
21.	Resuspend cells in 100 µl sterile water and plate all on a selective plate.
22.	**Days 6-7** Restreak single colonies onto fresh selective plates.

8.2 Microscope setup

The microscope and its hardware was constantly expanded. Based on a Nikon Ti-Eclipse body, I use a LumecorSpectraX light source, with 540/25 filter for GFP and 670/35 filter for RFP, a fluorescence emission filter wheel, which became obsolete with the installation of a beam splitter from Hamamatsu and a second ORCA Flash 4. Due to software problems with the open source program MicroManager, the installation and operation of the dual camera setup had to be shifted to Nikon NIS Elements. The microscope body was placed inside a incubator box (LifeImagingServices) and kept at 30 °C. The white light source was focused each time following Köhlers illumination. The experiments for the transcription dynamics were conducted with a 100 x 1.45 NA objective from Nikon, 9 Z-slices in 500 nm spacing, 200 ms exposure in a 10 s interval with 18 % power of the Lumencor. Bright-field images with 10 ms exposure time and an CooLED intensity of 5%. The order of the experiment was: Position > Time > Channel > Slice. For the experiments for the co-operations, the 60 x 1.40 NA objective is used with 5 Z-slices in 1μm spacing. Imaging interval for the ROC'n'Ribo part was every 5 min and for the PLY part every 15 min. The order of the experiment was: Time > Position > Channel > Slice.

8.3 Chemical competent *E.coli*

All transformations with *E.coli* strain DH5α are performed following the protocol of zymo research *Mix & Go* kit. Cells are grown to mid-log phase (OD_{600} = 0.4 - 0.6) and placed on ice for 10 min . The cells are harvested by centrifugation at 2500 g for 10 min in a cooled centrifuge. Supernatant is removed and 5 ml ice-cold wash buffer is added to the pellet. Cells are re-pelleted again. The final step is the suspension in 5ml pre-chilled competent buffer. The cells are stored in a -80°C refrigerator.
The transformation includes thawing on ice, addition of 1-5 μL circular DNA, an hour of incubation at 37 °C under constant shacking and last the spreading onto pre-heated agar plates with the corresponding resistant marker. In most cases Ampicillin is used as selective pressure.

8.4 DNA Manipulation

Polymerase Chain Reaction on Stem Loops

The polymerase chain reaction (PCR) is used to produce fragments of double stranded DNA. Several polymerases have been tested for their ability to process repetitive sequences like the PP7 stem loop. Q5© and Phusion© from New England Biolabs, Taq and AccuPrime™ GC-Rich DNA polymerase from Invotrogen™ and Herculase II from Aligent have failed to produce exclusively the correct size band with several sets of primers (located up to 200 bp away from the stem loops). With the available templates from Daniel Zenklusen (addgene # 45162 and # re45163) [37] the PCR shows all kinds of short fragments with the length difference of one repeat (Fig. 4.8). The concentration of the right sized product is the lowest because of the time to produce it. Smaller fragments can be amplified several times in the time one full length product. Despite the lack of melting during the productive cycle, primers can bind on multiple sites on the plasmid to produce shortened fragments. Aborted and truncated side products can be used as wrong templates in the next cycle. The addition of DMSO (0.5-5% in 0.25% steps, example is shown in Fig. 4.9) and Mg^{2+} (0.1 mM to 2 mM in 0.1 mM steps) reduces the shortened fragments, but annealing of the repeated region with itself is not prevented, resulting in a blur above the correct size.

Restriction Site Cloning to modify pJD121

The PCR based modification of plasmids with repeated units of stem loops does not result in an usable construct. I designed and ordered a whole integration cassette that could be directly digested with enzymes and used for integration. The plasmid pJD121 contains [Gal10 overlap - 10x PSL - GC66 - 18x PSL - loxP- HIS3 - loxP -Gal10 overlap] and

is digested with HindIII and XhoI to cut out the 18x PSL stem loops. The gap is filled with 24x repeats of the MSL with suitable ends for ligation (pJD122). Multiple clones are sequenced to validate the correct ligation. Using pJD122, the GC rich sequence is replaced by an AT rich one with the restriction sites NotI and BamHI (pJD123). The AT and GC rich DNA fragments are produces based on a Matlab script, with restricted codon usage in order to maintain a uniform GC content distribution.

Gibson Assembly of gRNA Plasmids

To target the Gal10 gene at the specified sequence 3 different gRNA containing plasmids based on pJD165 are constructed, following the protocol of Gibson *et al.* [30]. Through PCR, the whole plasmid is amplified with overlapping sequences at each end. This sequence contains the gRNA and is used in a Gibson assembly to correctly close the plasmid. The sequences for the gRNA are:

```
2097 GGAAAATCTGTAGACAATCTTGG 2075
1275 GCGCCTATATAAGCACTATCAGG 1252
 818 CCTTCATTTTCATTGTAGG 800
```

Each of the gRNA sequence is included in the forward primer (5'gRNA-GTTTTAGAGC TAGAAATAGC-3') and reverse primer (5'-gRNA rc-GATCATTTAT CTTTCACTGC GGAG-3').

After transformation of the gibson mix, plasmids are amplified and purified with a MiniPrep kit from Qiagen

Colony PCR with *E. coli* and Yeast

The search for the right clone is started with a colony PCR on the insert. For the cloning in *E. coli*, a colony is diluted in 25 μL water and 1 μL is transferred into the PCR master mix.

The access of DNA directly from yeast requires an opening of the cell wall. A colony is transferred into either 20 mM NaOH, 0.2% SDS or a zymolase digest (4 units/μL) to reach the DNA inside. SDS and NaOH require a boiling step before a PCR can be performed. Colonies in zymolase solution are incubated at 37°C for 20 min.

Agarose Gel-Eletrophoresis

Agarose gels are a basic tool to separate DNA based on its size. I used TAE buffer and for all gels 130V for 45 min. Staining is done in an ethidium bromide bath and images are taken with a digital camera from INTAS. Fig. 4.8 and 4.9 are examples of PCRs with

8.5 Microfluidic chips

8.6 Fabrication of master wafer

The layout was designed in AutoCAD (AutoDesk) and were ordered as foil plot for channels or chrome plot, for traps. The fabrication of a wafer was performed as depicted in Table 8.2:

As last step and before using the wafer, a coating with a highly hydrophobic chemical is performed. The wafer are placed inside a vacuum chamber, together with 100 μL of 1H,1H,2H,2H-Perfluorodecyltrichlorosilane sealed and vacuum applied. Through evaporation, a thin layer is applied to the silane oxid wafer. This coating increases the re-usability, as the pouring agent can be better removed during chip fabrication. Vacuum is applied over night, and the treatment is repeated after the casting of 25 chips.

8.7 Casting of microfluidic chips

A two compound silane, polydimethylsilane (PDMS) is poured into weighting trays for mixing and degassing. For the control layer, 40 g compound and 4 g catalyst are mixed. The flow layer has a ratio of 20:1. After 2 h of degassing, the PDMS is cast onto the wafer; for the flow layer this is directly done in the spin coater. The thickness of the flow layer is adjusted by rotation at 400 rpm for 30 s, then accelerated to 1700 rpm for 60 s and slowly stopped during a 30 s deceleration. Both wafers are degassed again for 1 h before the curing at 75°C is performed in an preheated oven. The flow layer is already taken out after 13 min to have partially cured compound for a better bonding of the two layers. After additional 30 min, the control layer is removed from the oven and cooled down at room temperature. The cutting in the actual chip size is followed by the alignment of the corresponding control part into the right area on the flow layer. Alignment marks (a big and a small cross that fall together) are used for visual control of the process under a stereo microscope. The chips are now cured over night to firm the connection between the layers in the oven. A 0.5mm thin puncher makes the connection from the channels through the PDMS possible. The prepared PDMS chips have to be bond to cover slips for the usage of high magnification oil objectives. A plasma oven treatment for 10 s activates the surface of the glass and the PDMS to metal plugs connect the tygon tubes with the chip.

8.8 Experiments for transcription dynamics

Preparation

The preparation of an experiment with the microfluidic chip includes the following:

	Layer 1: Cell traps 7.5 µm with SU8-2005
1.	Wash wafer with 1. acetone, 2. isopropanol
2.	Coating with TI Prime “Haftvermittler”
3.	Spinning: 500 rpm, 10 s; 1500rpm, 30 s
4.	Softbake: 5 min to 50 °C Exposure: 120 mJ/cm^2 at 365 nm
5.	Exposure mode: hard contact Alignment gap: wafer and mask 15 µm
6.	Resting for 10 min
7.	Post exposure bake (PEB) 3 min to 50°C, 3 min to 65°C, 3 min to 95°C
8.	Developing: 40 s
9.	Rinse with isopropanol, than water
10.	Hard bake: 5 min at 130 °C
	Layer 2: Media in- and outlets 20 µm with SU8-25
1.	Wash wafer with 1. acetone, 2. isopropanol
2.	Coating with TI Prime “Haftvermittler”
3.	Spinning: 500 rpm, 10 s; 4000rpm, 30 s
4.	Softbake: 5 min to 50 °C, 5 min to 65 °C, 5 min to 95 °C Exposure: 300 mJ/cm^2 at 365 nm
5.	Exposure mode: hard contact Alignment gap: wafer and mask 40 µm
6.	Resting for 10 min
7.	PEB: 6 min to 95°C
8.	Developing: 3:30 min
9.	Rinse with isopropanol, than water
10.	Hard bake: 2 h to 130°C, then slow cooldown
	Layer 3: Round channels 10 µm with AZ4562
1.	Wash wafer with 1. acetone, 2. isopropanol
2.	Coating with TI Prime “Haftvermittler”
3.	Spinning: 500 rpm, 10 s; 1700rpm, 30 s
4.	Softbake: 10 min at 100C Exposure: 700 mJ/cm^2 at 405 nm
5.	Exposure mode: hard contact Alignment gap: wafer and mask 40 µm
6.	Resting for 5 min
7.	Developing: 8 min
8.	Rinse with isopropanol, than water
9.	Hard bake: 15 min at 130C, then slow cool down

Table 8.2: Fabrication protocol for silicon wafers.

1. filling of reservoirs
2. cutting 30 cm of tygon tube for the connection from reservoir to the chip
3. connecting water filled tubes to the valve inlets
4. flushing the chip with media
5. testing of the on-chip-valves
6. calculating flow rate
7. adjusting pressured air for flow rate

Loading

Cells are grown overnight before the experiment and diluted in the morning to an OD_{600} of 0.01. After reaching an OD_{600} of 0.05-0.075, the cells are loaded by hand with a syringe into the chip. The applied pressure on the media is reduced during the loading to 0.5 bar in order to expose the cells to lower hydro-static stress. On average 150 μL of cell suspension is injected to yield a loading efficiency of 70%. The loading itself is performed with an increasing pressure on the syringe by hand. The cell chamber is constantly monitored to review the process. If bigger cell aggregates get stuck inside the chamber, pulses of high pressure are applied to the syringe to flush those out. It is recommended to tape the loading tube to the stage of the microscope to prevent unwanted movement of the chip.

Temporal Ordering of Positions

In the experiments with PCP-GFP, I encountered strong photobleaching. To prevent bleaching of positions, that are yet to record, an optimization on the temporal order of the experiment is performed (Fig. 8.1). Rules for the optimization is a general maximization of the time difference of two neighboring positions being recorded. For all experiments of the transcription dynamic, the 100 x objective is used and the optimization has to be adapted to the 60x chip and 100x chip. In the 60x chip, the FOV with the 100x is only occupying a subset of the area intended to be observed (47 μm^2 compared to 17 μm^2).

8.9 Image procession

yJD - Transcription
Each position of an experiment is scanned manually in the maximum projection for cells that show a spot. The spot has to appear and disappear within the recording. True positive spots move slightly during the acquisition inside the nucleus, so stationary spots are excluded. A rectangle is drawn around each cell and the 4D data (xyzt) is copied for this region of interest (ROI). Each cell is analyzed by an algorithm, that sums the intensity

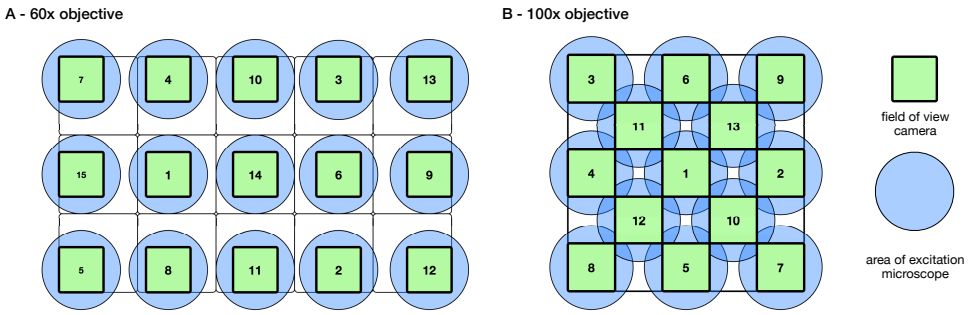

Figure 8.1: Temporal optimization of experiments with short recording interval. The order is optimized to minimize cross-bleaching from overlapping areas of excitation from the microscope.

with an 3D Gaussian fit under the assumption of a point light source. The background intensity is calculated by the pixels adjacent to the found spot voxel. If no spot is detected, the intensity is calculated on the last coordinate. Movement of the spot is limited by diffusion at XX mm/min. Subpixel coordinates are determined with the Gaussian fit for intensity calculation. In addition, a variable estimates the existence of a spot in each frame. Relevant files and processes are:

1. NIS_RFP_GFP_max_save.ijm
2. *hand selection of rois for cells*
3. Read_MAX_save_ROI.ijm
4. Duplicate_crop_NIS_GR.ijm
5. Christians code
6. statistics on spot intensity etc.
7. Leos Code

yCS - Logic Gate

The traces in Fig. 4.25 are based on a semi automatic code. In the first frame, the position of a cell is manually set. For the rest of the recording, the intensity of a circle around the position is summed for all Z slices and corrected for back ground intensity and normalized by the size of each cell.

ySH - Ion Channel

Each frame is divided in 256 ROIs, for each ROI, the summed intensity in Z is calculated over time. Each ROI is taken into account that has an intensity level about a threshold. In average a valid ROI has 1.2 cells. ROIs with jumps in the trace, which can only explained by a lost or new cell, are rejected. The signal is normalized by the number of pixels that are brighter than the threshold of 1500 units as measure for the size of a cell.

9 Appendix

9.1 Cirriculum Vitae

Personal Information

Place of birth	Frankfurt am Main
	31.05.1988
Contact	Koblenzer Str. 52
	60327 Frankfurt am Main
	jdiemer@gmx.de
	0151 22 39 21 29

Education

12/13-12/18	**PhD** at Technische Universität Darmstadt, Title: "Single cell measurements in microfluidic chips to determine the dynamics of transcription during induction", AG Koeppl
10/11-09/13	**Master of Science** at Technische Universität Darmstadt, Title: "Superresolution Microscopy with Blinking and Bleaching Fluorophores", AG Meckel
10/08-09/11	**Bachelor of Science** at Technische Universität Darmstadt, Title: "Determination of the catalytic activity of peptid based mimetica of nickel superoxid dismutase with stopped flow", AG Buntkowski
09/99-08/08	Higher education entrance qualification, Kurt-Schumacher-Schule, Karben; graduated with honours in physics

9.2 Supporting Material

All raw data and code of processing the images and files can be obtained upon request.

List of Figures

1.1 Galactose gene network, adapted from [40]. The GAL4 protein activates transcription of *Gal1*, *Gal2*, *Gal3*, *Gal5*, *Gal6*, *Gal7*, *Gal10* and *Gal80* (cyan arrow). The three core regulatory proteins are GAL3, GAL4 and GAL80. They activate and inactivate each other in turn and detect the presence of galactose (green lines). Dotted lines may indicate correlation of interactions. GAL1, 7 and 10 show the highest response on galactose induction. 3

1.2 Pol II assembly after Soutourina *et al.*, 2018 [87]. **A** Transcriptional activators bind on enhancer regions. **B** Chromatin remodelers take action and make the core promoter and transcription start site (TSS) accessible. **C** Mediator bridges the enhancer to the core promoter and recruits the GTFs. **D** RNAPII leaves the promoter into elongation. 4

1.3 Scheme of the torpedo termination mechanism. Upon reaching the polyA site (red cross), Rat1 cuts the RNA and starts to digest the newly synthesized RNA. The cleaved RNA gets its polyA tail and moves on to translation, while Rat1 collides with RNAP and stops its transcriptional activity. 7

1.4 Scheme of bursting. The number of RNAPs dictates the burst size, while the time between two bursts is considered as burst frequency. This stochastic behavior results in transcriptional noise and is amplified by translation. 8

2.1 Structure of the PP7 stem loop. The color indicates the strength of the binding from weak (blue) to red (strong). 12

2.2 Single transcription event. After promoter release, RNAP processes the stem loop region, resulting in an step-wise increase in fluorescence after binding of PCP-GFP. A plateau is reached, when RNAP transcribes the gene body. After release of the pre-mRNA from the TS, the signal is lost. Multiple polymerases generate a superimposition of this signal. 12

2.3 Collection of microfluidic traps. **A** Design taken from [76]. Yeast cells are trapped during assembly of the chip. **B** Picture taken from [20]. A two layer design traps the cells between the cover glass and the ceiling. **C** Design from [64]. Alcatras trap design helps to prevent clogging of the device as daughter cells are washed out. 16

4.1 Examples of fabrication errors. The desired trap design is highlighted in yellow for two different rounds of wafer producing under the same settings. 19

4.2 The second design iteration. Traps 1,2,6 are directly adapted from [15, 43]. Flow direction is from bottom to top. Scale bar equals 6 μm. 19

4.3 Sketch of the wafer (100 mm diameter) for the second design iteration. **A** The colors correspond to one layer of fabrication: cyan is the control layer, red is the AZ layer, black is the SU8 2025 and purple the cell traps with SU8 2005. 3 different chip layouts are tested: **B** 7 independent chambers **C** 6 x 2 dependent parallel chambers. The 6 x 2 chambers have in addition a valve to separate the two strains during the loading process. **D** The last design 3x3 independent parallel cell chambers 20

4.4 Montage of traps made with direct laser writing. Drawn in yellow are the specified designs, images taken at 100x . Resolution and section are matched. Scale bar equals 20 μm. 21

4.6 Histogram of the loading efficiency for the transcription experiments. In total, 5000 cells are observed in 100 FOVs. **A** 74 % of all traps are filled with cells. for the strain yJD, the most common caught cluster size is 2 cells (33%) followed by single cells (18%) **B** Statistics without empty traps. 25% of traps with cells are occupied by one single cell, 43% have caught a doublet. The remaining 32% of captured cell aggregates consists of 3 cell or more. 24

4.7 Scheme of experimental system. PCP-GFP and GEV-mCherry are constantly expressed and present in the nucleus and the cytosol respectively. Upon exposure with β-estradiol, GEV is relocated into the nucleus to start transcription on the GAL10 locus. This re-localization can be observed in the red channel (right side, top). Once transcription is activated, the stem loops are synthesized and can be bound by PCP. This event can be displayed as spot appearing inside the nucleus (right side, bottom). Following the intensity of the spot over time describes the dynamic of transcription. 27

4.8 Example of agarose gel with PSL product. The used primer pair binds 200 bp upstream and 300 bp downstream the PSL region to minimize side products. Annealing temperatures tested range from 52°C (right) to 66°C (left). The product is expected to run at 2100 bp. 27

4.9 Example of agarose gel with PSL product. The product is expected to run at 1600 bp. DMSO is tested ranging from 0% to 4% under 3 different Mg^{2+}-concentrations (0, 0.5 and 1 mM additional). 28

4.10 Montage of a cell in the channels bright-field, GFP and RFP from top to bottom. The appearance of a spot can be spotted in the GFP channel on the lower cell. The re-localization of the transcription factor is only visible at the end of the recording (after 20 min). 29

4.11 A typical measurement. Top row shows the cell in a maximum projection, cell membrane highlighted in white. The corresponding intensity values are beneath each frame and marked with a vertical line. Several transcription event can be observed in this cell. 29

4.12 Raw data wit different recording times. **A** Single cell traces for experiments with 10 s interval over the period of 1 h. Strong bleaching is visible and the last third of the trace hold little useful information. Two sequential time points are plotted**B** Pooled single cell traces for sequential recordings of 20 min. The total information captured in 1 h is increased with shorter movies at different locations. Each line represents a trace from a single cell and the different colors are for a better visualization. The fast time decay of the signal is caused by photobleaching. 30

4.13 Spot intensity by its duration with 50 nM of β-estradiol. **A** The accumulated spot intensity is increasing with the duration of the spot. **B** Even when normalized by its duration, a correlation at the beginning is observable. Shorter spots can include abortive events and those have a lower spot intensity as less stem loops are present. 31

4.14 Raw data from experiments with different β-estradiol concentrations. **A** 5 nM, **B** 50 nM, **C** 500 nM β-estradiol. The increasing number of responders is visible in the higher number of traces generated by one experiment (109 to 271 to 301). Each line represents a trace from a single cell and the different colors are for a better visualization. The fast time decay of the signal is caused by photobleaching. 32

4.15 Time-resolved dose-dependent transcriptional activity. The cumulative fraction of cells showing a spot per total cell count for the time since β-estradiol (time 0 min). The curve converges to the average for each dose, as all cells are taken into account that have shown activity from former positions. **A** Lower concentrations (0.5 to 5 nM β-estradiol, pink to orange). The curves show little influence of the inducer on the cells. The fraction of responders does not change significant around time point 0. Even before β-estradiol is introduced to the chip, cells show activity. **B** For 10 to 50 nM an increase at time 0 is visible. **C** The highest concentrations result in the highest number of induced cells (500 nM). Each point is averaged over at least 3 independent experiments. **D** Average fraction of active cells per dosage. A positive correlation is visible. 33

4.16 Dotplot of the integrated trace intensity in relation to other parameters. **A** The calculated GFP value for each cell is size corrected and rather uniform (one log difference). **B** Despite a size correction, a strong correlation between GEV and PCP expression level is observable for low expression levels. The amount of GEV in the cell (**C**) and its size (**D**) do not have an influence on the total spot intensity. These data are based on 3500 cells. . 34

4.17 Spot duration and intensity separated by dosage. **A** The calculated spot duration does not change with the concentration of β-estradiol. Cells that show a spot over 90% of imaging time are rare events (0.9%). **B** Average spot intensity shown for each dose. No correlation between the dosage and the spot intensity can be found. **C** Average pause duration between two spots per induction level is at ...± s. The number of events for each dose varies: 369, 487, 547, 470, 402, 607, 2335, 1627, 5164 and 2187 spot appearances have been observed from 6000 cells. 35

4.18 Dotplot of the integrated trace intensity in relation to other parameters. **A** The calculated GFP value for each cell is size corrected and rather uniform (one log difference). A slight correlation between the GFP level and trance intensity is observable for cells induced with more than 100 nM. Each set contains at least 800 cells. 35

4.19 Correlation matrix from variables. The GFP and RFP values are corrected for size, as a dependency is here expected. The correlation between the dose and RFP is based on the re-location of the transcription factor. As the maximum projection is used to estimate the RFP signal in the cell, the brighter nucleus does bias the data. 36

4.20 Properties of synthetic DNA. **A** Relative GC content in a 40 nt window. The calculated GC66 sequence (green) has a consistent GC content at 66% on average. The native Gal10 gene (orange) shows fluctuations, while the AT66 control sequence (lilac) is at 33% GC. **B** Estimated melting temperature with after SantaLucia (1998) [77] in a 14 nt window representing the transcription bubble. 38

4.21 Plasmid map of one of the final construct. The overlapping regions correspond to GAL10. 38

4.22 Genome map with annotated gRNAs. The already modified Gal10 locus with 14 PSL has 3 available gRNA targets (at position 800, 1252 and 2097) in its sequence. gRNA is searched with E-CRISP [34]. Below are the sequences of the 3 targets in 5' to 3' orientation. Decreasing number indicate the complementary DNA strand. 39

4.23 Exemplary inference result. **A** Based on the observations (red) a likely path of produced stem loops (black). **B** Starting times for each of the estimated 8 polymerases 40

4.24 Posterior distribution of selected parameters. **A** Estimated elongation rate. The blue curve shows the prior probability density, the red the posterior. The physiological regime is highlighted in gray. The data are best explained by a elongation rate of 30 nt/s and a termination rate of 0.033 events per second **B** Probability density for the termination rate 41

4.25 Qualitative population dynamics of the NeoTc gate. **A** Histogram of time-lapse flow cytometry data from the +2 h experiment. Uninduced cells (black) show a broader distribution than cells induced at 0 h (green) and cells induced at 0 h and repressed at +2 h (gray). **B** The picture series shows a representative cell trapped on the microfluidic chip and recorded over time by fluorescence microscopy. Scale bar 5 µm. **C** Density plots of segmented single-cell traces. Random examples for individual traces are highlighted in orange. Induction with 10 nM β-estradiol, repression with 100 μM neomycin and/or 250 μM tetracycline. Color becomes darker for more cell traces approaching the median. 43

4.26 Example images for the light induced gene expression of a potassium channel. The channel gets integrated into the endoplasmatic reticulum membrane around the nucleus. Scale bar equals 5 μm 44

4.27 Experiment with strain PLY240. Display of single cell traces from 3 different experiments. Left: Control experiment with no additional light excitation, images taken every 15 min. n=460 **B** Images in the GFP channel are taken every 5 min. This is enough light to trigger gene expression. n=389 **C** Illumination of cells for 3 minutes between microscopic image recordings for the first 30 min of the experiment. n=512 **D** Comparison of the dynamic of the GFP expression. For better visualization the mean of the single cell traces is plotted 45

4.28 Result of parallel experiment with 3 different strains. The light induction of gene expression is continuously on. Images are taken in a 10 min interval. **A** **y**east**o**ptimized KCV-yoEGFP, n=176 **B** KCV-yoEGFP, n=219 **C** KCV-EGFP as reference, n=481 **D** Comparison of relative intensity increase of the average cell per experiment. 45

4.29 Result of parallel experiment with 2 strains. **A** The strain with singleplasmid EGFP, contains the light sensitive TF and an EGFP as target gene. n=161 **B** spKCV-EGFP expressed after induction with light. n=383 **C** The individual dynamic of the plasmid is comparable, the drop of intensity results from out of focus time points. What is surprising though, is the total intensity after induction. Cells with KCV-EGFP express more functional EGFP than cells without the channel. 46

4.30 Experiment with stain FM4-64. Example cell with KCV-EGFP after 1h of light induction. The image montage shows the localization of the ion channel in the ER membrane and the dye FM4-64 in the membrane of the vacuole. The composite images shows no co-localization of the signals. Scale bar equals 5 μm. 46

8.1 Temporal optimization of experiments with short recording interval. The order is optimized to minimize cross-bleaching from overlapping areas of excitation from the microscope. 67

Bibliography

[1] K. P. Adamala, D. A. Martin-Alarcon, and E. S. Boyden. Programmable RNA-binding protein composed of repeats of a single modular unit. *Proceedings of the National Academy of Sciences*, page 201519368, 2016.

[2] E. A. Alcid and M. S. Jurica. A protein-based EM label for RNA identifies the location of exons in spliceosomes. *Nature Structural and Molecular Biology*, 15(2):213–215, feb 2008.

[3] B. L. Allen and D. J. Taatjes. The Mediator complex: a central integrator of transcription. *Nature reviews. Molecular cell biology*, 16(3):155–166, mar 2015.

[4] K. J. Armache, H. Kettenberger, and P. Cramer. The dynamic machinery of mRNA elongation. *Current Opinion in Structural Biology*, 15(2):197–203, 2005.

[5] E. Bertrand, P. Chartrand, M. Schaefer, S. M. Shenoy, R. H. Singer, and R. M. Long. Localization of ASH1 mRNA particles in living yeast. *Molecular cell*, 2(4):437–45, oct 1998.

[6] X. Bi. Heterochromatin structure: Lessons from the budding yeast, oct 2014.

[7] B. A. Blount, T. Weenink, and T. Ellis. Construction of synthetic regulatory networks in yeast. *FEBS letters*, 586(15):2112–21, jul 2012.

[8] A. R. Buxbaum, G. Haimovich, and R. H. Singer. In the right place at the right time: visualizing and understanding mRNA localization. *Nature Reviews Molecular Cell Biology*, 16(2):95–109, dec 2014.

[9] M. Chabert and J.-L. Viovy. Microfluidic high-throughput encapsulation and hydrodynamic self-sorting of single cells. *Proceedings of the National Academy of Sciences*, 105(9):3191–3196, mar 2008.

[10] J. A. Chao, Y. Patskovsky, S. C. Almo, and R. H. Singer. Structural basis for the coevolution of a viral RNA-protein complex. *Nature structural & molecular biology*, 15(1):103–5, jan 2008.

[11] J. Chappell, M. K. Takahashi, and J. B. Lucks. Creating small transcription activating RNAs. *Nature Chemical Biology*, 11(3):214–220, mar 2015.

[12] G. Charvin, F. R. Cross, and E. D. Siggia. A microfluidic device for temporally controlled gene expression and long-term fluorescent imaging in unperturbed dividing yeast cells. *PLoS ONE*, 3(1):e1468, jan 2008.

[13] A. M. Corrigan, E. Tunnacliffe, D. Cannon, and J. R. Chubb. A continuum model of transcriptional bursting. *eLife*, 5:1263–1271, feb 2016.

[14] A. Coulon, M. L. Ferguson, V. de Turris, M. Palangat, C. C. Chow, and D. R. Larson. Kinetic competition during the transcription cycle results in stochastic RNA processing. *eLife*, 3, 2014.

[15] M. M. Crane, I. B. N. Clark, E. Bakker, S. Smith, and P. S. Swain. A microfluidic system for studying ageing and dynamic single-cell responses in budding yeast. *PloS one*, 9(6):e100042, jan 2014.

[16] R. D. Dar, B. S. Razooky, A. Singh, T. V. Trimeloni, J. M. McCollum, C. D. Cox, M. L. Simpson, and L. S. Weinberger. Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proceedings of the National Academy of Sciences of the United States of America*, 109(43):17454–9, oct 2012.

[17] R. D. Dar, B. S. Razooky, L. S. Weinberger, C. D. Cox, and M. L. Simpson. The Low Noise Limit in Gene Expression. *PloS one*, 10(10):e0140969, jan 2015.

[18] P. E. de Martin, Y. Du, P. Novick, and S. Ferro-Novick. Ice2p is important for the distribution and structure of the cortical ER network in Saccharomyces cerevisiae. *Journal of Cell Science*, 118(1):65–77, 2005.

[19] K. Decker, 2018.

[20] D. Di Carlo, N. Aghdam, and L. P. Lee. Single-cell enzyme concentrations, kinetics, and inhibition analysis using high-density hydrodynamic cell isolation arrays. *Analytical Chemistry*, 78(14):4925–4930, 2006.

[21] J. E. Dicarlo, J. E. Norville, P. Mali, X. Rios, J. Aach, and G. M. Church. Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. *Nucleic Acids Research*, 41(7):4336–4343, apr 2013.

[22] G. Dieci, G. Fiorino, M. Castelnuovo, M. Teichmann, and A. Pagano. The expanding RNA polymerase III transcriptome, dec 2007.

[23] E. V. Dolgosheina, S. C. Jeng, S. S. S. Panchapakesan, R. Cojocaru, P. S. Chen, P. D. Wilson, N. Hawkins, P. A. Wiggins, and P. J. Unrau. RNA Mango aptamer-fluorophore: A bright, high-affinity complex for RNA labeling and tracking. *ACS Chemical Biology*, 9(10):2412–2420, oct 2014.

[24] A. Dvir, R. C. Conaway, and J. W. Conaway. A role for TFIIH in controlling the activity of early RNA polymerase II elongation complexes. *Proceedings of the National Academy of Sciences of the United States of America*, 94(17):9006–10, aug 1997.

[25] S. Egloff and S. Murphy. Cracking the RNA polymerase II CTD code. *Trends in Genetics*, 24(6):280–288, jun 2008.

[26] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain. Stochastic gene expression in a single cell. *Science (New York, N.Y.)*, 297(5584):1183–6, aug 2002.

[27] J. Fischer and M. Wegener. Three-dimensional direct laser writing inspired by stimulated-emission-depletion microscopy. *Optical Materials Express*, 1(4):614, aug 2011.

[28] F. Fröhlich, C. Petit, N. Kory, R. Christiano, H. K. Hannibal-Bach, M. Graham, X. Liu, C. S. Ejsing, R. V. Farese, and T. C. Walther. The GARP complex is required for cellular sphingolipid homeostasis. *eLife*, 4(September), 2015.

[29] C. W. Gardiner. Handbook of stochastic methods for physics, chemistry and the natural sciences, vol. 13 of springer series in synergetics, 2004.

[30] D. G. Gibson, L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison, and H. O. Smith. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, 6(5):343–345, may 2009.

[31] I. Golding, J. Paulsson, S. M. Zawilski, and E. C. Cox. Real-time kinetics of gene activity in individual bacteria. *Cell*, 123(6):1025–1036, 2005.

[32] F. Groher and B. Suess. Synthetic riboswitches - A tool comes of age. *Biochimica et biophysica acta*, 1839(10):964–973, oct 2014.

[33] S. Hahn. Structure and mechanism of the RNA polymerase II transcription machinery. *Nature Structural and Molecular Biology*, 11(5):394–403, may 2004.

[34] F. Heigwer, G. Kerr, and M. Boutros. E-CRISP: Fast CRISPR target site identification, feb 2014.

[35] S. W. Hell and J. Wichmann. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Optics Letters*, 19(11):780, jun 1994.

[36] N. Hernandez. TBP, a universal eukaryotic transcription factor? *Genes & development*, 7(7B):1291–308, jul 1993.

[37] S. Hocine, P. Raymond, and D. Zenklusen. Single-molecule analysis of gene expression using two-color RNA labeling in live yeast. *Nature* . . . , 10(2), 2012.

[38] T. Høiby, H. Zhou, D. J. Mitsiou, and H. G. Stunnenberg. A facelift for the general transcription factor TFIIA, jul 2007.

[39] R. M. Hughes, S. Bolger, H. Tapadia, and C. L. Tucker. Light-mediated control of DNA transcription in yeast. *Methods (San Diego, Calif.)*, 58(4):385–91, dec 2012.

[40] T. Ideker, V. Thorsson, J. A. Ranish, R. Christmas, J. Buhler, J. K. Eng, R. Bumgarner, D. R. Goodlett, R. Aebersold, and L. Hood. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science*, 292(5518):929–934, may 2001.

[41] T. Jakociunas, A. S. Rajkumar, J. Zhang, D. Arsovska, A. Rodriguez, C. B. Jendresen, M. L. Skjoedt, A. T. Nielsen, I. Borodina, M. K. Jensen, and J. D. Keasling. CasEMBLR: Cas9-facilitated multi-loci genomic integration of in vivo assembled DNA parts in Saccharomyces cerevisiae. *ACS Synthetic Biology*, page 150317152158005, 2015.

[42] C. Jeronimo, A. R. Bataille, and F. Robert. The writers, readers, and functions of the RNA polymerase II C-terminal domain code, nov 2013.

[43] M. C. Jo, W. Liu, L. Gu, W. Dang, and L. Qin. High-throughput analysis of yeast replicative aging using a microfluidic system. *Proceedings of the National Academy of Sciences of the United States of America*, 112(30):9364–9, jul 2015.

[44] H. N. Joensson and H. Andersson Svahn. Droplet microfluidics-A tool for single-cell analysis, dec 2012.

[45] I. Jonkers, H. Kwak, and J. T. Lis. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *eLife*, 3(e02407):1–25, sep 2014.

[46] C. D. Kaplan. Basic mechanisms of RNA polymerase II activity and alteration of gene expression in Saccharomyces cerevisiae. *Biochimica et biophysica acta*, 1829(1):39–54, jan 2013.

[47] A. S. Karim, K. A. Curran, and H. S. Alper. Characterization of plasmid burden and copy number in Saccharomyces cerevisiae for optimization of metabolic engineering applications. *FEMS yeast research*, 13(1):107–16, feb 2013.

[48] M. J. Kennedy, R. M. Hughes, L. A. Peteya, J. W. Schwartz, M. D. Ehlers, and C. L. Tucker. Rapid blue light induction of protein interaction in living cells. *Nature Methods*, 7(12):973–975, 2010.

[49] M. Kim, N. J. Krogan, L. Vasiljeva, O. J. Rando, E. Nedea, J. F. Greenblatt, and S. Buratowski. The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature*, 432(7016):517–22, nov 2004.

[50] B. Klauser, J. Atanasov, L. K. Siewert, and J. S. Hartig. Ribozyme-Based Aminoglycoside Switches of Gene Expression Engineered by Genetic Selection in S. cerevisiae. *ACS Synthetic Biology*, 4(5):516–525, may 2015.

[51] R. D. Kornberg. The molecular basis of eukaryotic transcription. *Proc. Natl. Acad. Sci. U. S. A.*, 104(32):12955–12961, aug 2007.

[52] D. Kostrewa, M. E. Zeller, K. J. Armache, M. Seizl, K. Leike, M. Thomm, and P. Cramer. RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature*, 462(7271):323–330, nov 2009.

[53] A. Lamberti, S. L. Marasso, and M. Cocuzza. PDMS membranes with tunable gas permeability for microfluidic applications. *RSC Adv.*, 4(106):61415–61419, nov 2014.

[54] D. R. Larson, C. Fritzsch, L. Sun, X. Meng, D. S. Lawrence, and R. H. Singer. Direct observation of frequency modulated transcription in single cells using light activation. *eLife*, 2:e00750, jan 2013.

[55] D. R. Larson, D. Zenklusen, B. Wu, J. A. Chao, and R. H. Singer. Real-Time Observation of Transcription Initiation and Elongation on an Endogenous Yeast Gene. *Science*, 332(April):475–478, 2011.

[56] P. J. Lee, N. C. Helman, W. A. Lim, and P. J. Hung. A microfluidic system for dynamic yeast cell imaging. *BioTechniques*, 44(1):91–95, jan 2008.

[57] T. L. Lenstra, A. Coulon, C. C. Chow, and D. R. Larson. Single-Molecule Imaging Reveals a Switch between Spurious and Functional ncRNA Transcription. *Molecular Cell*, 60(4):597–610, nov 2015.

[58] F. Lim and D. S. Peabody. RNA recognition site of PP7 coat protein. *Nucleic acids research*, 30(19):4138–44, oct 2002.

[59] Y. Lorch and R. D. Kornberg. Chromatin-remodeling and the initiation of transcription. *Quarterly Reviews of Biophysics*, 48(04):465–470, nov 2015.

[60] X. Lu, X. Zhu, Y. Li, M. Liu, B. Yu, Y. Wang, M. Rao, H. Yang, K. Zhou, Y. Wang, Y. Chen, M. Chen, S. Zhuang, L. F. Chen, R. Liu, and R. Chen. Multiple P-TEFbs cooperatively regulate the release of promoter-proximally paused RNA polymerase II. *Nucleic Acids Research*, 44(14):6853–6867, aug 2016.

[61] Z. Luo, C. Lin, and A. Shilatifard. The super elongation complex (SEC) family in transcriptional control, sep 2012.

[62] M. Manosas, X. G. Xi, D. Bensimon, and V. Croquette. Active and passive mechanisms of helicases. *Nucleic Acids Research*, 38(16):5518–5526, sep 2010.

[63] L. Mazutis, J. C. Baret, P. Treacy, Y. Skhiri, A. F. Araghi, M. Ryckelynck, V. Taly, and A. D. Griffiths. Multi-step microfluidic droplet processing: Kinetic analysis of an in vitro translated enzyme. *Lab on a Chip*, 9(20):2902–2908, oct 2009.

[64] R. S. McIsaac, S. J. Silverman, M. N. McClean, P. a. Gibney, J. Macinskas, M. J. Hickman, a. a. Petti, and D. Botstein. Fast-acting and nearly gratuitous induction

of gene expression and protein depletion in Saccharomyces cerevisiae. *Molecular Biology of the Cell*, 22(22):4447–4459, 2011.

[65] T. K. Meyvis, S. C. De Smedt, P. Van Oostveldt, and J. Demeester. Fluorescence recovery after photobleaching: A versatile tool for mobility and interaction measurements in pharmaceutical research, aug 1999.

[66] T. Muramoto, D. Cannon, M. Gierlinski, A. M. Corrigan, G. J. Barton, J. R. Chubb, D. Cannona, and M. Gierli\'nski. Live imaging of nascent RNA dynamics reveals distinct types of transcriptional pulse regulation. *Proceedings of the National Academy of Sciences of the United States of America*, 109(19):7350–7355, may 2012.

[67] T. Okamoto, S. Yamamoto, Y. Watanabe, T. Ohta, F. Hanaoka, R. G. Roeder, and Y. Ohkuma. Analysis of the role of TFIIE in transcriptional regulation through structure-function studies of the TFIIE?? subunit. *Journal of Biological Chemistry*, 273(31):19866–19876, jul 1998.

[68] R. C. Olsthoorn, G. Garde, T. Dayhuff, J. F. Atkins, and J. Van Duin. Nucleotide sequence of a single-stranded RNA phage from Pseudomonas aeruginosa: kinship to coliphages and conservation of regulatory RNA structures. *Virology*, 206(1):611–25, jan 1995.

[69] G. Orphanides, T. Lagrange, and D. Reinberg. The general transcription factors of RNA polymerase II. *Genes & development*, 10(21):2657–83, nov 1996.

[70] J. S. Paige, K. Y. Wu, and S. R. Jaffrey. RNA mimics of green fluorescent protein. *Science (New York, N.Y.)*, 333(6042):642–6, jul 2011.

[71] M. Pal, A. S. Ponticelli, and D. S. Luse. The role of the transcription bubble and TFIIB in promoter clearance by RNA polymerase II. *Molecular Cell*, 19(1):101–110, jul 2005.

[72] J. Paulsson. Models of stochastic gene expression, 2005.

[73] Z. C. Poss, C. C. Ebmeier, and D. J. Taatjes. Critical Reviews in Biochemistry and Molecular Biology The Mediator complex and transcription regulation The Mediator complex and transcription regulation. *Critical Review in Biochemistry and Molecular Biology*, 48(6):1549–7798, nov 2013.

[74] P. J. Robinson, M. J. Trnka, D. A. Bushnell, R. E. Davis, P. J. Mattei, A. L. Burlingame, and R. D. Kornberg. Structure of a Complete Mediator-RNA Polymerase II Pre-Initiation Complex. *Cell*, 166(6):1411–1422.e16, sep 2016.

[75] J. Russell and J. C. B. M. Zomerdijk. The RNA polymerase I transcription machinery. *Biochemical Society symposium*, 1(73):203–216, 2006.

[76] J. Ryley and O. M. Pereira-Smith. Microfluidics device for single cell gene expression analysis in Saccharomyces cerevisiae. *Yeast*, 23(14-15):1065–1073, oct 2006.

[77] J. SantaLucia. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proceedings of the National Academy of Sciences*, 95(4):1460–1465, feb 1998.

[78] C. Schneider, L. Bronstein, J. Diemer, H. Koeppl, and B. Suess. ROC'n'Ribo: Characterizing a Riboswitching Expression System by Modeling Single-Cell Data. *ACS Synthetic Biology*, 6(7):1211–1224, jul 2017.

[79] L. Schramm and N. Hernandez. Recruitment of RNA polymerase III to its target promoters, oct 2002.

[80] I. Schroeder, S. Gazzarrini, G. Ferrara, G. Thiel, U. P. Hansen, and A. Moroni. Creation of a reactive oxygen species-insensitive Kcv channel. *Biochemistry*, 52(18):3130–3137, may 2013.

[81] D. Schübeler. Function and information content of DNA methylation, jan 2015.

[82] A. Selimis, V. Mironov, and M. Farsari. Direct laser writing: Principles and materials for scaffold 3D printing. *Microelectronic Engineering*, 132:83–89, jan 2014.

[83] A. Senecal, B. Munsky, F. Proux, N. Ly, F. E. Braye, C. Zimmer, F. Mueller, and X. Darzacq. Transcription factors modulate c-Fos transcriptional bursts. *Cell Reports*, 8(1):75–83, jul 2014.

[84] M. E. Sherlock, N. Sudarsan, S. Stav, and R. R. Breaker. Tandem riboswitches form a natural Boolean logic gate to control purine metabolism in bacteria. *eLife*, 7:e33908, mar 2018.

[85] O. Shimomura, F. H. Johnson, and Y. Saiga. Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from the Luminous Hydromedusan,Aequorea. *Journal of Cellular and Comparative Physiology*, 59(3):223–239, jun 1962.

[86] S. T. Smale and J. T. Kadonaga. The RNA polymerase II core promoter. *Annual review of biochemistry*, 72(1):449–79, jun 2003.

[87] J. Soutourina. Transcription regulation by the Mediator complex, dec 2018.

[88] M. Sunbul and A. Jäschke. SRB-2: a promiscuous rainbow aptamer for live-cell RNA imaging. *Nucleic Acids Research*, jun 2018.

[89] S. Tan, Y. Hunziker, D. F. Sargent, and T. J. Richmond. Crystal structure of a yeast TFIIA/TBP/DNA complex. *Nature*, 381(6578):127–151, may 1996.

[90] K. Tantale, F. Mueller, A. Kozulic-Pirher, A. Lesne, J.-M. Victor, M.-C. Robert, S. Capozi, R. Chouaib, V. Bäcker, J. Mateos-Langerak, X. Darzacq, C. Zimmer, E. Basyuk, and E. Bertrand. A single-molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting. *Nature communications*, 7:12248, jul 2016.

[91] A. Teixeira, A. Tahiri-Alaoui, S. West, B. Thomas, A. Ramadass, I. Martianov, M. Dye, W. James, N. J. Proudfoot, and A. Akoulitchev. Autocatalytic RNA cleavage in the human beta-globin pre-mRNA promotes transcription termination. *Nature*, 432(November):526–530, nov 2004.

[92] T. O. Tollefsbol. *Handbook of Epigenetics*. Elsevier Inc., 2011.

[93] D. Tollervey. Termination by torpedo. *Nature*, 432(7016):456–457, nov 2004.

[94] E. Tutucci, M. Vera, J. Biswas, J. Garcia, R. Parker, and R. H. Singer. An improved MS2 system for accurate reporting of the mRNA life cycle. *Nature Methods*, 15(1):81–89, nov 2017.

[95] M. Unger. Interrogating the single cell: Computational and experimental methods for optimal live cell experiments, 2014.

[96] M. Valencia-Burton, R. M. McCullough, C. R. Cantor, and N. E. Broude. RNA visualization in live bacterial cells using fluorescent protein complementation. *Nature methods*, 4(5):421–7, may 2007.

[97] A. Vannini, R. Ringel, A. G. Kusser, O. Berninghausen, G. A. Kassavetis, and P. Cramer. Molecular basis of RNA polymerase III transcription repression by Maf1. *Cell*, 143(1):59–70, oct 2010.

[98] A. Veloso, K. S. Kirkconnell, B. Magnuson, B. Biewen, M. T. Paulsen, T. E. Wilson, and M. Ljungman. Rate of elongation by RNA polymerase II is associated with specific gene features and epigenetic modifications. *Genome Research*, 24(6):896–905, jun 2014.

[99] J. Venema and D. Tollervey. Ribosome Synthesis in Saccharomyces cerevisiae. *Annual Review of Genetics*, 33(1):261–311, dec 1999.

[100] J. D. WATSON and F. H. CRICK. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, 171(4356):737–8, apr 1953.

[101] F. Winston, C. Dollard, and S. L. Ricupero, Hovasse. Construction of a set of convenient saccharomyces cerevisiae strains that are isogenic to S288C. *Yeast*, 11(1):53–55, jan 1995.

[102] B. Wu, J. Chen, and R. H. Singer. Background free imaging of single mRNAs in live cells using split fluorescent proteins. *Scientific reports*, 4:3615, jan 2014.

[103] B. Wu, V. Miskolci, H. Sato, E. Tutucci, C. A. Kenworthy, S. K. Donnelly, Y. J. Yoon, D. Cox, R. H. Singer, and L. Hodgson. Synonymous modification results in highfidelity gene expression of repetitive protein and nucleotide sequences. *Genes and Development*, 29(8):876–886, apr 2015.

[104] Y. Xu, C. Kanagaratham, and D. Radzioch. Chromatin Remodelling During Host-Bacterial Pathogen Interaction. *Chromatin Remodeling*, pages 173–198, apr 2013.

[105] Q. Yan, R. J. Moreland, J. W. Conaway, and R. C. Conaway. Dual roles for transcription factor IIF in promoter escape by RNA polymerase II. *Journal of Biological Chemistry*, 274(50):35668–35675, dec 1999.

[106] J. Yao, M. B. Ardehali, C. J. Fecko, W. W. Webb, and J. T. Lis. Intranuclear Distribution and Local Dynamics of RNA Polymerase II during Transcription Activation. *Molecular Cell*, 28(6):978–990, dec 2007.

[107] R. A. Young. RNA Polymerase II. *Annual Review of Biochemistry*, 60(1):689–715, jun 1991.

[108] C. Zechner, M. Unger, S. Pelet, M. Peter, and H. Koeppl. Scalable inference of heterogeneous reaction kinetics from pooled single-cell recordings. *Nature methods*, 11(2):197–202, feb 2014.