UNIVERSITY OF COPENHAGEN
FACULTY OF SCIENCE

PhD Thesis

Kasper Graves Hvid

Experimental single molecule fluorescence microscopy for investigation of embryonic stem cell priming and transcription regulation

Supervisor: Poul Martin Bendix
Co-supervisor: Lene Broeng Oddershede
Co-supervisor: Joshua Mark Brickman

This thesis has been submitted to the PhD School of The Faculty of Science, University of Copenhagen on December 31, 2019

Abstract

Embryonic development is both fascinatingly complex yet highly robust in replicating morphology and expression to a precise timing. The underlying mechanisms and signaling pathways are diverse and many remain undiscovered, yet with advancements in fluorescence imaging technology, we are now able to observe nanoscale structures and follow individual molecules inside living embryonic stem cells. This has the potential to give us a highly detailed look at how sub-cellular mechanisms are involved in making cell-fate choices. In this thesis, I will describe how we have implemented single molecule fluorescence microscopy to track and measure how individual core pluripotency transcription factor binds to chromatin inside living cells. Next, we implemented stochastic blinking of individual fluorophores to surpass the diffraction barrier and image the actin cytoskeleton of naive stem cells at nanoscale resolution. Finally, I show how we are able to localize and quantify active sites of transcription by hybridizing fluorescent probes to single molecules of mRNA.

Acknowledgments

I would like to thank my supervisors Poul Martin Bendix and Lene Oddershede from the Niels Bohr Institute for all their help, guidance, and inspiration during this PhD project. I also wish to thank them for their patience and encouragement in allowing me to freely explore areas and fields of interest within not just the field of physics, but also molecular and cellular biology.

I also wish to thank my co-supervisor Josh Brickman from the Novo Nordisk Foundation Center for Stem Cell Biology (DanStem), for his help and guidance, and for introducing me to the forefront of the fields of stem cell and developmental biology research. His expertise in this field helped bring both this research but also my understanding to a much higher level of relevance in a biological context.

Thanks to Timothee Lionnet and Wang Shuoshuo from the Lionnet Lab at New York University for hosting me and teaching me mRNA FISH, and for discussions and inspiration for single molecule imaging.

Thanks to everyone from the Optical Tweezers lab at the Niels Bohr Institute, in particular Younes Farhangi Barooji for our collaboration on building the super resolution platform, and for not just breaking the resolution limit, but endlessly improving protocols and imaging, and being an invaluable source of help and expertise.

From the Brickman lab at DanStem, I also wish to thank everyone for fruitful discussions, for adopting me in their lab, and particularly for their patience and willingness to explain complex biological concepts and questions to the clueless physicist dumping into their lab without much notice.

A special thanks to Karolina Ditrychova for our close collaboration on much of this research, for teaching me so many things I never knew I didn't know about biology and stem cells. At the outset of this project and collaboration, nobody would have foreseen how close we would become, but I am endlessly happy and grateful for it.

Summary

Murine embryonic stem cells (ESC) are cultured cells derived from the early mouse embryo, which retain the fundamental properties of *pluripotency* and *self-renewal*. Potency describes the ability of a cell to differentiate and give rise to distinct cell fates, and pluripotent cells are able to develop into any cell of the adult body. Self-renewal describes the ability to divide and propagate indefinitely. ESC are cultured in various conditions to recapitulate different stages and cell types of early embryonic development, and can be stimulated to initiate differentiation or to maintain a state of naive pluripotency. Stem cells provide a powerful platform to investigate how cells make lineage fate choices, and what changes cells undergo when they make these decisions.

This study is a biophysical investigation of ESC through fluorescence microscopy. The scope involves both the dynamics of individual molecules in living stem cells, and actin cytoskeleton organization with nano-scale resolution. In the first part of the project, a fluorescence microscopy platform was constructed, implementing tilted illumination modes, fluorophore photo-activation, and highly sensitive detection for single molecule imaging in stem cells. In the second part of the project, single molecule imaging was applied to track individual molecules of the pluripotency transcription factor Sox2 in living ESC. We were able to detect and track single molecules as they bound to chromatin, and discovered a sub-population of highly stable binding events, which previously published models of transcription factor binding do not account for. We also found that while Sox2 binding was not affected by the priming of stem cells towards differentiation, the stable binding events were reduced by inhibition of the transcription co-factor BRD4, suggesting a relationship with liquid-phase condensates around super-enhancers.

Next, we induced the stochastic blinking of fluorophores on fixed structures in order to localize and map each individual molecule with high precision, and in this way overcome the diffraction resolution limit of light microscopy. By linking fluorophores to actin filaments in ESC, we determined the shape and structure of the actin cytoskeleton with nanoscale resolution. Comparing the actin fine-structure in two common conditions of ESC culture, we observed a disordered mesh of actin running along the cortex in more naive pluripotent cells, minimizing their surface attachment, while less naive primed cells revealed a tendency to spread along the surface, with actin organized in bundles of stress fibers attaching at the surface towards cell periphery.

In conclusion, the application of single molecule fluorescence microscopy to the study of ESC allowed new detailed insight into the underlying molecular mechanisms of both transcription and cytoskeleton organization. While the biological scope of this work was limited to just a single transcription factor, and one component of the cytoskeleton, the methods developed and described can easily be applied to other proteins. This study therefore also lays a foundation for future work using single molecule fluorescence on alternative models.

Resumé på dansk

Embryonale stamceller (ESC) er kultiverede celler etableret fra den indre cellemasse af et tidligt udviklingsstadie af mammale embryoner, i denne afhandling fra mus. Stamceller er udifferentierede celler, som kan udvikle sig til forskellige specialiserede celletyper, og ESC er såkaldt pluripotente celler som kan specialisere sig til alle kroppens celletyper. Udover evnen til at differentiere sig, har stamceller også egenskaben selvfornyelse, hvilket betyder at stamceller kan dele og kopiere sig selv uafbrudt.

ESC kan kultiveres under forskellige vilkår, som hver afspejler forskellige cellestadier i den tidlige embryonale udvikling. Ved at ændre kultiveringsbetingelserne fx ved at tilsætte forskellige inhibitorer eller ved genmodificering, kan ESC manipuleres til at differentiere sig i en bestemt retning eller forblive i deres pluripotente stadie. På denne måde udgør ESC et vigtigt og effektivt redskab til at forstå både embryonal udvikling, men også hvordan og hvornår de enkelte celler bestemmer sig for at differentiere sig, samt hvilke ændringer dette medfører i cellerne.

I denne afhandling anvendes biofysiske metoder og avanceret fluorescens mikroskopi til at undersøge transskriptionsfaktorer og cytoskelettet i embryonale stamceller på nanoskala. Den første del af projektet omhandler opbygningen af et fluorescens mikroskop, der implementerer vinklet laser-illumination, fotoaktivering af fluoroforer, og et følsomt kamera til detektion af enkelte fluorescerende molekyler i stamceller. I den anden del af projektet anvendtes enkelt-molekyle mikroskopi til at følge enkelte molekyler af den pluripotens-associerede transskriptionsfaktor Sox2 i levende ESC. Vi var i stand til at detektere og følge enkelte Sox2 molekyler, samt måle deres binding til kromatin. Vi observerede en population af meget stabile interaktioner, som ikke kunne forklares ud fra tidligere beskrevne modeller. Vi fandt også at Sox2 bindinger til DNA ikke var påvirket af stimulering mod stamcelle differentiering. Derimod så vi en reduktion i stabile bindinger ved inhibering af transskriptions cofaktoren BRD4, hvilket antyder en sammenhæng med væske-fase kondensater omkring super-enhancere.

I den næste del inducerede vi stokastisk blinken i fluoroforer bundet til actin, med det formål at kortlægge de individuelle molekyler position med høj præcision, og på denne måde overkomme den teoretiske grænse for lysmikroskopets opløsning. På denne måde beskrev vi på nanoskala cytoskelettet i ESC under to forskellige kultiveringsbetingelser. Ved at sammenligne finstrukturen i de to typer stamceller, så vi at mere naive og pluripotente celler udviste en tendens til at minimere deres fastgørelse til underlaget med actin orienteret parallelt med overfladen, mens mindre naive celler spredte sig ud med actin samlet i bundter vinklet mod underlaget.

Konklusionen er, at anvendelsen af enkelt-molekyle fluorescens mikroskopi til studiet af ESC har muliggjort ny detaljeret indsigt i de underlæggende molekylære mekanismer bag både transskription og cytoskelettets organisering. Mens dette studies fokus var begrænset til en enkelt transskriptionsfaktor og en enkelt komponent af cytoskelettet, er de udviklede og beskrevne metoder nemt overført til andre proteiner. Dette studie fungerer derfor også som en grundsten for videre og opfølgende studier med enkelt-molekyle fluorescens i alternative modeller.

Contents

List of Abbreviations

The following is a list of abbreviations that will be used in the main body of the report.

AIC	Akaike information criterion
AICc	Sample size corrected AIC
AOTF	Acousto-optic tunable filter
BRD4	Bromodomain-containing protein 4
CDF	Cumulative distribution function
CHIR	GSK3 inhibitor CHIR99021
DNA	Deoxyribonucleic acid
EMCCD	Electron-multiplying charge-coupled device
ERK	Extracellular signal-regulated kinase
ESC	Embryonic stem cell
ETOH	Ethanol
FGF	Fibroblast growth factor
FGFR	FGF receptor
FISH	Fluorescence in situ hybridization
HILO	Highly inclined and laminated optical sheet
HT	HaloTag
ICM	Inner cell mass
JF549/646	Janelia Fluor 549/646
LIF	Leukemia inhibitory factor
MED24	Mediator 24
MEK	MAPK/ERK Kinase

mRNA	Messenger RNA
NA	Numerical aperture
NLS	Nuclear localization sequence
OHT/4-OHT	(Z)-4-Hydroxytamoxifen
p300	Histone acetyltransferase p300
p300i	p300 inhibitor
PA-JF646	Photo-activatable Janelia Fluor 646
PALM	Photo-activated localization microscopy
PD03	MEK inhibitor PD0325901
PD17	FGFR inhibitor PD173074
pERK	Phosphorylated ERK
PIC	Transcription pre-initiation complex
PSF	Point spread function
RNA	Ribonucleic acid
RNAP	RNA polymerase II
SMLM	Single molecule localization microscopy
Sox2	Sex-determining region Y-box 2
SPT	Single particle tracking
STORM	Stochastic optical reconstruction microscopy
TF	Transcription factor
TIRF	Total internal reflection fluorescence

Objectives and overview of the thesis

This PhD project was structured according to the University of Copenhagen integrated PhD and MSc programme; the first two years of the PhD project was done concurrently with the MSc studies, and the results of this first part of the work was published in a thesis functioning as a qualifying exam for the final two years of the PhD. Therefore, this thesis will mainly focus on the later part of the PhD work, only giving a brief description of the work already published in the MSc thesis, while referring to it [1].

The objectives of the PhD project were the following:

Firstly, to construct a super-resolution fluorescence microscopy platform suitable for single molecule imaging of both fixed and live stem cells.

The second objective was to implement methods of both super resolution imaging and single molecule tracking using the constructed microscope, and apply these methods to the investigation of embryonic stem cells.

The third objective was to image and track individual transcription factor binding of a core pluripotency factor on the single molecule level, and quantify this behavior in different conditions and in naive or primed stem cells.

The final objective was to take advantage of the super-resolution methods to map the nanoscale actin cytoskeleton structure in stem cells, and compare sub-cellular structural differences between stem cells in states of either naive pluripotency or primed pluripotency.

The thesis is structured as a series of chapters.

The first chapter is intended as an introduction to core concepts and literature in the fields of stem cell biology and super resolution fluorescence imaging.

The second chapter contains a brief summary of the first objective of the thesis; the construction of a super-resolution setup, already described in the MSc thesis.

The third chapter describes the implementation of the setup for single-molecule tracking in embryonic stem cells, and is written in the style of a manuscript, which is in preparation for submission.

The fourth chapter describes implementing super-resolution imaging to study nanoscale actin structure in two defined conditions of stem cell culture. This work is also being prepared as a manuscript for submission.

The fifth chapter contains preliminary work with single-molecule mRNA fluorescence in situ hybridization, which is intended to be implemented both as a quantification of allele transcription activity, but also to be combined with transcription factor tracking, in order to relate dynamics to specific genetic loci.

Chapter 1

Introduction

1.1 The developing pre-implantation embryo

Early embryonic development starts with fertilization of the egg resulting in a single cell zygote stage. As the cell undergoes a series of divisions, the embryo initially consists of indistinguishable and totipotent cells, which can give rise to any and all cells of the later embryo. This stage we call morula stage (around day 2.5 on figure 1.1). The first cell lineage segregation occurs shortly after, where cells facing outwards differentiate into trophectoderm (in green, figure 1.1), while cells inside form the inner cell mass (in purple). This stage is called blastocyst, and trophectoderm of this stage will eventually contribute to the placenta, while the ICM will give rise to all embryonic and extraembryonic tissue.

As blastocyst grows, formation of blastocyst cavity appears inside the trophectoderm. Between day 3.5 and 4.5 of embryonic development, a second lineage division occurs, where cells of ICM decide between epiblast and primitive endoderm (red and blue on figure 1.1 respectively).

During the first days of pre-implantation embryonic development, as cells start to segregate and differentiate into separate cell fates, a remarkable and drastic divergence appears in the expression pattern of core transcription factors that eventually become exclusive to specific cell fates. The first cell line segregation event of the trophectoderm and ICM formation is clearly visualized by marking the differential expression of the transcription factors Nanog and Cdx2 (figure 1.2). Prior to the segregation event, all cells of the morula express a heterogeneous and overlapping mix of the two transcription factors. However, as the trophectoderm forms, Cdx2 becomes exclusively expressed in the trophectoderm, while Nanog is

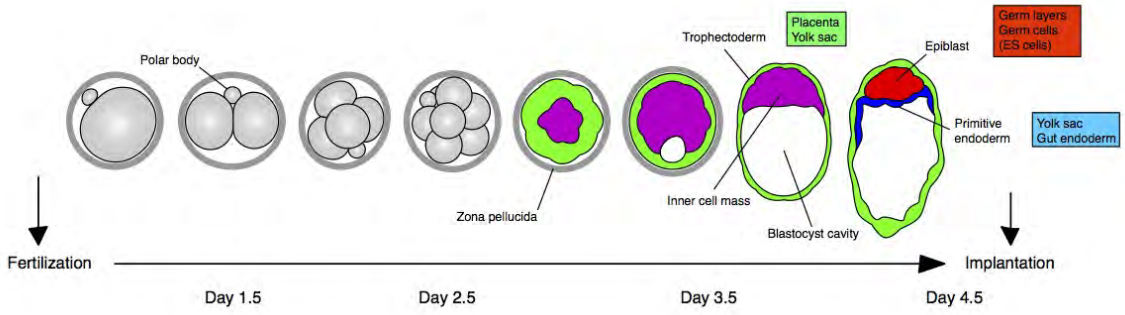

Figure 1.1: Figure from Saiz (2014)[2].

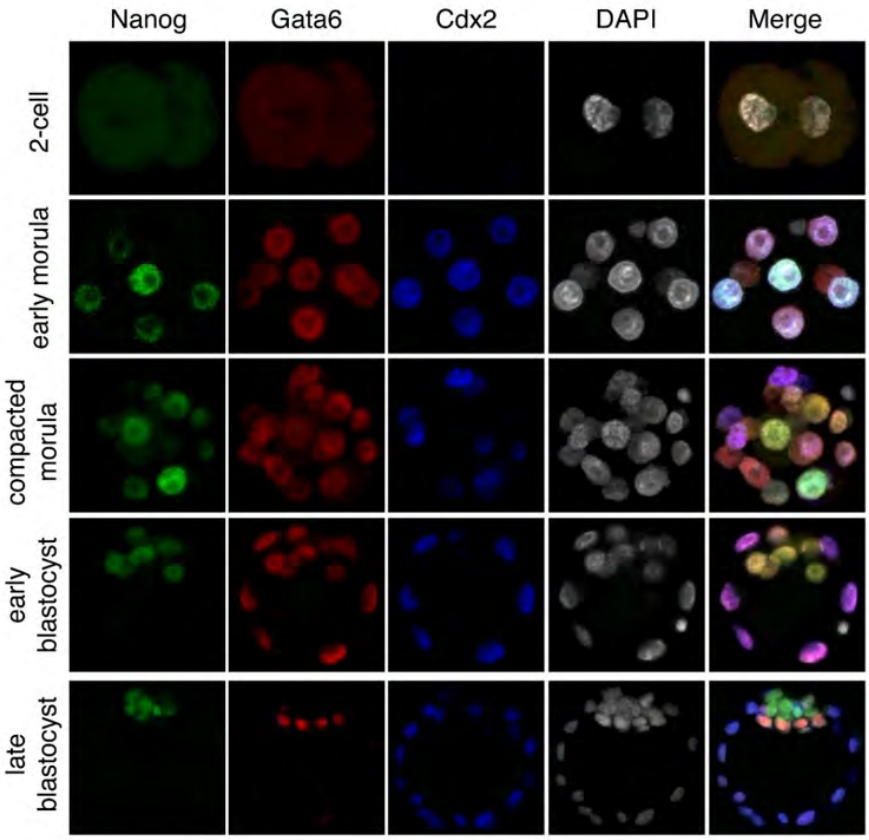

Figure 1.2: Heterogeneous expression of transcription factors in embryonic development separate cell fates. Before the blastocyst stage of development, transcription factors Nanog, Gata6, and Cdx2 are expressed heterogeneously but without any clear pattern. In the early blastocyst, segregation of Cdx2 (a trophectoderm marker) from Nanog (an epiblast and pluripotency marker) begins to appear, with Nanog concentrating in the inner cell mass (ICM). While cells in the ICM in the early blastocyst stage show co-expression of Nanog and Gata6 (a primitive endoderm marker), this changes in the late blastocyst, where cells differentiate into either Nanog-expressing epiblast or Gata6-expressing primitive endoderm. Figure from Morgani and Brickman (2015)[5].

exclusively expressed in the ICM.

As the blastocyst matures, the second cell lineage segregation event of the ICM is similarly identified by the divergence of transcription factor expression, this time of Nanog and Gata6. The initial co-expression is disrupted and segregated into a randomly distributed "salt-and-pepper" pattern of mutually exclusive Nanog or Gata6 expressing cells, forming precursors (or cells primed towards) the epiblast and primitive endoderm respectively. This is then followed by sorting of the cells into an outer primitive endoderm and inner epiblast layer, through a combination of cell movement, apoptosis and amplification of primed cell identity [3, 4]. By the late blastocyst stage, the pre-implantation embryo has matured into three completely segregated cell lineages; trophectoderm, epiblast, and primitive endoderm, as identified by their exclusive lineage markers; Cdx2, Nanog, and Gata6 respectively (figure 1.2).

The transition from homogeneous co-expressing cells in the ICM of the early blastocyst, to the segregated and exclusively Nanog or Gata6 expressing cells of the epiblast or primitive endoderm respectively in the late blastocyst stage, has been shown to critically depend on the FGF/ERK intercellular signaling pathway, involving fibroblast growth factor 4 (FGF) and extracellular signal-regulated kinase (ERK) [3, 4, 6–11].

Inhibition of FGF/ERK signaling prevents the formation of salt-and-pepper pattern in embryos, and suppresses the differentiation of pluripotent naive epiblast cells to primitive endoderm lineages [3, 6, 11, 12]. On the other hand, over-stimulation of the pathway with exogenous FGF drives the ICM to form a disproportionate amount of primitive endoderm cells at the cost of epiblast [3, 6, 12].

FGF/ERK signaling combines extracellular signals through the secretion and receptor binding of FGF, with intracellular regulation as an activated FGF receptor (FGFR) initiates a cascade of phosphorylations culminating in phosphorylated ERK entering the nucleus and regulating the transcription of core pluripotency transcription factors (figures 1.3 and 1.4)[9, 11, 13, 14]. Phosphorylated ERK directly represses Nanog expression, while Nanog and Gata6 are mutually repressing transcription factors. This explains how over-stimulation of the FGF/ERK pathway drives ICM to repress Nanog, and consequently up-regulate Gata6, pushing the balance towards primitive endoderm cell fates.

Furthermore, Nanog stimulates the secretion of FGF, while Gata6 induces FGFR expression. As such, cells already high in Gata6 have an increased sensitivity to FGF/ERK signaling, and have a propensity towards further primitive endoderm commitment (figure 1.3). We call this state the "primed" state of stem cells, as they are prone to differentiate towards a certain fate, however they are not yet fully committed. While primed stem cells have a higher propensity towards a certain cell fate, they still have the potency to differentiate towards other cell fates or remain pluripotent. In this way, through the emergence of stem cells primed towards either epiblast or primitive endoderm fates, small differences in expression and plasticity between the two cell types are amplified and propagated through FGF/ERK signaling. This differential sensitivity to the signaling is necessary for the segregation of primitive endoderm and epiblast [4, 15].

The reliance on FGF/ERK to exit the naive epiblast stage can be exploited for *in vitro* culture of embryonic stem cells (ESC); inhibition of the pathway by a MEK inhibitor (PD03, figure 1.4) in combination with a GSK3 inhibitor (CHIR) has been shown to be sufficient to maintain cultured ESC in a homogeneous pluripotent state resembling the naïve early epiblast cells (2i media, [7, 16]).

We can recapitulate and investigate single-cell responses to FGF/ERK signaling in cultured ESC, for example by directly activating MEK as in Hamilton and Brickman (2014)[14] (figure 1.4), inducing primitive endoderm priming and differentiation.

Hamilton and Brickman showed that the activation of ERK in cultured ESC quickly represses core pluripotency factors and epiblast markers, including Nanog, which in turn allow for primitive endoderm markers to be expressed. However, while cells were pushed towards the primitive endoderm lineage and showed identical characteristics of primitive endoderm primed cells, they required prolonged ERK signal to fully commit. Indeed within the first 8 hours, priming could still be reversed by inhibiting ERK signaling, and cells would regain expression of naive cells.

Recently Hamilton *et al.* (2019)[17] furthermore showed that this priming of ESC towards primitive endoderm was closely related to transcription factor occupancy on core pluripotency genes. As ERK activates immediate early response genes as its canonical targets, it also specifically decommissions genes associated with pluripotency. ERK signaling quickly and directly induces the repression of these genes, through the dissociation of RNA polymerase II (RNAP) and other transcription pre-initiation complex (PIC) components, from their promoter sequences. Data from Hamilton et al. suggest, that ERK phosphorylates specific components of PIC and suggests Mediator 24 (MED24), as the main target giving the specificity to the pluripotency enhancers [17].

Interestingly, it was also found, that other core factors necessary for transcription at these genes, namely specific transcription factors (TFs) Sox2 and Essrb, remained bound and present at both the promoter and enhancers. The fact that in particular Sox2, but also other TFs', protein stability is not affected by ERK, but their transcription is, may suggest protein lifetime as a possible mechanism determining the transition from priming to commitment. The final commitment

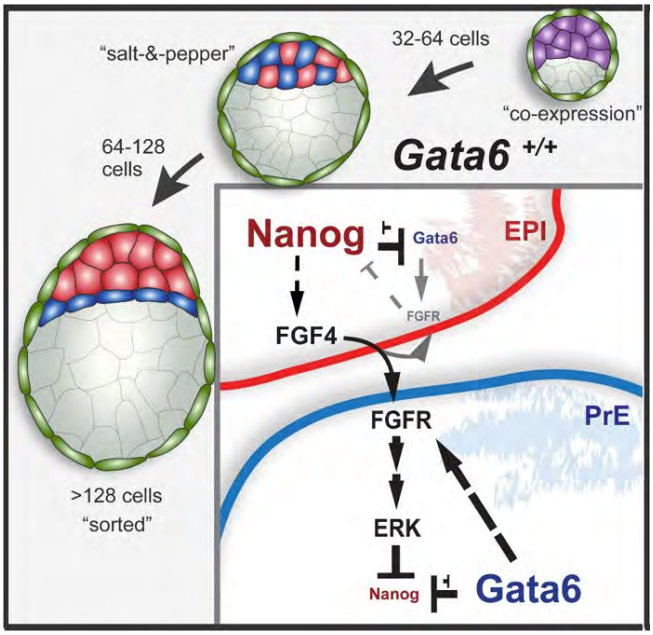

Figure 1.3: Mutual repression between Nanog and Gata6 leads to differential receptivity to extracellular fibroblast growth factor 4 (FGF) signaling. Epiblast cells (EPI) have high levels of Nanog, which promotes the expression and excretion FGF, while repressed levels of Gata6 and in turn FGF receptors (FGFR) limits its sensitivity to extracellular FGF signal. In Gata6 high cells of the primitive endoderm (PrE), the repression of Nanog is further enhanced by sensitivity to FGF signaling which activates the FGF/ERK signaling cascade, which represses the expression of Nanog and other pluripotency factors. Figure modified from Schrode *et al.* (2014)[11].

of primed stem cells to a lineage may only be after degradation of the TFs at the gene [14, 17].

Moreover, the unchanged stability of TF occupancy even after the dissociation of other components of PIC, raises interesting questions about TF binding to DNA. Is TF binding stability independent of cooperativity with PIC components, or does it rely on other mechanisms such as chromatin conformation, liquid phase condensation, or potentially on other undiscovered PIC factors?

Overall, this system recapitulates the mechanisms of stem cell priming and differentiation, and is a powerful tool to investigate cell fate choice and mechanisms of transcriptional activation and repression, with precise temporal control.

1.2 Embryonic stem cells

Stem cells are characterized by two defining properties; they have the capability to renew and proliferate, and they have the potential to differentiate or specialize into various cell types. The ability to indefinitely proliferate is called self-renewal, while we define the potential to differentiate into distinct cell fates potency, for example cell line capable of differentiation into 2 cell lines would be bipotent.

Murine embryonic stem cells (ESC) are karyotypically normal cells capable of indefinite propagation in cell culture conditions, retaining the feature of self-renewal. Meanwhile, ESC are also pluripotent, which means they can be differentiated to most of the cell lineages of the embryo. As they are derived from the inner cell mass (ICM) of the blastocyst, they can be differentiated into both epiblast and primitive endoderm cell lineages, that the ICM give rise to [18–20]. While primitive endoderm is cell line supporting embryonic development and contributes to placenta formation, epiblast gives rise to embryo proper in later embryonic development, and ESC can thus be used as a model of differentiation into various

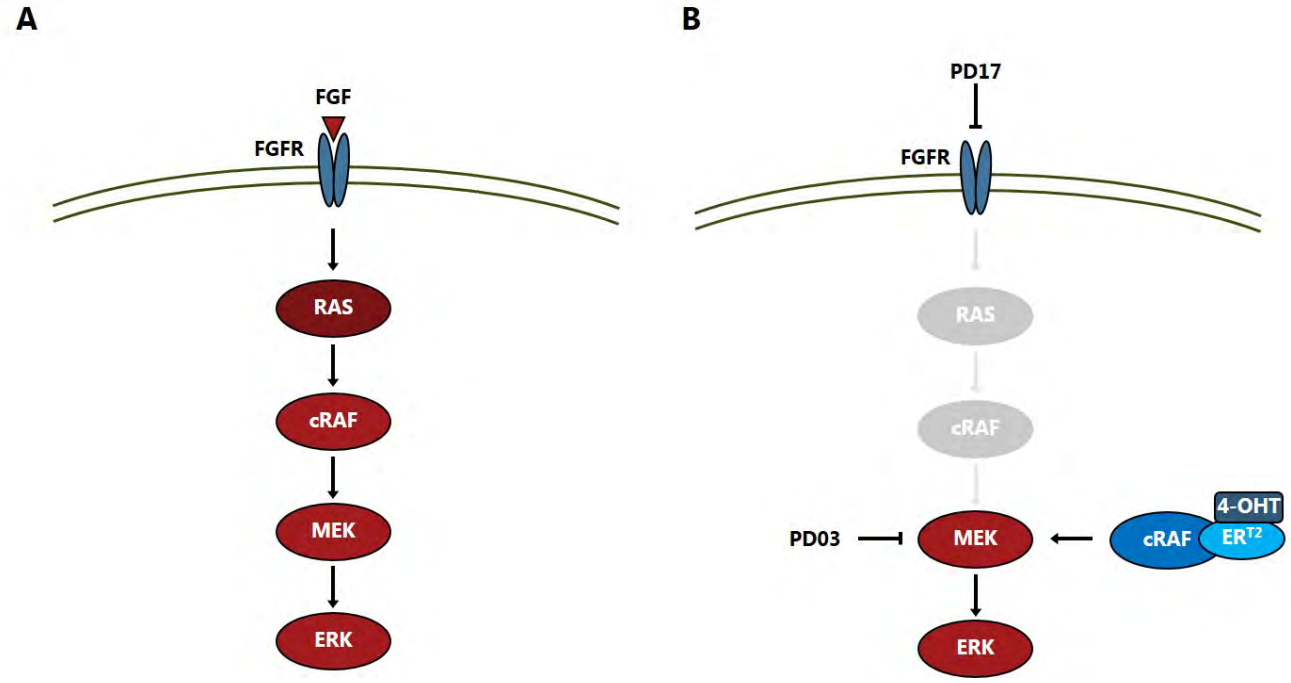

Figure 1.4: Simplified schematic of the FGF/ERK signaling cascade including the direct ERK activation system used in Hamilton and Brickman (2014)[14], and Hamilton *et al.* [17]. **(A)** Fibroblast growth factor 4 (FGF) binds and activates FGF receptor (FGFR), which initiates a signaling cascade starting with RAS, and continuing through the protein kinases cRAF, MEK, and ERK. Protein kinases are enzymes which phosphorylate their targets, extracellular signal-regulated kinase ERK in its phosphorylated, active form will enter the cell nucleus and regulate transcription of core pluripotency factors including Nanog [14]. ERK itself is in turn phosphorylated by the protein kinase MEK (MAPK/ERK kinase). Finally cRAF is responsible for phosphorylating MEK. **(B)** In the direct ERK activation system, FGFR is blocked by the FGFR inhibitor PD17, which stops the FGF/ERK signaling cascade from initiating. Instead, ERK activation is regulated by the introduction of a constitutively expressed construct containing cRAF fused to a ligand binding domain of estrogen receptor (ER^{T2}) which binds tamoxifen (4-OHT). In this way, the ERK signaling pathway can directly be activated by introducing 4-OHT to bind the cRAF construct. Reversely, the signaling pathway can be inhibited by introduction of the MEK inhibitor PD03.

cell lines, for example neural stem cells. Together, self/renewal and pluripotency make ESC a very powerful tool to study early embryonic development, allowing the rapid expansion of many cells for investigation, while recapitulating important features from the developing embryo.

While we can keep stem cells growing at the same stage of development by the inhibition or activation of certain signaling pathways, by similar mechanisms it is also possible to initiate differentiation by changing media conditions. Early ESC culturing conditions relied on growing cells on mouse embryonic fibroblasts, until leukemia inhibitory factor (LIF) was identified as the crucial secreted cytokine which blocks the Stat3 pathway, necessary for maintaining self-renewal. Since, a variety of culturing conditions have been published, in order to maintain and investigate ESC in different stages of development; serum-LIF is a common culturing media maintaining cells in a dynamically heterogeneous mix of epiblast-primitive endoderm primed states [21, 22]; in another common culturing condition, serum-free 2i media, inhibitors of both Gsk3 and MEK act to prevent ESC from exiting the naïve early epiblast stage [16, 23], keeping cells in a more homogeneous and pluripotent state than serum-LIF, which was demonstrated in its increased functional contribution to mouse chimeras [22]. Although serum-LIF and 2i culturing conditions recapitulate cells at two stages that are very close in the course of embryonic development, they already reveal very different levels and patterns of transcription factor expression and morphology. While several studies focused on a relationship of cytoskeleton to ESC morphology and properties, description of the properties of ESC in these stages of development is missing so far [24–26].

Overall, ESC allow us to investigate development at the cellular or molecular level, and at the same time it provides a unique system to study and understand the molecular mechanisms underlying cell fate choices and transcription regulation.

1.3 Transcriptional activation

Transcription is the process of copying DNA-sequences into mRNA, which for protein-encoding genes is translated into proteins. Gene expression is fundamental for cellular life, and the differential expression of different genes defines cellular identity and function. Developmental processes are fundamentally dependent on tight regulation of gene expression, requiring specific levels of expression at precise timings in each particular cell type. In embryonic development spatio-temporal dysregulation of expression can lead to lethality or disorders, but also in tissue homeostasis the misexpression of genes is associated with various disorders and diseases, including cancer.

Gene expression can be regulated on various levels, from transcription initiation and elongation, through mRNA stability to protein translation or degradation. As the first step of gene expression, transcription initiation is a fundamental step in this process and a crucial target for regulation.

While the importance of transcription initiation is clear, the process itself is extremely complex and our understanding of it is still far from complete.

Transcription of mRNA in mammalian cells occurs when RNA polymerase II (RNAP) binds to promoter sequences upstream of the gene-encoding DNA sequence and gains access to single-stranded DNA (initiation), reads the sequence by complementing the DNA bases on the 3' to 5' strand with their RNA base complementary pairs (elongation), and finally, as the termination sequence is read by RNAP, the mRNA strand is detached and transcription ends (termination).

This description is of course overly simplified. We know that RNAP in eukaryotes does not simply bind the promoter by itself, but requires the sequential binding of several general transcription factors, which all together form the transcription pre-initiation complex (PIC) directing RNAP to the transcription start site [27]. However, PIC assembly has only been shown to explain a low, basal level of expression *in vitro* by itself. For higher levels of expression, cells rely on the binding of additional classes of transcription factors and complexes in order to access and prepare regulatory elements for binding

of the transcription machinery (figure 1.5).

Pioneer transcription factors such as Sox2, Oct4, and Klf4, are proteins which can access and bind to nucleosomal DNA - DNA usually 'hidden' from binding by being wrapped around histones. By binding nuclesomal DNA, they induce regions of open or nucleosome free DNA (NFR on figure 1.5). This in turn recruits more factors and larger complexes such as Mediator and p300 to bind the region, further inducing chromatin accessibility through histone acetylation, chromatin remodelling, and interaction with the PIC.

We still do not have a clear understanding of how enhancers regulate transcription at promoters. Even defining and identifying promoters and enhancers elements is complex [28], and indeed these regulatory elements have been shown to exhibit both functions, suggesting they are not necessarily distinct, but instead a common class of regulatory elements with varying degrees of promoter and enhancer activity [29].

A common model suggests direct interaction through chromatin looping [27], however recent studies in *Drosophila* have demonstrated enhancer 'promiscuity', in the sense that a single enhancer coordinated simultaneous transcription bursting from two separate promoters [30]. Heist *et al.* (2019)[31] measured the distance between two such activated promoters to be more than 100 nm during transcription, suggesting a different mechanism than direct enhancer-promoter interaction. Clustering of RNAP was observed in live cells to be directly correlated with transcriptional activity [32] and perhaps molecular clustering prevents direct interaction.

However, chromatin 3D conformation has shown close correlation with gene-expression, and some degree of proximity does seem necessary for transcription [33, 34]. Recent studies observed large-scale clustering of transcription co-activators and RNAP, implicating liquid-phase separation by large aggregations of intrinsically disordered protein domains as a mechanism of enhancing recruitment and stability of transcriptional machinery to a cluster of promoters [35–37].

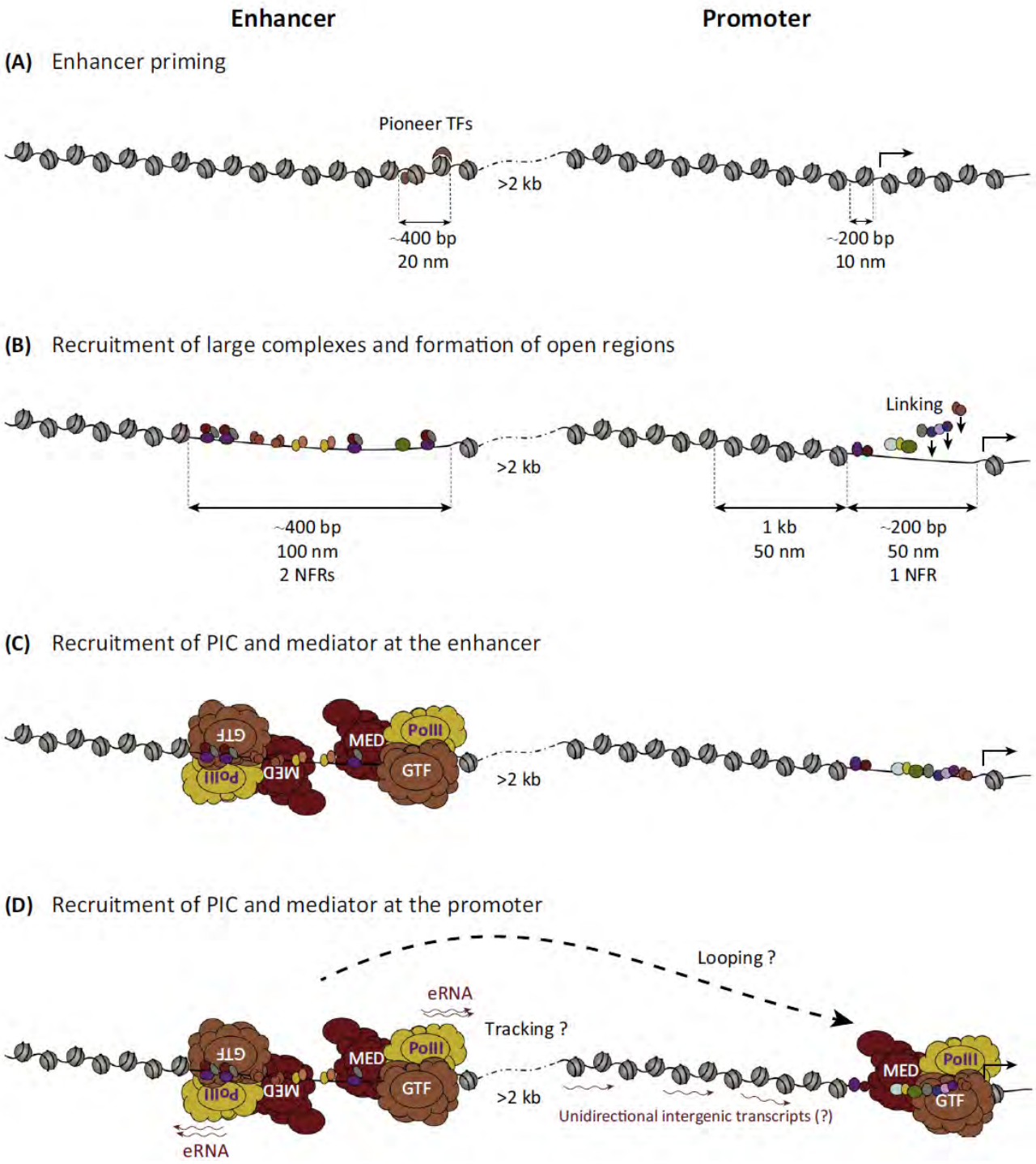

Figure 1.5: Model of transcriptional activation. **(A)** Pioneer transcription factors prime enhancers for activation by binding to nucleosomal DNA. **(B)** Pioneer transcription factors free up regions of DNA from nucleosomes allowing the binding of additional transcription factors and complexes. Similarly the promoter is bound by a sequential recruitment of activators and transcription factors. **(C-D)** At both enhancer and promoter, the transcription pre-initiation complex (PIC) and Mediator are eventually recruited. The active enhancer regulates the promoter through an as of yet unknown mechanism, possibly looping. Figure from Vernimmen and Bickmore (2015)[27].

1.4 Fluorescence microscopy

For the study of the function and organization of specific molecules in the cell, fluorescence microscopy is a versatile and powerful technique. Fluorescence microscopy is effective due to two fundamental properties. Firstly, fluorescence is the emission of longer wavelength light from molecules excited by shorter wavelength light, meaning that by filtering out the excitation wavelength(s) from the detection pathway, we detect only the emitted light from fluorescent molecules, resulting in very high contrast. Secondly, fluorescent molecules may be specifically connected to proteins or other molecules of interest, for example through fusion of a fluorescent protein DNA sequence to the gene of interest, or through fluorescently labeled antibodies that bind to specific proteins. Another highly useful feature of fluorescence microscopy is the possibility of labeling separate proteins simultaneously with spectrally distinct fluorophores, allowing the independent detection, (co)localization, and ratio of multiple molecules of interest.

While the versatility of fluorescence microscopy is apparent by the sheer number of fluorescence images in most biological journals, it is not without restrictions. One of the major restrictions, if one is interested in sub-cellular processes, is the fact that the spatial resolution of a light microscope is fundamentally limited by the diffraction of light. In a conventional fluorescence microscope, light emitted by a point source at the sample will be only partially collected by the objective lens, due to the finite size of the aperture. The restriction of light through an aperture causes diffraction, and as a result, once the image of the point source is focused onto the camera chip, what appears is not a point, but an Airy pattern defined as the point spread function (PSF) of the system. The width of the PSF is dependent on the size of the numerical aperture of the objective lens (NA), and on the wavelength of light (λ) following the equation $r_{psf} = 0.61\lambda/NA$ where r_{psf} is the radius from peak to the first minimum of the Airy pattern as shown in figure 1.6. The resolution of a microscope can be defined as the highest spatial frequency detectable in a sample, or as the smallest possible separation between two point emitters that still renders them distinguishable. The physical limitation on the resolution can be formulated as r_{psf} as above, which is called the Rayleigh resolution limit, basically stating that for two point emitters to be distinguished, the PSF of one emitter cannot be closer than the first minimum of the other emitter's PSF (figure 1.6). Perhaps more famous is the earliest definition of the resolution limit, formulated by Ernst Abbe in 1873 and defined as $d = 0.5\lambda/NA$. Regardless of which definition is used, the resolution limit is typically on the order of $\sim$200 nm for state-of-the-art conventional microscopes. The resolution limit is thus not a problem for the study of cellular behavior (on the order of tens of microns), but to investigate individual molecules or to image sub-cellular protein localization and organization of e.g. actin networks, it quickly becomes a limiting factor (figure 1.7).

1.4.1 Super resolution microscopy

In order to overcome the diffraction limit in light microscopy, a variety of methods have been developed in the past three decades, resulting in the awarding of the 2014 Nobel Prize in Chemistry to Eric Betzig, Stefan W. Hell, and William E. Moerner for their roles in developing super resolution fluorescence microscopy [38]. While many methods and adaptations exist, super resolution microscopy can generally be divided into three main techniques; stimulated emission depletion (STED) [39], single molecule localization microscopy (SMLM) [40–42], and structured illumination microscopy (SIM) [43].

Stefan W. Hell is credited with the invention of Stimulated Emission Depletion (STED) microscopy [39], which as the name suggests, takes advantage of the fact that fluorophores in their excited state can be stimulated to emit light of

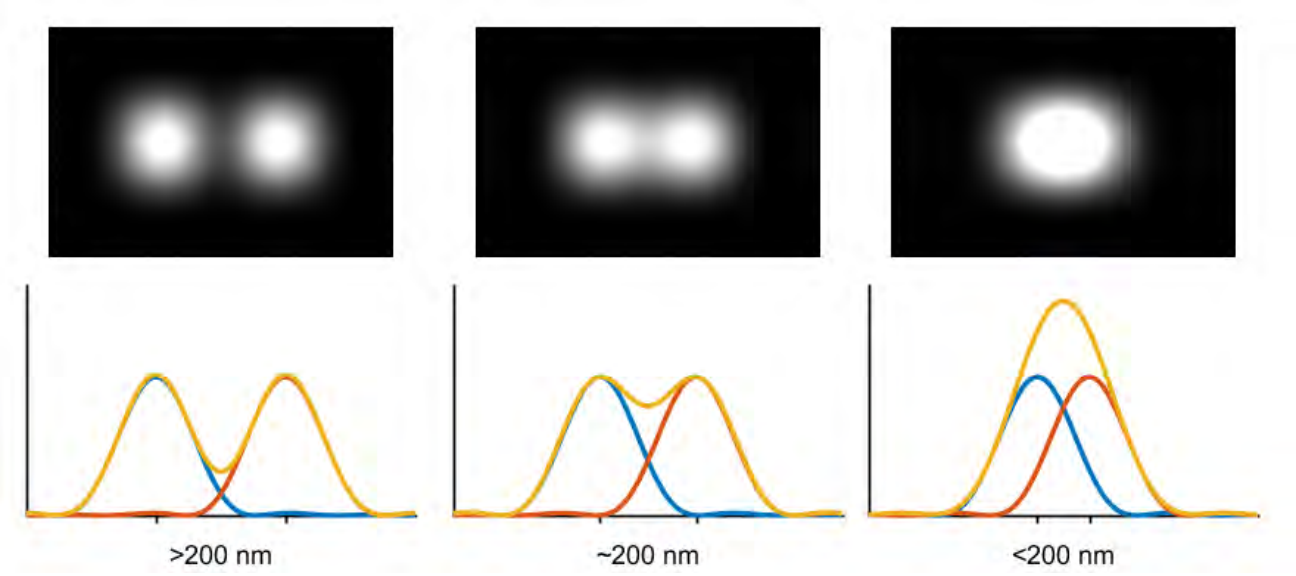

Figure 1.6: Illustration of the Abbe diffraction limit. As two single molecule emitters are moved closer together, their point spread functions overlap and at some distance become indistinguishable.

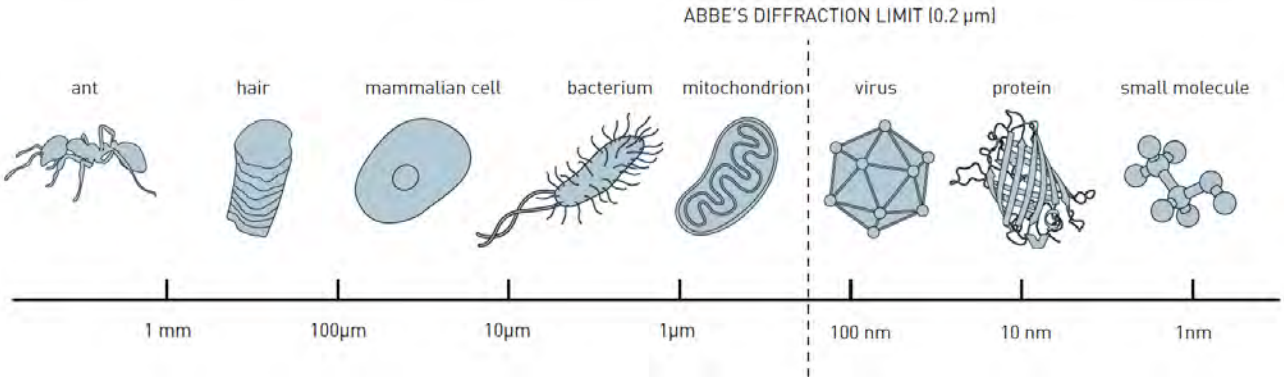

Figure 1.7: Conventional visible light microscopy is limited by diffraction to resolutions lower than the Abbe diffraction limit. This means that spacial features smaller than ∼200 nm cannot be distinguished. Figure from the 2014 Nobel Prize in Chemistry [38].

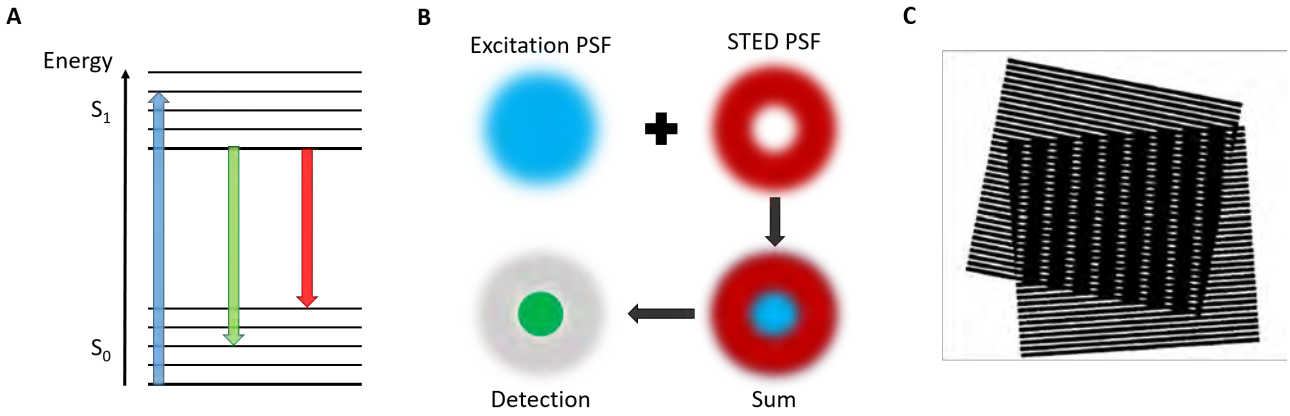

Figure 1.8: **(A)** Principle of stimulated emission. Fluorophores can be put in an excited state by excitation light (blue), and would eventually return to the ground state by emission of light according to their fluorescence emission spectra (green). If irradiated by a certain wavelength light (red), the excited fluorophores can however be stimulated to emit light at that same wavelength and transition to the ground state. **(A,B)** adapted from [1]. **(C)** The principle of structured illumination microscopy. Overlapping two high spatial frequency patterns (e.g. fluorescent sample and structured excitation PSF) reveals a Moiré pattern with much courser features. Knowing the exact structure of one of the patterns, one can determine high frequency information about the other from the Moiré pattern. **(C)** adapted from Gustafsson (2000)[43].

a specific wavelength when irradiated with that wavelength (figure 1.8A). STED microscopy is based on laser scanning confocal microscopy, where fluorophore excitation is achieved point by point, by scanning a diffraction limited laser focus through the sample, and a pinhole is used to eliminate out of focus light from the detection. In STED however, stimulated emission is exploited to narrow the effective excitation PSF beyond the diffraction limit by introducing a donut-shaped PSF of stimulated emission wavelength light around the excitation PSF (figure 1.8B). This STED beam depletes the irradiated volume of fluorophores in the excited state, by stimulating them to emit light of the STED wavelength, which is then filtered out of the detection pathway. By narrowing the STED "donut", the effective excitation PSF can thus be decreased beyond its original diffraction limited size. As the intensity of the STED beam is increased, the donut shape will swell and overlap more of the excitation PSF, thus increasing the imaging resolution. In theory, the STED beam could be increased infinitely, reducing the excitation PSF an infinitesimal volume, however in practice the resolution is limited by the laser power, signal-to-noise, and fluorophore stability - and viability if investigating live cells. Typically STED imaging can reach resolutions in the tens of nanometers, however to achieve this, it also needs peak intensities of up to 1 $\mathrm{GW/cm^2}$ [44], which would induce significant photodamage in living cells [45], and it is mostly a method reserved for fixed cell imaging. However, as with all super-resolution methods, the technique is constantly evolving, and a recent application adapted STED imaging with target-locking to an active transcription site and measuring transcription factor presence and accumulation on a specific gene in live cells [46].

Structured illumination microscopy (SIM) is another way of manipulating the excitation PSF in order to break the resolution limit [43]. In SIM, the excitation PSF is structured in a line pattern with a known spatial frequency and orientation, which is then rotated and phase-shifted a certain number of times depending on the implementation. If we consider a fluorescently labeled sample as a structure containing certain spatial frequencies and orientations, we can imagine that we are thus superimposing two such patterns. Similar to Moiré patterns, this gives rise to interference patterns which have larger features than each of the underlying patterns. Knowing one pattern (the structured illumination), and detecting the

coarser interference pattern, it is possible to calculate higher frequency information about the underlying sample than is possible within the resolution limit. Combining the structured illumination patterns and the images from each rotation and phase-shift, it is possible to reconstruct the sample image with a resolution typically on the order of 100 nm, surpassing the diffraction limit by a factor two [43]. Key advantages of SIM are that it works with any fluorophores, does not require high excitation intensities and is relatively fast. However, the main drawback is that the resolution gain beyond the diffraction limit is relatively small compared to STED or SMLM.

1.4.2 Single Molecule Localization Microscopy

Single molecule localization micrscopy (SMLM) covers a range of super resolution microscopy techniques which all are based on the principle of reconstructing an image from a set of precise localizations of individual fluorescing molecules. These include among others, photo-activated localization microscopy (PALM)[40], stochastic optical reconstruction microscopy (STORM)[41], and direct STORM (dSTORM) [42, 47]. These methods all overcome the resolution limit by utilizing the fact that an individual point emitter, even in a diffraction limited system, can be localized to a much higher precision than the resolution limit (figure 1.9). The center of the PSF can be well approximated by a Gaussian function, and as such the error on the center of the signal ($\sigma_{x,y}$ is given as [40];

$$\sigma_{x,y} \simeq s/\sqrt{N} \tag{1.1}$$

Where s is the standard deviation of the Gaussian fit to the emission signal, and N is the number of photons detected. As such, the precision of the localization is theoretically only limited by the number of photons collected, but in practice, sample drift and the limited photon budget of fluorophores typically limit SMLM to resolutions on the order of 10s of nanometers [40–42].

For SMLM to work, one needs a way to image subsets of fluorophores in way that each PSF can be individually fitted with a Gaussian distribution. SMLM techniques achieve this by separating the active fluorophores in time, gathering many thousands of frames and composing an image by overlaying plotting every fluorophore localization (figure 1.9). The main difference between the different SMLM techniques lies in the way they induce this switching of fluorophores. PALM as the name suggests uses photo-activatable fluorophores and short bursts of photo-activation light to switch on small random subsets of fluorophores which are then localized [40] and bleached. Photo-activatable fluorophores are only excitable by their excitation wavelength after irradiation with a shorter wavelength light, for example the photo-activatable PA-GFP which is excitable at 488 nm only after activation with 405 nm light.

The original implementation of STORM used activator-reporter dye pairs of Cy3-Cy5, but was otherwise the procedure was similar to PALM, including rounds of photo-activation of a stochastic subset of fluorophores, excitation, and localization [41]. Finally, the more recent dSTORM technique does not rely on photo-activation to switch on molecules, but instead uses specialized buffer conditions to drive organic dyes to a non-fluorescing triplet state [42, 47]. Under standard imaging conditions, organic dyes such as Cy5 and Alexa Fluor 647 undergo cycling between the ground state, the excited state and a non-fluorescing triplet state. The triplet state is very short-lived, but presence of β-mercaptoethylamine (MEA) drives fluorophores in the triplet state to a more stable radical anion state, which has to react with oxygen to return to the ground state. This can be further stabilized by adding an oxygen scavenging system to the buffer (commonly glucose and glucose oxidase), keeping most fluorophores in a state where they can not be excited or fluoresce. Finally since fluorophores can only enter the triplet state from the excited state, the laser power can be adjusted to tune the density of actively fluorescing molecules in the sample.

However, it is not uncommon to see the terms used interchangeably, and indeed the underlying principle of single-molecule localization is identical in all the methods.

1.4.3 Total Internal Reflection Fluorescence

Single molecule localization microscopy is highly reliant on maximizing signal-to-noise ratio in order to be able to detect and fit individual fluorescing molecules, as well as limiting the number of emitters within the field of view. One way of

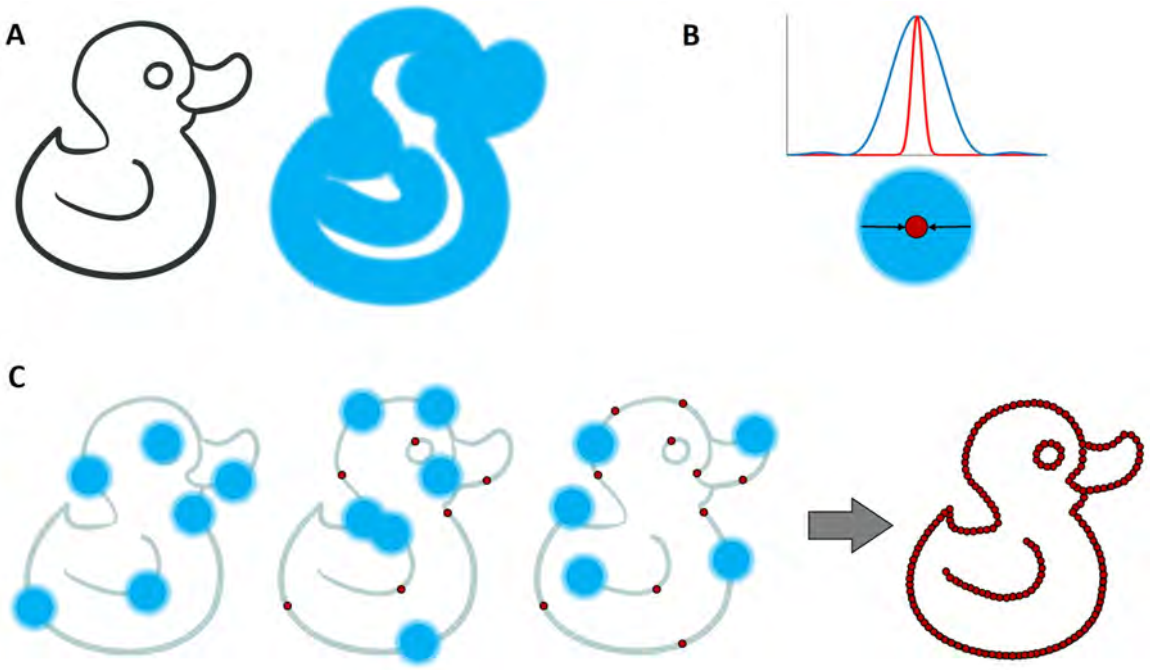

Figure 1.9: Principle of single molecule localization microscopy. **(A)** Due to the finite size of the PSF in diffraction limited microscopy, the image of a sample is the convolution of the shape with the PSF, and will thus appear blurry. **(B)** The location of the center of individual fluorescent molecules can be determined with much higher precision than the width of the signal, by fitting the PSF with a Gaussian function. **(C)** By exciting just a small subset of fluorophores at a time, as long as their PSFs do not overlap, the center of each emission can be precisely localized. Repeating this procedure many times, one can eventually reconstruct the underlying shape with a resolution much better than the diffraction limit. Figure from [1].

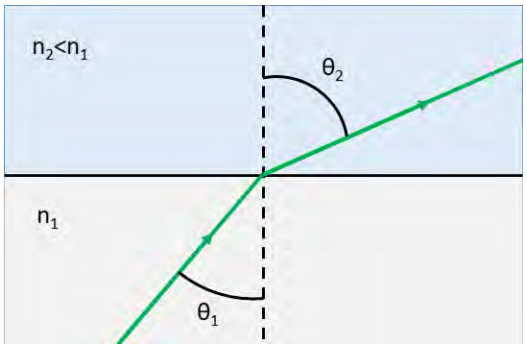

Figure 1.10: A beam entering a lower refractive index medium is refracted at an angle higher than the angle of incidence in accordance with Snell's law.

both improving the signal-to-noise ratio and reducing the active emitters to a narrow volume is by angling the excitation beam out of the objective.

In conventional epifluorescence microscopy, the illumination consists of a collimated beam exciting the objective along the optical axis, illuminating everything in the z-direction. This results in significant background fluorescence, which can be avoided by restricting the illumination to a thin volume above the sample coverslip, using total internal reflection fluorescence (TIRF) microscopy [48].

Snell's law describes the path of a propagated beam refracted by a change in refractive index:

$$n_1 \sin(\theta_1) = n_2 \sin(\theta_2) \tag{1.2}$$

where θ_1 and θ_2 are the angles relative to the interface normal of incidence and refraction respectively, and n_1 and n_2 the refractive indices of the two media. If we consider the excitation beam from a fluorescence microscope encountering the interface between coverslip glass ($n_1 = 1.52$) and the sample media ($n_2 = 1.33$), the angle of the refracted beam will be at a higher angle than the incident beam (figure 1.10). If we increase θ_1 it is apparent that θ_2 will eventually reach $90°$ and not enter the sample. This occurs at the critical angle, defined from Snell's law as $\theta_c = \arcsin(n_2/n_1)$, beyond which the incident beam is completely reflected by the interface. However, while the beam is reflected, some energy enters the sample as an evanescent field, occurring close to the surface and falling off exponentially with z as [49]:

$$I(z) = I(0)\, e^{-z/d} \tag{1.3}$$

Here d is defined as the penetration depth, given by:

$$d = \frac{\lambda}{4\pi\sqrt{\left(n_1^2 \sin^2(\theta_1) - n_2^2\right)}} \tag{1.4}$$

where λ is the wavelength of the light. This evanescent field can be used to excite fluorophores close to the surface, which is exactly the principle of TIRF microscopy [48].

Considering that numerical aperture (NA) of an objective is defined as $n \sin(\theta_{max})$ where θ_{max} is the maximum incidence angle, we see that the critical angle can only be reached if the NA of the objective exceeds the refractive index of the sample, necessary for TIRF. Furthermore, the larger the difference between between NA and refractive index of the sample, the smaller the minimum possible penetration depth. The microscope constructed and used in later chapters of this thesis is equipped with 1.45 NA objective, which for 640 nm light and sample refractive index of $n_2 = 1.33$ yields a minimum penetration depth of $d_{min} = 88\text{nm}$.

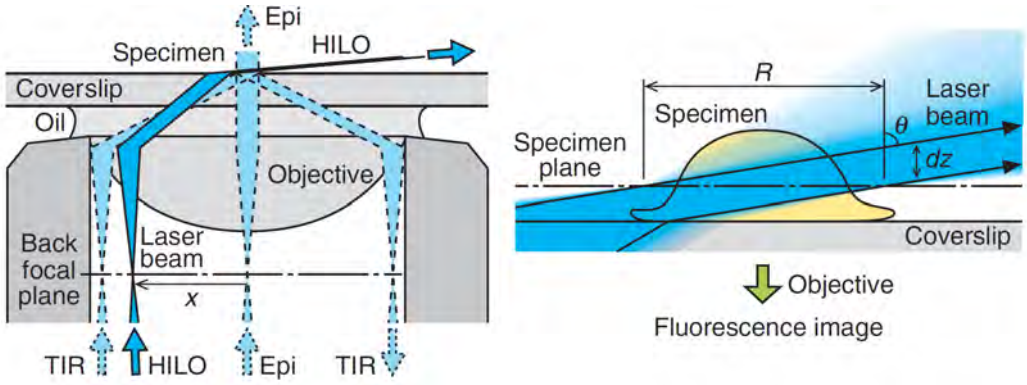

Figure 1.11: Figure from Tokunaga *et al.* (2008)[48]

An alternative, but related, method of illumination is termed highly inclined and laminated optical sheet (HILO) [48] illumination. To achieve HILO illumination, the incident excitation beam must enter the sample at a high angle, far from epifluorescence but without reaching the critical angle, as shown in figure 1.11. In this HILO mode the sample is illuminated in a thin sheet, thicker than TIRF, but with the advantage of being able to go several μm into the sample instead of being restricted to the surface.

Chapter 2

Construction of fluorescence microscope for sub-cellular imaging

This PhD project was structured as an integrated MSc and PhD degree with 2 years for each half. The first two years concerned constructing a microscopy platform for TIRF/STORM imaging and single-molecule tracking, in addition to implementing super-resolution imaging for stem cell imaging at the sub-cellular level. This work was described in detail and submitted as a MSc thesis, functioning as a qualifying exam for the final two years of the PhD. In this chapter, I will therefore give only a brief summary of the construction while referencing the previously published thesis [1].

2.1 TIRF/STORM microscope setup

To achieve sub-cellular resolution for fluorescence imaging of stem cells, it was decided to implement single molecule localization microscopy (SMLM) as the best compromise between resolution, speed, and photo-toxicity.

While SIM offers fast imaging at relatively low illumination intensities, the resolution gain is only limited, and SIM would not provide an ideal platform for single-molecule tracking, which was another criteria. STED provides high resolution, and is generally faster than SMLM, however the high light-dose required for imaging would severely limit its application for live-cell imaging and molecule tracking. While SMLM certainly also has its disadvantages, such as requiring long time-lapse acquisitions and high laser power for super-resolution imaging, which typically limits it to fixed-cell imaging of structures, it is however easily adapted to allow single-particle tracking in live cells. The crucial features of an imaging platform required for SMLM imaging are identical for single particle tracking; a sparse density of fluorophores to ensure non-overlapping PSFs, high signal-to-noise ratio to detect and localize individual molecules with high precision, and sensitive detection.

For single-particle tracking, a TIRF microscope in HILO illumination mode will reduce out-of-focus light, increasing signal-to-noise ratio, while reducing the number of active emitters by narrowing the excitation volume. Additionally, with the photo-activation platform needed for many iterations of SMLM, it is possible to easily control the density of active emitters for tracking, by tagging them with a photo-activatable fluorophore. By using short and low-power bursts of activation light, it is possible to activate just a small sub-population for tracking, allowing the tracking of individual molecules even within a crowded population [32, 35, 50].

In the following, the setup is first briefly described, along with a list of the components involved.

Description of the setup

The setup is based on the Olympus inverted microscope IX83, with additional modules to correct for z-drift and to control the excitation angle out of the objective (figure 2.1). The setup is in large part inspired by the many previously published SMLM setups, particularly the dSTORM setup described in van de Linde *et al.* (2011)[42]. Z-drift is controlled by the Olympus Z-Drift Controller (ZDC2, Olympus), which utilizes a low power (240 µW) 790 nm laser reflecting off the interface between coverslip and sample, and corrects the Z-position of the sample in real-time by measuring the reflected beam relative to a calibrated value set at the start of acquisition. The TIRF illumination module allows the motorized changing of excitation beam angle out of the objective, by adjusting the focus position of fiber-launched lasers on the back-focal plane (BFP) (figure 2.2). Two independently controlled fibers are connected to the TIRF module, one containing 405 and 488 nm lasers, and the other 561 and 640 nm lasers. Since the penetration depth of the evanescent field of TIRF illumination is wavelength dependent (eq. 1.4), to illuminate the same volume for e.g. activation (405 nm) and excitation (most commonly 561 or 640 nm), the two fiber positions thus have to be adjusted independently.

The setup is set up with 4 laser lines as mentioned; a 40 mW 405 nm laser for photo-activation and imaging of DAPI, and three 200mW lasers of 488 nm, 561 nm, and 640 nm for excitation of a variety of fluorophores. As shown in figure 2.1, the four laser lines are individually directed through the aperture of the acousto-optic tunable filter (AOTF), which allows the fast shuttering of selected wavelengths, by applying an acoustic wave across a material which changes its optical properties when exposed to an acoustic wave (an acousto-optic crystal). Applying an acoustic wave at a certain frequency to the crystal creates a fluctuating index of refraction across the crystal, which diffracts light within a narrow range of frequencies. By tuning the frequency of the acoustic wave, we can control which wavelengths to diffract and direct into the optical fibers. The AOTF is connected via the Olympus real-time controller (RTC) using a binary on/off 5 V TTL signal to synchronize the shuttering of lasers with the camera shutter, eliminating illumination of the sample outside of the camera exposure window.

To have single molecule sensitivity, the camera choice for SMLM microscopes typically stands between sCMOS cameras and electron-multiplying charge-coupled device (EMCCD) cameras. While both are feasible, we here chose an EMCCD camera as they generally have higher sensitivity at low-light applications such as single-molecule detections, at the cost of speed and smaller field of view. The EMCCD camera has chip size of $8.19\mathrm{mm} \times 8.19\mathrm{mm}$, and a pixel size of $16\mu\mathrm{m} \times 16\mu\mathrm{m}$.

As described in the TIRF section of the introduction, for TIRF to even be achievable, it is necessary to have an objective with NA higher than the refractive index of the sample (typically $n = 1.33$). For this reason, the setup is equipped with a 150X NA 1.45 oil immersion objective.

A final important consideration of a SMLM setup is sufficient sampling of the PSF for single-molecule detection. The Nyquist-Shannon sampling theorem states that for a function of frequency f to be determined, one needs to sample it every $1/(2f)$ steps. In our case the PSF is the function to be sampled, which has a typical width of around 250 nm depending on wavelength. In order to adequately determine and fit the PSF, we therefore need a pixel width of less than 125 nm. In the case of our 150X objective and 16 µm pixel width, each pixel corresponds to $106.7\mathrm{nm} \times 106.7\mathrm{nm}$ in the sample, meaning the PSF is sufficiently sampled by our setup.

Once images are acquired, the final step of SMLM is to detect individual molecules and generate a map of localizations to reconstruct the underlying sample. To do this, we use the open-access ThunderSTORM plugin for ImageJ [51–53].

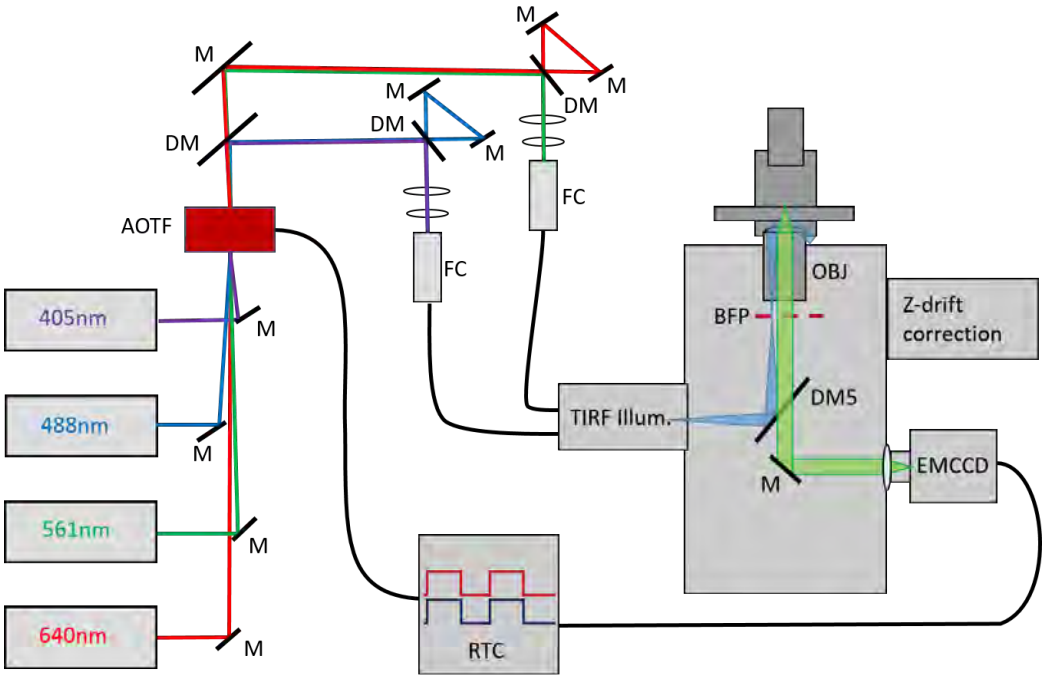

Figure 2.1: Schematic of the STORM/PALM super resolution setup. Four lasers of 405, 488, 561, and 640 nm wavelength are directed using mirrors (M) through an acousto-optic tunable filter (AOTF), which functions as a fast shutter for each of the wavelengths independently. After the AOTF, 405 and 488 nm lasers are separated from 561 and 640 nm lasers by a long-pass dichroic mirror (DM). The lasers are coupled in pairs into two separate single-mode fibers by the use of mirrors (M), long-pass filters (DM) and an optomechanical fiber coupler (FC). The fibers attach to the TIRF illumination module on the IX83 inverted microscope, where the lasers are focused onto the back focal plane of the objective (BFP) by adjusting the position of the fiber launches (see also figure 2.2). Fluorescence excitation and emission are filtered by quad-band emission filter (DM5), blocking wavelengths around 405, 488, 561, and 640 nm from the detection path. For 3D STORM imaging a cylindrical lens is inserted just before the EMCCD camera. The EMCCD camera and the AOTF are synchronized by the Olympus Real-Time Controller (RTC), shuttering the lasers in between camera exposures. The Z-drift correction module corrects the z-position of the objective in real-time based on a reflected laser beam.

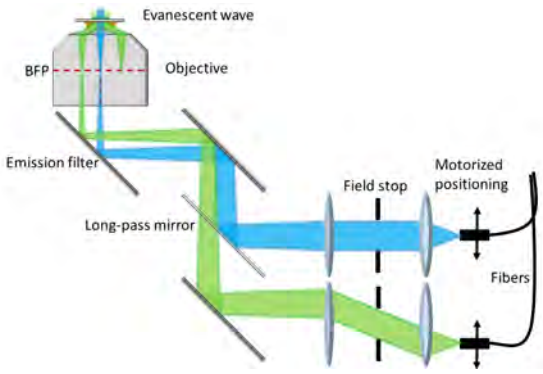

Figure 2.2: Schematic of the TIRF illumination module in (figure 2.1). The fiber output is focused onto the back focal plane by horizontal movement of the two fiber launches. Displacing the fibers vertically then shifts the focus point away from the optical axis, angling the beam out of the objective away from epifluorescence illumination. As the angle is increase, it eventually reaches first HILO illumination before entering the TIRF regime. The two fiber outputs are separately calibrated for TIRF or HILO, and combined by a long-pass dichroic mirror. Figure from [1].

Components of the setup

- Inverted microscope (IX83, Olympus)
- 150x, 1.45 NA oil-immersion objective (UAPON OTIRF, Olympus)
- Acousto-optic tunable filter (AOTFnC-400.650, AA Opto Electronic)
- Lasers
    - 405 nm, 40 mW (FV-LD405, Olympus)
    - 488 nm, 200 mW modulated laser diode (MLD 488, Cobolt)
    - 561 nm, 200 mW diode-pumped solid state laser (Jive, Cobolt)
    - 640 nm, 200 mW diode laser (iBeam smart, Toptica)
- Mirrors
    - Quad-band filter (405/488/561/640m-TRF, Chroma)
    - Long-pass filter (593 LP BrightLine HC, AHF Analysentechnik)
    - Dichroic Beamsplitter 473-491 (LaserMUX, AHF Analysentechnik)
    - Dichroic Beamsplitter 375-415 (LaserMUX, AHF Analysentechnik)
    - Dichroic Beamsplitter (HC BS R488, AHF Analysentechnik)
    - Long-pass filter (AT 515LP, AHF Analysentechnik)
    - Short-pass filter (532 SP Edge Basic, AHF Analysentechnik)
    - Mirrors (MaxMirror UV-VIS-NIR $\lambda/10$, AHF Analysentechnik)
- Opto-mechanical fiber couplers (kineMATIX, Qioptic)
- Single-mode fibers (Qioptic)
- EMCCD camera (ImagEM X2, Hamamatsu)
- TIRF Illuminator (cellTIRF, Olympus)
- Cylindrical lens, 1m focal length
- Z-drift controller (ZDC2, Olympus)
- Real-time controller(RTC, Olympus)
- 4066 quad bilateral switch

Chapter 3

Single molecule tracking in mESC reveals highly stable binding of Sox2 dependent on BRD4

Kasper G. Hvid[1,3], Karolina Ditrychova[2,3], Lene B. Oddershede[1], Joshua M. Brickman[2]

3.1 Introduction

Transcription of genes is one of the most fundamental concepts in cellular biology, and forms the basis of all life as we know it. The central dogma of biology says that DNA is transcribed to mRNA which in turn is translated into proteins, the molecular motors of life.

While many parts and mechanisms involved have been extensively studied, for example the initiation and movement of RNA polymerase in transcribing DNA [32, 35, 46, 54–56], it is also clear that the process is incredibly complex and involves countless parts and pieces, that all have to play together in order to give rise to both the incredible complexity and diversity found in living organisms, while robustly replicating the tightly regulated transcription patterns and timings of embryonic and organism development [57].

The dependence of transcription on specific transcription factors (TFs) binding to gene regulatory sequences is well documented, and mis-expression or absence of even single core pluripotency factors disrupts proper embryonic development [6, 11]. Conversely the over-stimulation of adult cells with a cocktail of pluripotency factors can reverse differentiation and induce pluripotency in already differentiated cells [58].

However, the mechanism of how TFs are recruited to the regulatory sequences, how their binding is regulated, and crucially how they regulate transcription is not yet well understood. Many genes are activated through interaction with a cis-regulatory enhancer element, which may be several mega-bases from the promoter, however, how exactly is this interaction formed, how dynamic is it, and how is it affected by transcription, all remain open questions.

We know that transcription factors are binding both promoters and enhancers, and large transcription initiation complexes may form between the two, perhaps stabilizing the interaction. Perhaps transcription factors recruit the enhancer to

[1]Niels Bohr Institute, University of Copenhagen
[2]Novo Nordisk Foundation Center for Stem Cell Biology, University of Copenhagen
[3]Shared first authors

the promoter, or perhaps the enhancer-promoter interaction is what stabilizes the transcription initiation complex [27].

Recent studies have shown that transcription factors can remain bound at both enhancers and promoters of inactivated genes [14, 17] for hours, and other studies revealed the promiscuity and dynamical movement of enhancer elements relative to their associated promoters [30].

Recent advancements, particularly in the field of single molecule fluorescence microscopy, including the development of alternative modes of illumination [48, 59, 60], and improved, photo-activatable fluorescent probes [61–63], have allowed the detailed study of the movements of individual TFs in living cells [64–67]. Previous studies have mainly focused on describing in detail the dynamical movement and mechanisms of TF binding in a few, simple biological conditions, rather than map their changes over complex biological transitions.

Here we investigate a recently described model of transcriptional priming in ESC, from a state of pluripotency to primitive endoderm primed cells, by Hamilton *et al.* [14, 17]. This model system replicates *in vitro* the developmental stage of primitive endoderm priming and differentiation from naive epiblast in pre-implantation embryos, as induced by FGF/ERK signaling, by directly inducing ERK phosphorylation with a small-molecule drug [14, 17]. Transcriptional priming in this way, involves quickly repressing the expression of core pluripotency factors regulated by Sox2, as observed by a loss of RNAP from their promoters. Interestingly, Hamilton *et al.* [17] recently showed that the decommissioning of these genes was not accompanied by loss of TFs including Sox2 from promoter or enhancer. TFs would stay bound for hours after initiation of primitive endoderm priming, even after losing RNAP and Mediator complex cooperativity.

In this study, we apply single molecule tracking of the core pluripotency transcription factor Sox2 in live embryonic stem cells to elucidate Sox2 binding and residence time in dependency on various transcription co-factors.

3.2 Methods

Cell culture

Murine embryonic stem cells (ESC) E14 were cultured at 37 °C, in humidified incubators with 5% CO2, in Glasgow's Minimum Essential Media (GMEM, Sigma-Aldrich) supplemented with 10% Fetal Bovine Serum (FBS)(Gibco), 2 mM L-Glutamine (Gibco), 1 mM Sodium Pyruvate (Gibco), 0.1 mM Non-Essential Amino Acids (Gibco), 1000 U/mL Leukemia Inhibitory Factor (LIF)(DanStem, University of Copenhagen), and 0.1 mM 2-mercaptoethanol (Sigma), in culture flasks (Corning) coated with 0.1% gelatin. Non-endogenous cell lines were cultured in media containing puromycin or geneticin depending on the resistance.

Stable cell line generation

An endogenous Sox2-HaloTag reporter (Sox2-HT) was established by CRISPR-Cas9 induced homologous repair in the parental E14 line. The plasmid pUC19-Sox2-HaloTag-IRES-Puro was generated by modifying pUC19-Sox2-EGFP-IRES-Puro (obtained from F. Roske) by substitution of EGFP with HaloTag. Using Lipofectamine 2000, the plasmid was co-lipofected with CRISPR-Cas9 plasmid and gRNA targeting the Sox2 locus (obtained from F. Roske) in the ratio 4:1:1, respectively. Puromycin (1 µg/mL) was added 2 days after lipofection and individual colonies were picked and expanded following 7 days of selection. Clones were selected by Western blot and tested by Southern blot, colony formation assay and Sox2 locus sequencing. As a last test, selected clones were labeled with Janelia Fluor 646 or 549, and nuclear

localization of the dye was checked on a confocal microscope (Leica SP8). Clone Sox2HT5 was used for the consequent experiments in this study.

H2B-HT cell line was generated by electroporation of LZ10-PBREBAC-H2BHalo (gift from James Zhe Liu; Addgene plasmid #91564: ref. [68]) into the parental E14 line. Geneticin (150 ug/ml) was added 24 hours after electroporation and individual colonies were picked and expanded. Selected clones were labeled with Janelia Fluor 549 and nuclear localization of the dye was checked on a confocal microscope (Zeiss 780). NLS-HT cell line was generated from parental E14 cell line. Plasmid pCAG-NLSHT-IRES-Neo was generated by cloning NLS-HaloTag gblock (IDT DNA) to pCAG backbone containing IRES-Neo. The plasmid was electroporated into the E14 cell line, and geneticin (300 µg/ml) was added 24 h after transfection for 7 days and individual clones were picked and expanded. Selected clones were labeled with Janelia Fluor 646 or 549, and nuclear localization of the dye was checked on a confocal microscope (Leica SP8). An Erk-inducible variant of the Sox2HT reporter line (Sox2HT-Erk) was generated by electroporation of the plasmid pCAG-cRAF-ERT2-IRES-Puro from W. Hamilton [17] into parental Sox2HT5 line. Clones were selected with 300 ng/mL geneticin for 7 days starting 24 hours post-transfection.

Sample preparation for imaging

For live-cell imaging, the cells were plated on fibronectin-coated 8-well chambered #1.5H glass bottom slides (Ibidi) at a density of approximately 50000 cells/cm^2. Fibronectin coating consisted of covering glass slides in 10 µg/mL fibronectin in DPBS for at least 1 hour prior to plating. Samples were left to grow overnight before staining. For single-molecule detection without photo-activation, cells were labeled with 5 nM of Janelia Fluor 646 (JF646)[62] for 15 minutes, before washing 3 times waiting 5 minutes between each wash. For imaging the cells were transferred to medium without phenol red, identical to the Serum+LIF culture media, except GMEM was exchanged with Fluorobrite DMEM (A1896701, ThermoFisher). For experiments with photo-activation, 50 nM of photo-activatable Janelia Fluor 646 (PA-JF646)[63] was used to label the cells, but otherwise the protocol was identical.

ERK activation

For ERK activation experiments (figures 3.3 and 3.4), cells were plated at a density of 25000 cells/cm^2, in normal media supplemented with MEK inhibitor PD17 (PD173074, Sigma), and left to grow for 24 hours. After 24 hours (i.e. time 0 on figure 3.3), cells were either ERK-activated by addition of 250 nM tamoxifen (4-OHT) ((Z)-4-Hydroxytamoxifen, Sigma), or kept as a control by addition of an equal volume of ethanol (ETOH). Cells were imaged after 2 or 8 hours, or after 10 hours with addition of 1 $\mu\mathrm{M}$ MEK inhibitor PD03 (PD0325901, Sigma) for the final 2 hours of the experiment. Each timepoint consisted of 3 biological replicates, and control and activated samples were always prepared and imaged concurrently.

JQ1 and p300i treatment

Samples for inhibitor treatment were prepared as normal, with the non-photoactivatable Janelia Fluor 646 (JF646). The inhibitor JQ1 was added to samples 7 hours before imaging, and an equal volume of DMSO was added to a control group. The p300 acetyltransferase inhibitor (p300i) was added to samples 1 hour before imaging.

Single molecule imaging

Single-molecule imaging was performed on an Olympus IX83 microscope equipped with a 150X oil-immersion objective (NA=1.45, Olympus), an EMCCD camera (ImagEM X2, Hamamatsu), Z drift compensation system (IX3-ZDC2, Olympus), and a quad-band emission filter (405/488/561/640m-TRF, Chroma). For imaging with JF646, we used a 640 nm laser (iBeam smart, Toptica Photonics) fiber coupled to a motorized TIRF module (cellTIRF, Olympus), allowing us to control the angle of the excitation beam out of the objective, which was set to HILO illumination [48], generally at a depth of 5-7 $\mu\mathrm{m}$ into the sample. Time-lapse imaging was acquired for 400 frames at 500 ms exposure time for a total of 4-5 timelapses per condition for two conditions per experiment.

Single molecule tracking

Cell nuclei were manually identified and outlined using the average projection of the timelapse image-stack in Fiji [52]. For each cell nuclei, individual fluorescent spots were localized using the Thunderstorm plugin for ImageJ [53]. A table detailing the position, frame, and intensity of each localized molecule was exported and analyzed by a custom written MATLAB script calling the simpleTracker program designed by Jean-Yves Tinevez [69]. The simpleTracker program first links points between subsequent frames using the hungarian algorithm, minimizing pair distance over all particles. A maximum linking distance of 450 nm between two frames is enforced, assuming a maximum diffusion rate of 0.1 $\mu\mathrm{m}^2/\mathrm{s}$ as described in Chen *et al.* (2014) [65]. Finally the program will go through the individual tracks and link two tracks if the end of the first is close to the beginning of the next, allowing a gap of up to 2 frames. Mean square displacement was calculated using the msdanalyzer class for MATLAB by Tarantino *et al.* (2014) [70].

Residence time fitting

Considering each track to represent a molecule bound to DNA, we used a custom written MATLAB script to analyze and fit the residence time of molecules. First we generated histograms of track lengths and produced the cumulative distribution function (CDF). We fit 1-CDF with exponential decay models containing either 2, 3, 4, or 5 populations with independent lifetimes of the form:

$$f_n = \sum_{1}^{n} a_i \cdot \exp(-t/\tau_i) \tag{3.1}$$

$$1 = \sum_{1}^{n} a_i \tag{3.2}$$

To validate our results and analysis, various methods previously described in single-molecule tracking literature were implemented, including photobleaching corrections [65, 71], fitting thresholds [72], in addition to binning and model comparison by calculating sample-size corrected Akaike information criterion (AICc)[73, 74]. Histogram binning, fitting, and all further analysis was done in custom written MATLAB scripts.

To correct for photobleaching, we calculated the total number of detections per frame over time, $N(t)$, and fit it with a single exponential decay function $N(t) = N_0 \cdot \exp(-t/\tau_B)$, where $1/\tau_B$ is an estimate of the photo-bleaching rate.

Bootstrapping was performed using the MATLAB function bootstrp, resampling the residence time data by randomly pulling 1000 data samples with replacement. Each new data sample was then fitted with an exponential model as normal, and the parameters of each fit displayed in a box-plot. The standard error on each parameter was calculated from the standard deviation of the bootstrap distribution, and the probability of the parameters being equal (the p-value), was calcu-

lated from the difference in means and the error propagation of the standard error of each mean, assuming approximately Gaussian distribution.

RNA extraction and qRT-PCR

RNA was extracted using RNeasy Mini kit (Qiagen). Adherent cells were washed with 1X PBS and collected into RLT buffer (supplement of the kit). RNA extraction thereafter followed manufacturer's protocol. 1 µg of RNA was used for reverse transcription using SuperScriptIII (ThermoScientific) according to manufacturer's instructions, cDNA product was diluted 1:10 and either stored in -20 degrees, or used for qRT-PCR. qRT-PCR was prepared and detected using Universal Probe Library Assay (Roche) system on Light Cycler 480 (Roche). Expression data was normalized using geometric mean of Sdha and Gapdh. Primers were designed using Universal Probe Library Assay Design Centre (Roche).

Western blot

Adherent cells were washed with 1X PBS and lysed in 2X Laemmli buffer (4% SDS, 10% glycerol, 120mM Tris-HCl pH 7.4). Samples were then sonicated for 10 seconds at 20% amplitude using Branson probe sonicator and boiled for 5 min at 70 °C. The concentration of sample was estimated using NanoDrop 2000 spectrophotometer (wavelength 280 nm). 20-40 µg of protein samples was then mixed with 1M DTT and bromphenol blue and boiled for 5 min at 70 °C. Samples were then loaded on nuPAGE 4-12% Bis-Tris Gel (Invitrogen) and run in NuPAGE MES SDS buffer (Invitrogen) for 1 h at 190 V. Gel was transferred to Amersham Protran Premium 0.45 notrocellulose membrane (GE Healthcare) at 400 mA for 1.5 h. Blot membrane was consequently blocked using 10% skim milk (Sigma-Aldrich) for 0.5 h at room temperature and washed 5x with 1X TBST. Primary antibody was diluted in 1X TBST supplemented with 5% bovine serum albumin (Sigma-Aldrich) and incubated overnight at 4 °C. Blot was washed 3x with 1X TBST and incubated for 2 h in room temperature with secondary antibody diluted in 1X TBST supplemented with 5% skim milk. After final washing with 1X TBST, blots were analyzed using ChemiDoc MP system (BioRad). Primary antibodies used: histone H3 (Abcam; 10799; 1:2000), Sox2 (Santa Cruz; sc17320; 1:200), Erk1/2 (p44/42 MAPK; Cell Signalling; 4696S; 1:1000), pErk XP (p44/42 MAPK T202/Y204; Cell Signalling 4370S; 1:1000). Secondary antibodies used: Donkey anti-Mouse Alexa Fluor 488 (ThermoFisher; A21202; 1:2000), donkey anti-Rabbit Alexa Fluor 647 (ThermoFisher; A31573; 1:2000), donkey anti-Goat Alexa Fluor 546 (ThermoFisher; A11056; 1:2000).

Immunofluorescence staining

ESC were grown in respective conditions on micro 8-well slides (Ibidi) in the concentration of 50 000 – 100 000 of cells/cm^2. Cells were washed with 1X PBS and fixed by 4% (w/v) paraformaldehyde (ThermoFisher) for 10 min in room temperature. Samples were washed three times with 1X PBS and fixed with cold methanol for 10 min in -20 °C. Cells were consequently washed three more times with 1X PBS. Samples were blocked with a buffer containing 0.3% Triton X-100 (TBST) and 5% donkey serum in 1X PBS for either 2h at room temperature or overnight in 4 °C. Samples were then incubated overnight at 4 °C with primary antibodies diluted in buffer containing 0.3% Triton X-100 (TBST) and 1% BSA in 1X PBS and washed three times with 1X PBS. Secondary antibodies were diluted (1:2000) in the same buffer and incubated with the sample for 2 h at room temperature. Samples were consequently washed three more times with 1X PBS, incubated with 1 µg/mL DAPI for 5 min at room temperature and washed three times with 1X PBS again. Samples were kept at 4 °C until confocal imaging (Leica SP8). Primary antibodies used: pErk XP (p44/42 MAPK T202/Y204;

Cell Signalling 4370S; 1:200). Secondary antibody used: donkey anti-Rabbit Alexa Fluor 647 (ThermoFisher; A31573; 1:2000).

3.3 Results and discussion

Ultra-low labeling concentration allows for single molecule tracking

By using a labeling concentration of 5 nM of Janelia Fluor 646 (JF646), in combination with HILO illumination angle [48], we were able to achieve single molecule resolution of Sox2 in the nuclei of living stem cells (fig 3.1). Cells were labeled as described in Methods above, and the final optimal labeling concentration of 5 nM was empirically determined by titration of the dye concentration between 100 nM and 0.5 nM. Optimal labeling was achieved when individual emitters could be distinguished and separately localized, while still maintaining a sufficient level of fluorescence to allow fine-tuning of the focus and HILO angle. Detected Sox2 molecules displayed a variety of behaviors, some staying stably bound within a small volume for many seconds, and others binding shortly before diffusing away or moving out of focus. As is also clear from figure 3.1, photobleaching also was observed as the intensity of the fluorescent signal decreases over time. As a control, we compared time-lapse imaging of Sox2 with relatively immobile histone H2B, and diffusing nuclear localization sequence (NLS) (figure 3.1). With the 500 ms exposure time, the diffusing Halo-tagged NLS molecules mostly appear as a blur, with a few molecules detected as immobile. In contrast, the nucleosomal histones appear as fixed bright spots, showing minimal diffusion and seemingly disappearing from view mainly through photobleaching.

Sox2 residence time reveals stable sub-populations of binding-events

Residence time fitting from transcription factor single molecule traces has previously been described extensively [65–67, 71, 72, 75–77], however the procedure and details vary greatly between studies. The most common procedure is to fit the 1-CDF with the double-exponential decay model, with the argument that this accounts for two modes of transcription factor binding to DNA, specific and non-specific [65, 76]. This model fits our data well on shorter timescales (<25 s, figure 3.2), and yields comparable residence times as previously published; fitting just the first 25 seconds yields $\tau_1 = 0.61 \pm 0.06$s and $\tau_2 = 13 \pm 2$s, similar to Chen *et al.* (2014)[65]. However, this model does not account well for the small but significant portion of single molecule traces that last much longer, some up to hundreds of seconds. We hypothesized that perhaps additional modes of TF binding exist, and that the data would be more adequately fitted by additional residence time components. In figure 3.2 we expanded the model to include between 3 and 5 exponential decay terms (eq. 3.1), and fit them to the logarithmically binned data for model comparison. For the model comparison, we decided to exclude single-molecule tracks consisting of only one detection, since these are more prone to noise and false-positive detections, as described in Mazza *et al.* (2012) [34, 72] (figure 3.9). As an evaluating criteria for the models, we calculated the Akaike information criterion (AIC) and the sample size corrected AIC (AICc) [73, 74] for each fit. AIC and AICc reward fitting the data points while penalizing a model for a higher number of parameters, with the lower AIC value indicating a better fit. Based on the calculated AIC values, we found that the 2-component model is a poor fit of the residence time data, when compared to the three component model, which incorporates a third highly stable sub-population of binding events ($\tau_3 = 55 \pm 6$s) in addition two shorter residence times ($\tau_1 = 1.01 \pm 0.06$s, $\tau_2 = 5.7 \pm 0.8$s). However, we also found that the best model selected based on AIC was highly dependent on the fineness of the histogram binning (figure 3.2B and C). As we increased the number of bins, the four component model became a better fit, which was directly comparable to the three-component model, but with an extra sub-population of $\tau = 17 \pm 4$s. This does compromise our model selection, and we cannot say with confidence whether the three or four component models are a better fit, however both models include the highly stable binding sub-population with comparable residence times ($\tau_4 = 64 \pm 3$s). Regardless of the binning, however, the 5-component, and indeed the 2-component model never yielded an improved fit to either 3 or 4 component

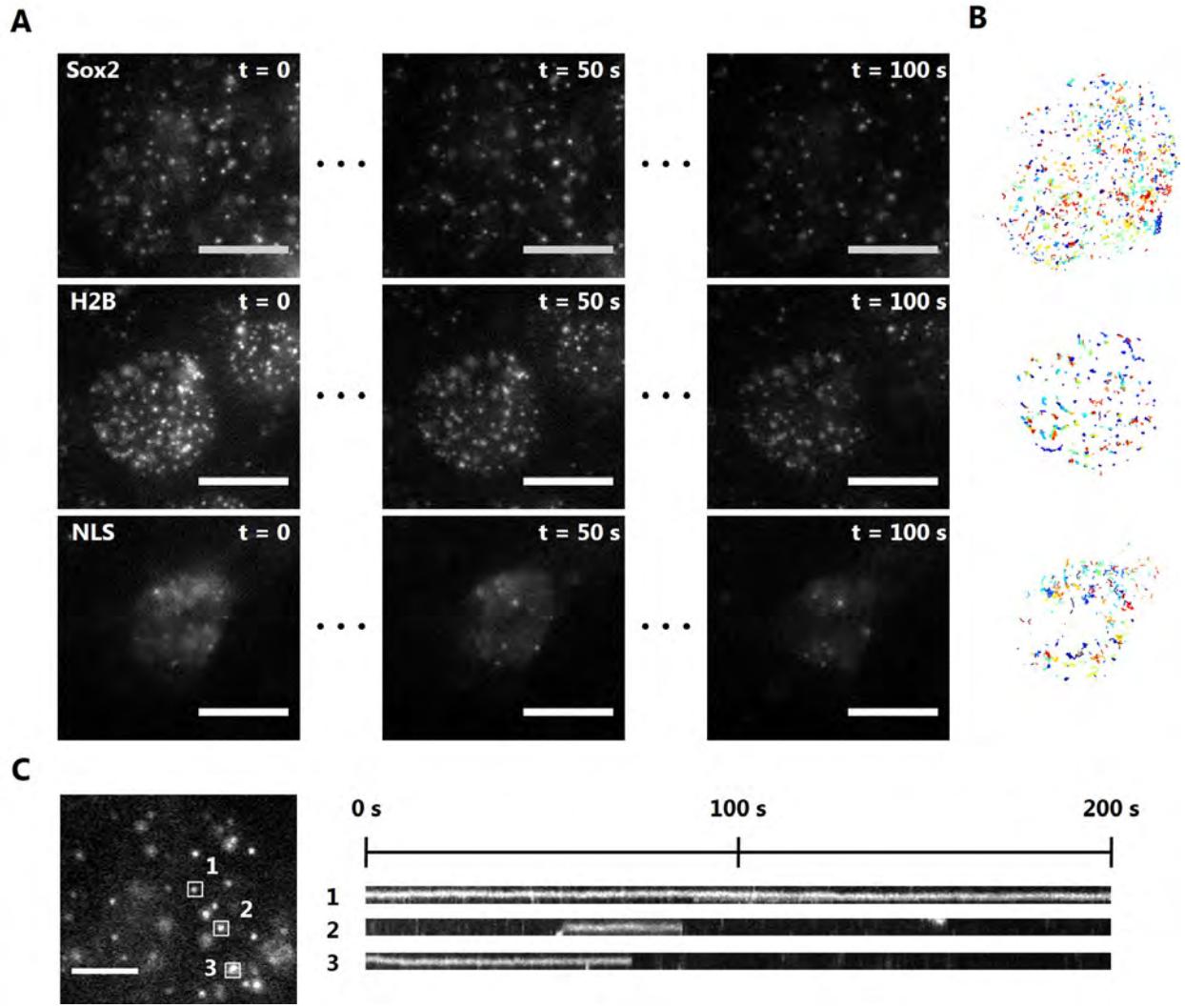

Figure 3.1: Example images from single molecule imaging of Sox2-HT, H2B-HT, and NLS-HT, using 5 nM concentration of the non-photoactivatable fluorophore Janelia Fluor 646, and 500 ms exposure time. **(A)** Frames from single-molecule timelapse imaging of Sox2-HT, H2B-HT, and NLS-HT at times 0, 50 s, and 100 s. **(B)** Traces determined by performing the tracking routine in MATLAB using localizations from ThunderSTORM. **(C)** Three regions with detected Sox2-HT molecules, as indicated by the white boxes, are plotted as a function of time, showing both very stable bindings and shorter traces. Scale bars in **(A)** are 10 μm, and 5 μm in **(C)**.

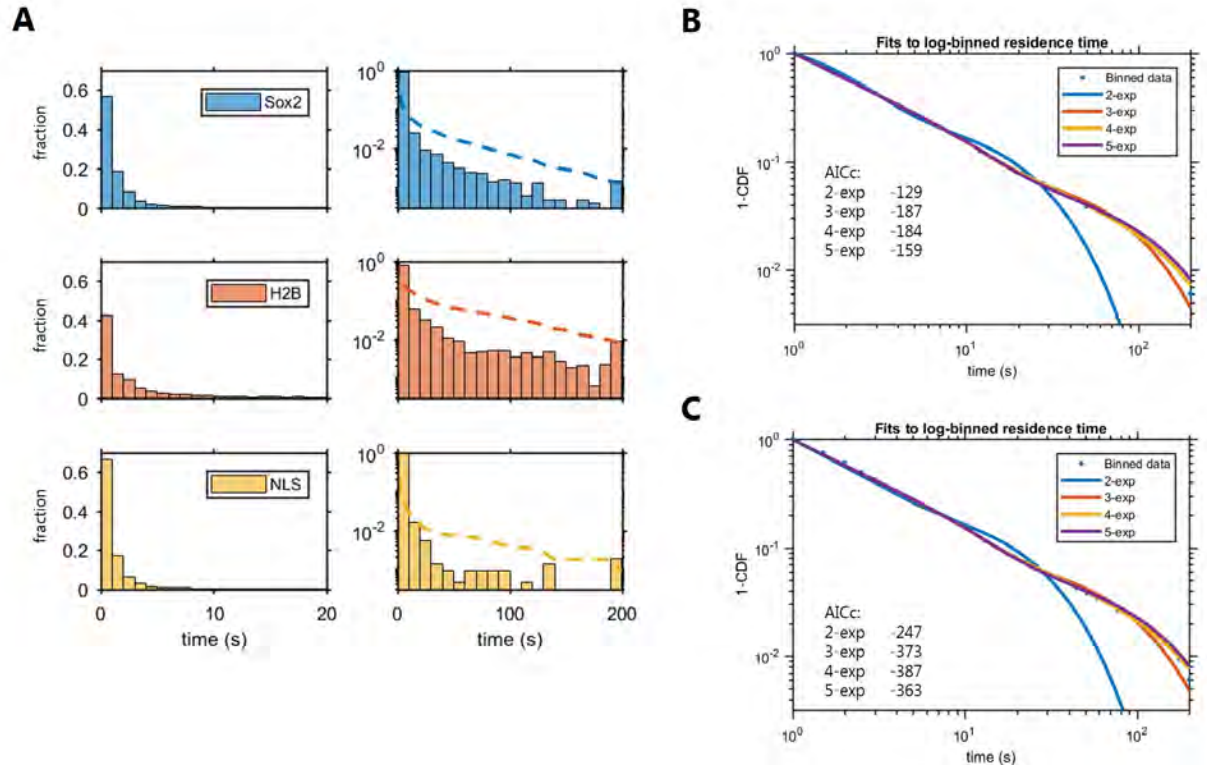

Figure 3.2: Single molecule tracking residence time histograms and 1-CDF fitting. **(A)** Histograms of the residence times (length of each immobile trace within the tracking parameters) plotted on linear and logarithmic y-axes, for Sox2-HT, H2B-HT, and NLS-HT. Consistent with what was shown in e.g. Chen *et al.* (2014)[65], residence times are predominantly within the first 10-20 seconds, however, from the logarithmic plot it is also clear that a small but significant portion of molecule traces last much longer. **(B,C)** Fitting of exponential decay models to the 1-CDF of Sox2-HT. In **(B)** the 1-CDF is binned into 20 logarithmically spaced bins and fitted with exponential models containing between 2 and 5 decay terms (eqn. 3.1), and the sample size corrected Akaike information criterion (AICc)[74] is calculated for each. **(C)** Same as **(B)**, but using 40 logarithmically spaced bins instead. The best model fit contains either 3 or 4 exponential decay terms depending on the binning parameters, however, the 2 or 5 term exponential models never yielded better fits regardless of binning parameters. N = 8 cells, 7260 molecule traces.

models. We can therefore still conclude, that the models including a highly stable sub-population of binding events was consistently a better fit, regardless of analysis. For the further analysis in this paper, we compared Sox2 residence time in different conditions using the 3-component exponential model, since it is the simplest of the two models, which both can fit the data.

Interestingly, a highly stable fraction of Sox2 binding has not been reported in previous studies, and generally studies on TF residence time do not describe comparison of models beyond two-component exponential decay. Papers on TF residence time measurements typically only show the two-component, or sometimes compare the two-component model to a single-component exponential decay [34, 65, 72, 75]. Recently a power-law distribution of TF residence times was suggested by Garcia *et al.* [71], as similar to this study, they found that fitting the two-component model to GR transcription factor binding did not yield a good fit. The power-law model was argued to encompass a spectrum of binding affinities instead of a few distinct modes of binding, however this model did not yield a better fit to our data.

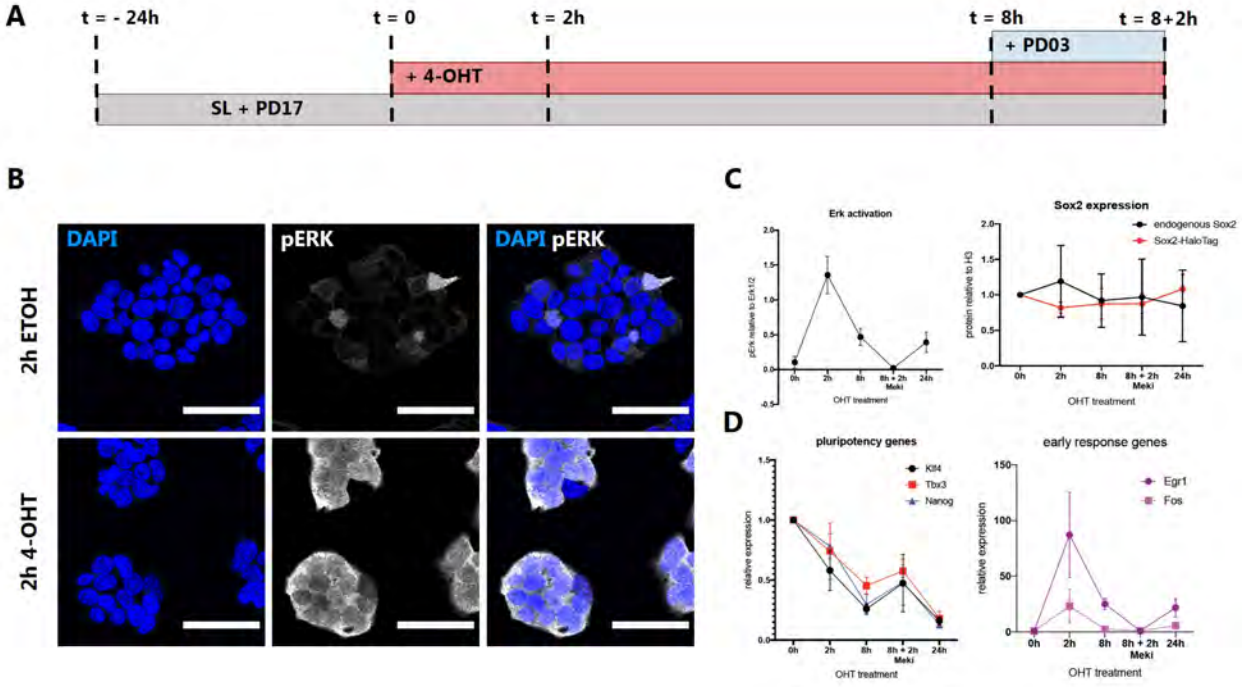

Figure 3.3: Erk activation in Sox2-ERK cell line by tamoxifen (4-OHT). **(A)** Schematic of the ERK activation experiment. FGF signalling is blocked for 24 hours prior to (and during) the experiment by addition of FGFR inhibitor PD17 to the media. At time t=0, ERK activation is started by addition of 4-OHT to the cell media, and cells are imaged after 2 and 8 hours. After 8 hours, the ERK signalling pathway is blocked by inhibition of MEK with PD03 for 2 hours, and a final timepoint is taken at 8+2 hours. **(B)** Immunostaining for phosphorylated ERK in ERK-activated cells (4-OHT) and control cells (ETOH) after 2 hours of treatment. **(C)** pERK and Sox2 levels as determined by Western blots. On the left, the level of pERK relative to ERK1/2 before and after 4-OHT treatment, including after 4-OHT treatment followed by MEK inhibition. To the right is shown how the level of Sox2 expression during the experimental timecourse remained unaffected by ERK-activation. **(D)** qPCR measurements of expression levels of pluripotency genes (left) and early response genes after ERK activation as well as MEK inhibition. A marked decrease in pluripotency genes was seen after ERK activation, which was slightly rescued after MEK inhibition.

Direct activation of ERK shows no significant effect on Sox2 binding

We next sought to answer if directly activating the priming of stem cells towards primitive endoderm would have an effect on Sox2 binding dynamics, detectable by our setup. As described in Hamilton and Brickman (2014) [14], introducing an estrogen receptor activated cRAF construct in stem cells, allows the direct activation of the MAPK/ERK pathway by phosphorylation of ERK via the small molecule drug tamoxifen (4-OHT). This initiates a priming of stem cells towards the primitive endoderm cell lineage as observed in the reduction of pluripotency and epiblast markers, including Nanog, Klf4 and Tbx3 (figure 3.3), followed by an increase in primitive endoderm markers such as Gata6 and Pdgfra. From comprehensive studies by Hamilton *et al.* [14, 17] we know that this ERK activation leads to a drastic reduction of RNAP from the promoters of core pluripotency genes, as determined by chromatin immunoprecipitation followed by sequencing (CHIP-seq). These genes are regulated by Sox2 binding to both promoter and enhancer elements, and while Sox2 stays bound during the priming, other factors of the transcription pre-initiation complex, including RNAP and Mediator dissociate from the region. This suggests that the stability of the complex, and by consequence Sox2 binding, would be affected by transcriptional priming. This was however not observed in CHIP-seq data, on the contrary, it was found that Sox2 and other transcription factors remained bound on regulatory elements of inactivated genes [17].

Notably, these studies were bulk studies of the ensemble behavior of chromatin binding averaged across the whole population of cells and molecules. Thus it is important to address whether the finding that transcription factor binding is unaffected also holds true on the cellular, and indeed molecular level in living cells.

We introduced the ERK activation construct used in [17] to our Sox2-HT cell line, generating the Sox2-ERK cell line. Because cells are grown in block of FGF receptors with the small molecule drug PD17 prior to the experiment, endogenous phosphorylated ERK is kept at minimal level. We could then directly initiate the ERK signaling pathway by activating the cRAF construct with 4-OHT, phosphorylating MEK, which in turn phosphorylates ERK. This gives us precise temporal control of the initiation of primitive endoderm priming, resulting in a homogeneous and consistent response, as shown in figure 3.3B, where a clear and rapid induction of phosphorylated ERK is apparent already after 2 hours of 4-OHT treatment.

We thus performed single molecule imaging of ERK-activated cells at timepoints 2 and 8 hours after activation with 4-OHT, and for each timepoint comparing with a control sample where ethanol (ETOH) is added instead of 4-OHT. Until 8 hours post-activation, the response of cells to ERK-activation can still be reversed if the pathway is inhibited [14], and we therefore added a third timepoint where we follow 8 hours of 4-OHT activation with 2 hours of inhibition of MEK by the inhibitor drug PD03 (figure 3.3A).

Since the time course of the ERK-activation experiment is relatively long (more than 34 hours between cell plating and imaging for the longest timepoint), we had difficulties achieving consistent single-molecule imaging conditions for every timepoint using the labeling strategy used in the first experiments with Janelia Fluor 646 (JF646). However, by labeling Sox2 with the photo-activatable version of the dye, PA-JF646, we were able to obtain consistent single-molecule imaging at all time-points. Switching to the photo-activatable dye had the advantage that the number of fluorescing molecules could be tuned by adjusting either duration or intensity of the activation pulse with (405 nm). Furthermore, since both activation and excitation of fluorophores is done in HILO illumination, background emission is further reduced, and contrast is increased. On the other hand, this complicates photo-bleaching correction, as molecular diffusion will also contribute to the decrease in apparent number of fluorescent molecules in the field of view over time. Only molecules inside the HILO light sheet at the time of photo-activation will be excitable, and in contrast to imaging without photo-activation, new molecules diffusing into the light sheet will not fluoresce.

We observed no clear differences between the residence time distributions when plotted with 3-exponential fitting, averaged over three biological replicates for each timepoint (3.4A). This was even more apparent when we excluded traces of only one detection from the fitting (figure 3.4C). In order to estimate the error on each fit, and estimate if any were significantly different, we performed bootstrapping, i.e. random sampling with replacement, of the residence time data and fitting procedure (figure 3.4B and D). For the raw data including all tracks, we found differences between some conditions of ($0.01 < p < 0.05$), however, as seen from figure 3.4B, only one of these differences were actually found between a timepoint and its control group; the 8+2 hours of 4-OHT and PD03 treatment showed a slight difference in the fraction of stable bindings (**c**). If we compare with the fits done on the data without the shortest tracks, we also observe that this difference is gone.

Although we do detect slight differences between single-molecule experiments, we do not detect any larger differences between a stimulated sample and the corresponding control, than between groups that were not directly compared. We therefore postulate that the activation of ERK does not significantly alter the Sox2 residence time distribution in ESC at a level we can detect.

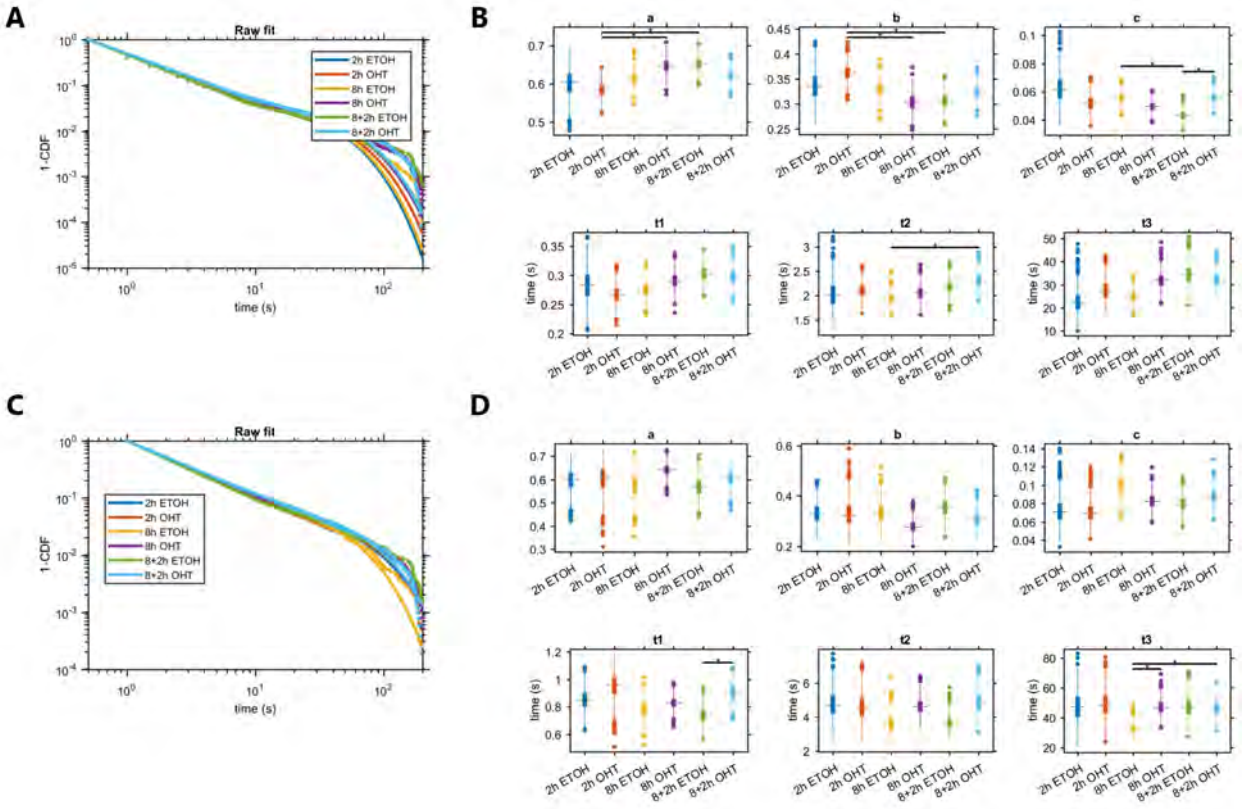

Figure 3.4: Sox2 residence time after activation of ERK **(A)** 1-CDF with 3-exponential fits for each condition, using all traces, including single point detections. **(B)** Fitting of the residence time 1-CDF in **(A)** was bootstrapped (n=1000) to estimate the error on the fits. (a,b,c) represent the fractions of the three populations, and (t1,t2,t3) are their associated characteristic residence times. While all fits remained within the same range of values, some variability was observed both between control groups and different experiments. * indicates $p<0.05$ as calculated from the standard deviation of bootstrap values. **(C)** 1-CDF with 3-exponential fits for each condition, using only traces of 2 or more connected points. **(D)** Bootstrap (n=1000) of the fitting procedure using residence time data plotted in **(C)**. Number of cells; 17, 29, 28, 27, 18, 20 from left to right on the box-plots.

Disruption of transcription initiation complex reveals Sox2 binding dynamics dependence on BRD4, but not p300 acetyltransferase activity

Recent studies have suggested the presence of highly stable transcription clusters, forming liquid phase condensates around large aggregates of transcription co-activators including BRD4 and Mediator [35–37]. In Cho *et al.* (2018)[35], they discovered highly stable clusters of both RNAP and Mediator lasting longer than 100-150 seconds. They found that loss of BRD4 from enhancers through treatment with the small molecule inhibitor JQ1, disrupted these super-stable clusters.

We hypothesized, that perhaps the highly stable events we see in our data could at least in part be explained by large-scale clustering of transcription co-factors, and perhaps liquid-phase condensation. If stable Sox2 binding is related to these highly stable condensates, it would be interesting to see how and if Sox2 binding is affected by BRD4 loss from enhancers. Since dissociation of RNAP and Mediator from chromatin at Sox2 regulated genes as observed with ERK activation did not cause significant change in Sox2 binding, it is not clear how Sox2 would be affected by RNAP and Mediator loss through BRD4 inhibition.

We probed Sox2 residence time with and without inhibition of BRD4 by JQ1, and found a significant and clear change

in the distribution (figure 3.5). While we observe small changes in both the short and medium binding populations, the stable sub-population decreased significantly in terms of fraction of molecules, but not in characteristic residence time.

Although dissociation of RNAP and Mediator by endogenous ERK signaling did not have an effect on Sox2 binding, it appears that dissociation through BRD4 inhibition does have a significant effect. This suggests that Sox2 is not stabilized by RNAP and Mediator at the pre-initiation complex, but that BRD4 does have a stabilizing role, perhaps through its role in forming liquid-phase condensates as shown by Sabari *et al.* [37].

Furthermore, we examined how acetylation activity of acetyltransferase p300 (p300) would affect Sox2 residence time. p300 dependent acetylation of histone H3K27 is an important marker of enhancer activity, keeping the chromatin open for transcription factor binding. We asked whether the inhibition of p300 acetylation activity[78] would affect Sox2 residence time. Interestingly we don't see any change in Sox2 binding (figure 3.6), however this may be due to Sox2's role as a pioneer factor, enabling it to bind to closed chromatin [27].

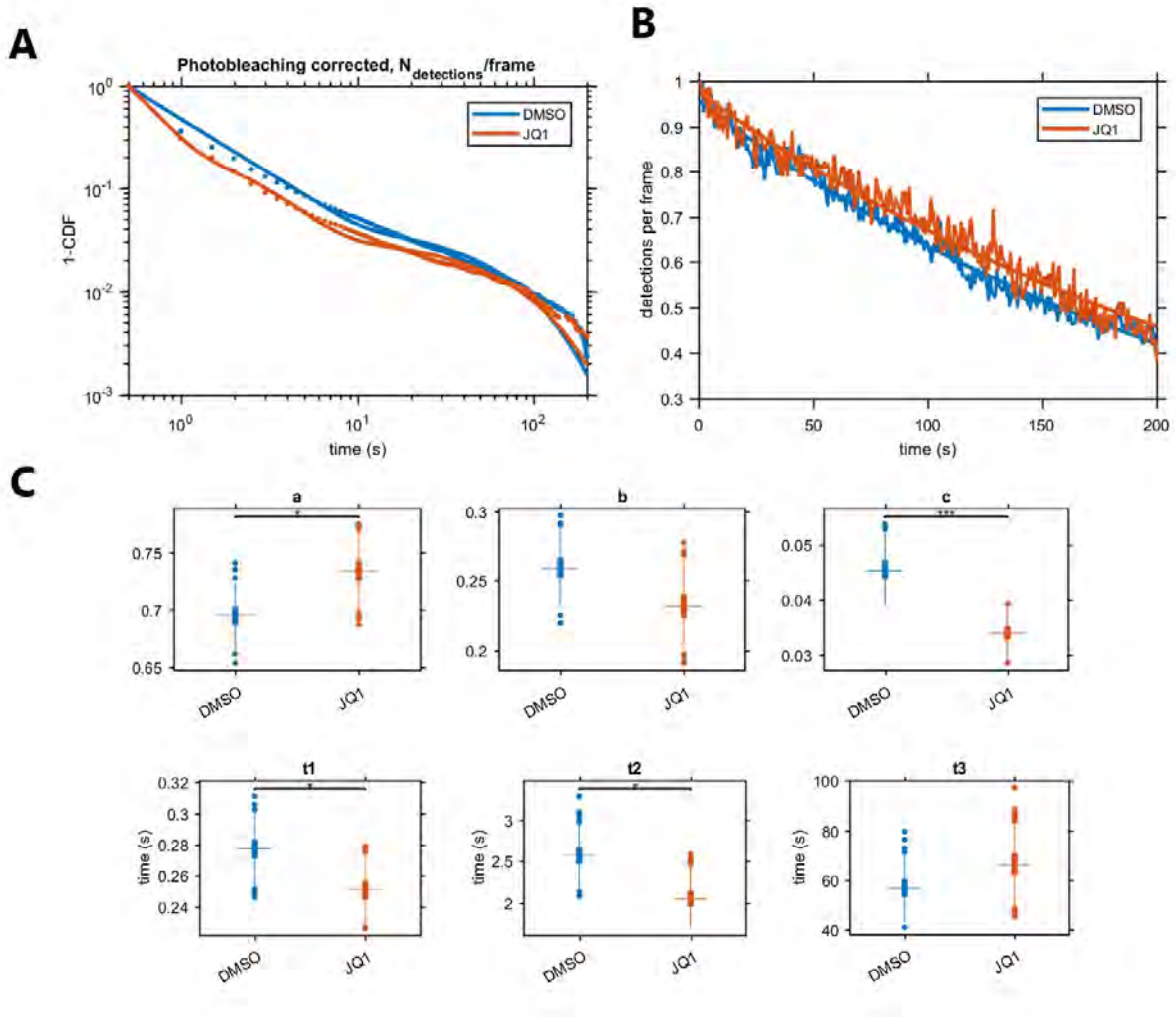

Figure 3.5: Sox2 residence time after inhibition of BRD4 by JQ1 **(A)** 1-CDF with 3-exponential fits for each condition, using traces of at least 2 connected points. **(B)** Photobleaching correction using number of detections per frame. **(C)** Bootstrap fitting (n=1000) of JQ1 and DMSO residence times (* indicates p<0.05, *** indicates p<0.001). (n=19 cells for DMSO, n=16 cells for JQ1).

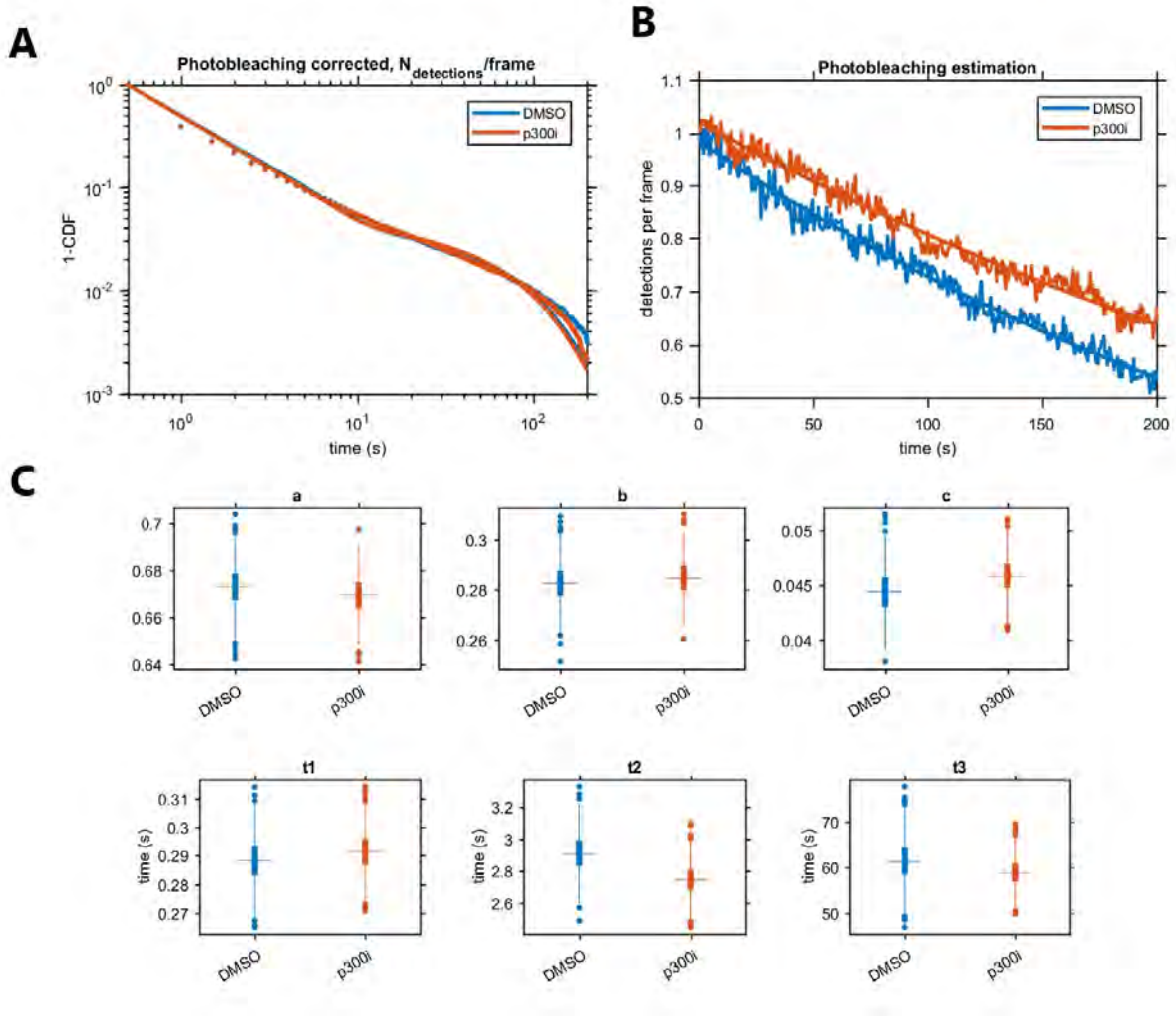

Figure 3.6: Sox2 residence time after inhibition of p300 acetylation by p300i **(A)** 1-CDF with 3-exponential fits for each condition, using traces of at least 2 connected points. **(B)** Photobleaching correction using number of detections per frame.**(C)** Bootstrap fitting (n=1000) of p300i and DMSO residence times. (n=23 cells for DMSO, n=28 cells for p300i).

3.4 Conclusion

In this study, we have performed single-molecule tracking of endogenously tagged Sox2 in living embryonic stem cells. We showed that the residence time distribution of Sox2 binding is not consistent with previously described two-component models of Sox2 binding. An extra term was needed to account for the small but consistent population of highly stable binding events of Sox2.

Interestingly, we found that the direct activation of the ERK pathway, inducing primitive endoderm priming in ESC did not have a clear effect on Sox2 binding, despite decommissioning of pluripotency genes and dissociation of RNAP and Mediator. The inhibition of acetylation activity of p300 did not have any detectable effect on Sox2 residence times, indicating that chromatin accessibility by acetylation is not an obstacle for Sox2 binding as a pioneer factor.

Finally, we found that the inhibition of BRD4, implicated in formation and stabilization of large-scale clustering of transcription co-activators forming liquid phase condensates, had a significant effect on the fraction of stable Sox2 binding events.

Author Contributions

K.G.H. wrote the first draft. K.G.H. built the super-resolution setup, and designed the single-molecule imaging procedure. K.D. designed and cloned the cell lines. K.D. and K.G.H. performed the single-molecule imaging. K.G.H. performed the data analysis of single-molecule tracking, K.D. performed data analysis of western blots and qPCR data. L.B.O. and J.M.B. supervised the project. K.G.H. and K.D. edited the manuscript.

Acknowledgments

This work is financially supported by Danish National Research Foundation, grant DNRF116 (StemPhys).

3.5 Supplementary figures

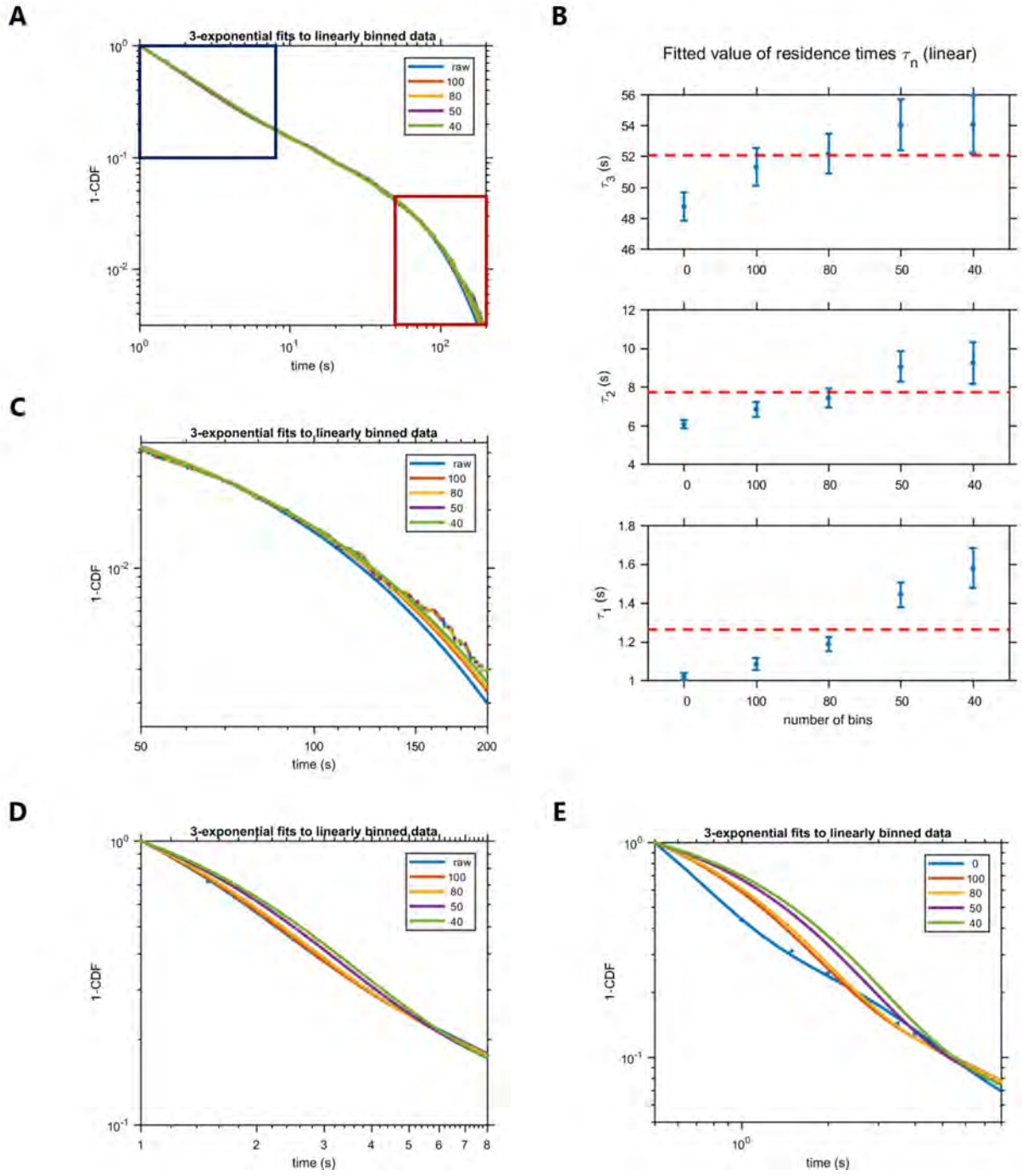

Figure 3.7: Effect of linear binning on 3-exponential fitting. The 1-CDF is calculated from the tracking data residence times excluding traces of only one point, and is either plotted as the raw function values, or binned into N = 40,50,80,100 linearly spaced bins. **(A)** The 1-CDF overlaid with 3-exponential fits to each of the binned data sets. **(B)** The fitted values of the three residence times determined from the fitting in **(A)** (0 bins represents the raw 1-CDF values). The red line indicates the mean of the fitted values, and errorbars represent 95% confidence intervals. **(C)** Same as **(A)**, but zoomed in on the later part of the fit, represented by the red box in **(A)**. **(D)** As in **(C)**, but this time zoomed in on the short residence times indicated by the blue box in **(A)**. **(E)** Same as **(D)**, but instead the 1-CDF is calculated also including traces of just one detection.

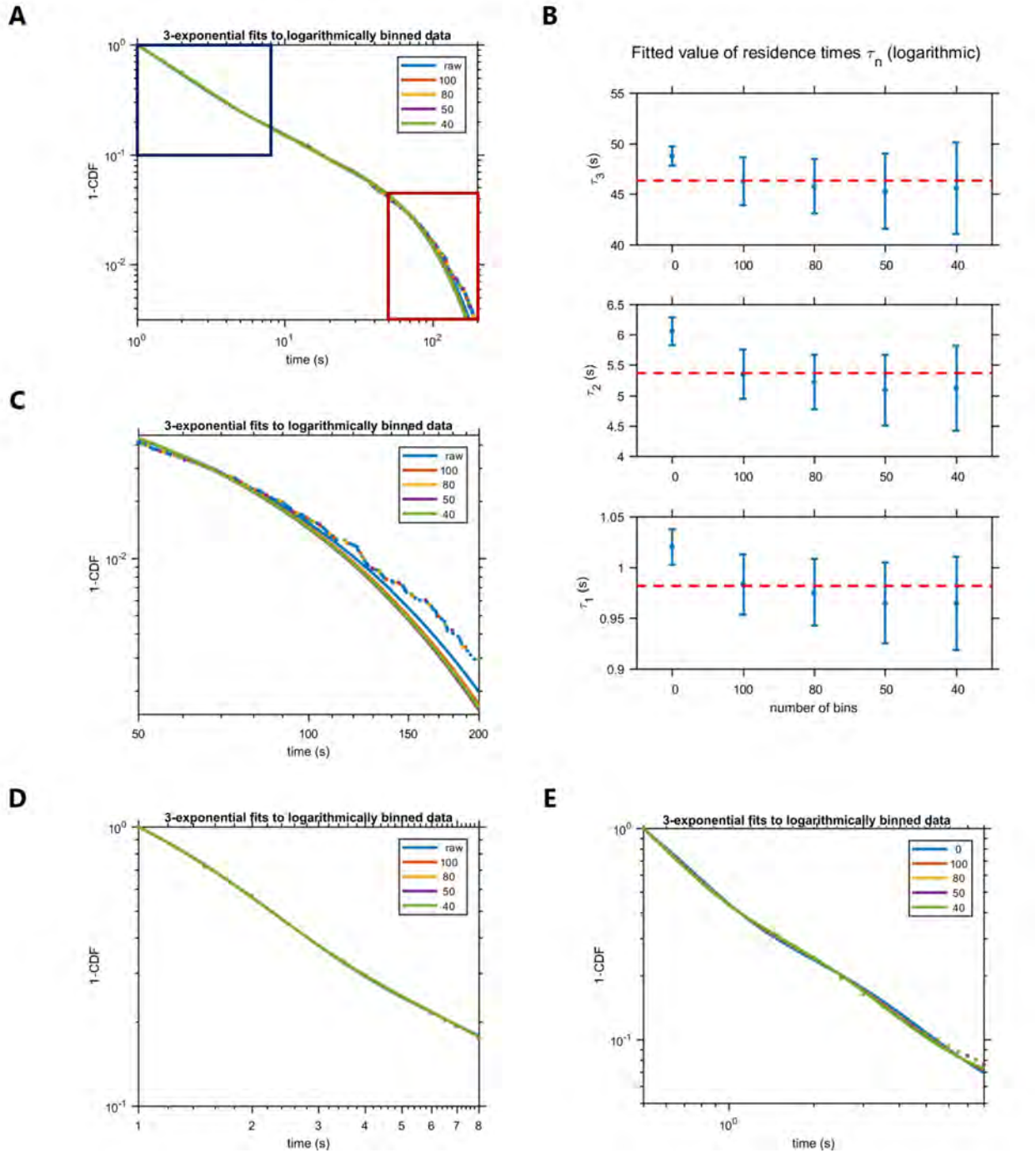

Figure 3.8: Effect of logarithmic binning on 3-exponential fitting. The 1-CDF is calculated from the tracking data residence times excluding traces of only one point, and is either plotted as the raw function values, or binned into N = 40,50,80,100 logarithmically spaced bins. As in figure 3.7, **(A)** shows the 1-CDF overlaid with 3-exponential fits to each of the binned data sets, with zoom-ins on the late times (red box) in **(C)** and the early times (blue box) in **(D)**. **(B)** The fitted values of the three residence times determined from the fitting in **(A)** (0 bins represents the raw 1-CDF values). The red line indicates the mean of the fitted values, and errorbars represent 95% confidence intervals. **(E)** Same as **(D)**, but instead the 1-CDF is calculated also including traces of just one detection.

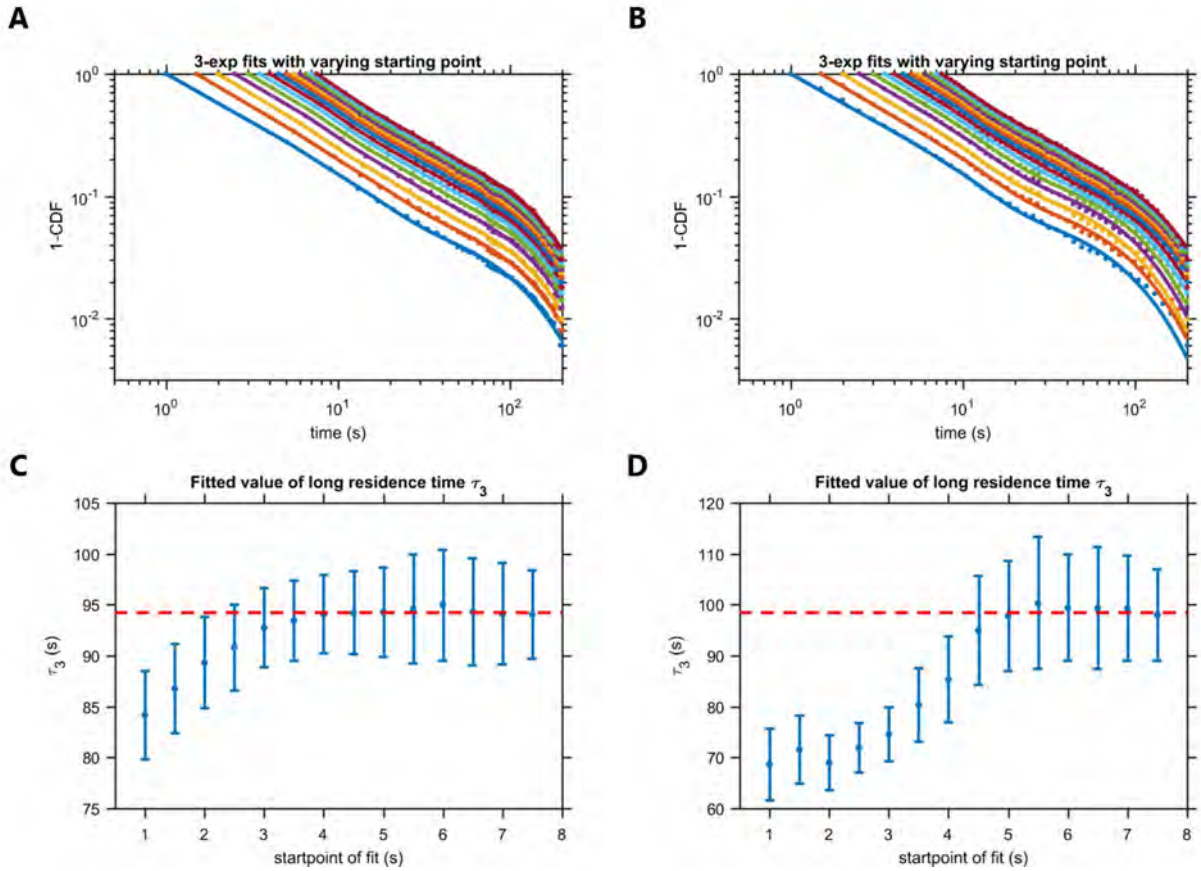

Figure 3.9: Effect of varying the fitting startpoint on 3-exponential fitting as described in Mazza *et al.* (2012)[72] and Hansen *et al.* (2017)[34]. A-B) Fits to 1-CDF of Sox2 residence time, fitting only traces above a threshold minimum length. Data is linearly binned (A) or logarithmically binned (B) with a bin number of 50. C) Shows the long residence time τ_3 resulting from the fits in (A) with the red line indicating the mean value from $t = 3.5$s to $t = 7.5$s, where the fitted value reaches a plateau. D) Fitted long residence times τ_3 from the fits in logarithmically binned (B). Red line indicates mean of fit values after they reach a plateau around $t = 4.5$s. Errorbars in (C) and (D) are 95% confidence intervals from MATLAB's nlparci function.

Chapter 4

Probing the actin cytoskeleton and viscoelastic properties of stem cells in two early stages of differentiation at sub-cellular resolution

Kasper G. Hvid[1], Younes F. Barooji[1], Irene Isturiz[1], Lene B. Oddershede[1], Poul M. Bendix [1]

4.1 Abstract

The cellular cytoskeleton provides the cell with mechanical rigidity and allows mechanically interaction between cells and with the extracellular environment. The actin structure plays a key role in regulating cellular behaviors like motility, cell sorting or cell polarity. From the earliest stages of development, in naive stem cells, the critical mechanical role of the actin structure is becoming recognized as a vital cue for correct segregation and lineage control of cells and as a regulatory structure that controls several transcription factors. The ultrastructure of the earliest embryonic stem cells has not been investigated in living cells despite the well known morphological shape changes cells undergo during the earliest stages of development. Here, we provide the first 3D investigations of the actin cytoskeleton of naïve mouse embryonic stem cells (mESCs) using super resolution optical reconstruction microscopy (STORM) and give a quantitative description of the morphological, cytoskeletal and mechanical differences appearing between cells at the earliest stages of inner cell mass differentiation, as recapitulated by cells cultured under two media conditions 2i-LIF and serum-LIF. The high resolution images of stem cells show that the peripheral actin structure undergoes a dramatic change between the two media conditions and changes from being predominantly oriented parallel to the cells surface in 2i-LIF medium to a more radial orientation in serum-LIF. Finally, using microrheology on living stem cells we also detect mechanical differences in the cell cortex between the cells cultured in these two media.

[1]Niels Bohr Institute, University of Copenhagen

4.2 Introduction

Cells sense their physical environment (rigidity, composition), and biomechanical signals and properties play an important role in cell organization, migration and differentiation. During development, the physical properties of stem cells change drastically in terms of viscoelasticity and morphology [79], and the rigidity of the environment has been shown to be crucial in the regulation of stem-ness of cells and for the development of organoids [80]. The viscoelasticity of cells is also important in disease, for example in cancer where invasiveness has been linked to stiffness, and adaptability to extracellular matrix stiffness [81, 82].While the exact mechanisms and effects of both mechanical properties and mechanosensing among cells are not yet well understood, it may be just as important to map the changes in physical properties of cells during differentiation and development as it is to map the differential expression of transcription factors.

The cytoskeletal nano–architecture and viscoelastic properties of differentiating and committed differentiated cells has been extensively investigated [25, 79, 83, 84]. Much less is known about the cytoskeletal structure of early embryonic stem cells despite the fact that stem cells are known to be mechanically softer and could therefore be structurally different [79]. In particular embryonic stem cells (ESC) exhibit dramatic changes in morphology and mechanics during the first few days after fertilization. Cultured ESCs can be maintained in a ground state of more and less naive pluripotency by maintaining certain culture conditions, and their exact stage can be detected by associated expression levels of transcription factors [5, 85].

Cells cultured in serum-free media with the addition of two small molecule inhibitors of GSK3 and MEK along with leukemia inhibitory factor (LIF) resemble the early naive epiblast of the pre-implantation embryo, around 3.5 days post fertilization [7, 16, 22]. These cells exhibit a greater level of pluripotency and homogeneity than cells grown in serum media supplemented with LIF [7, 21, 22] and exhibit a remarkably different morphology when grown on a substrate. To our knowledge no study has quantitatively investigated the structural and physical properties of cells cultured in these two widely used culture conditions.

Here we investigate mouse embryonic stem cells, which recapitulate cells from the developing embryo, however, can also be cultured and expanded in vitro, and studied in single cell detail. To address early development, we here culture ESC in two well-described stem cell media conditions; a complete media based on leukemia inhibitory factor (LIF) and fetal bovine serum (FBS) referred to as Serum-LIF (SL) and the second a serum-free media based on the addition of two small molecule inhibitors of GSK3 and MEK along with LIF, known as 2i-LIF (2i)[7, 16].

Both conditions mimic cells in the inner cell mass (ICM) of the mouse blastocyst, about 3.5 days post-fertilization and just before the ICM starts differentiating into epiblast and primitive endoderm lineages (figure 4.1B). However, cells grown in 2i media show biological characteristics consistent with more naïve stage than cells grown in SL [22, 86]. This specific stage of blastocyst development is particularly interesting from a biomechanical perspective, as the differentiation into epiblast and primitive endoderm is accompanied by a physical segregation of the cell lineages into an inner and outer layer. In this study we aimed to describe whether on the single cell level, there would be differences in either actin cytoskeleton structure or viscoelastic properties of cells grown in either 2i or SL.

To address these questions, we used both confocal imaging combined with super resolution fluorescence imaging of the actin structure. Confocal imaging is used to characterize the 3D shape of ESC colonies grown under these culture conditions. But while confocal imaging gives important information on the morphology of cells and the overall distribution of actin in the cell, the optical resolution restricts it from being able to resolve the detailed structure of crowded filament networks. To overcome this limitation and describe the nanoscale structure of subcellular actin, we implemented stochastic

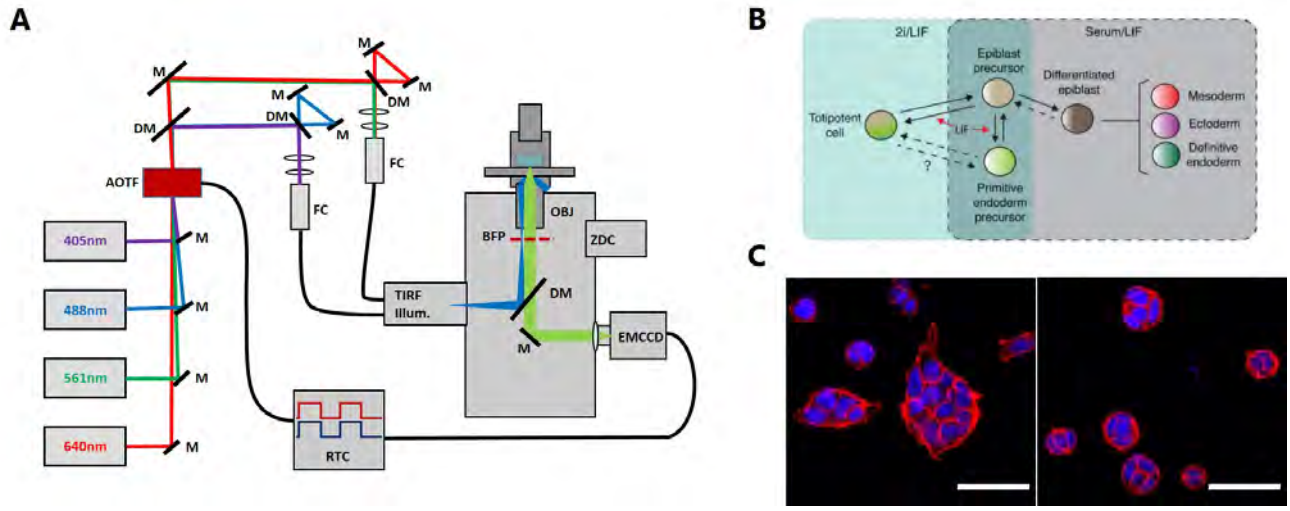

Figure 4.1: Schematic of the experimental setup and biological system investigated. **(A)** Schematic of the custom built TIRF setup for super-resolution imaging. *M*, mirror, *DM*, dichroic mirror, *AOTF*, acousto-optical tunable filter, *FC*, fiber coupler, *TIRF Illum.*, motorized TIRF illumination module, *BFP*, back focal plane, *OBJ*, objective, *ZDC*, real-time z-drift compensation module, *EMCCD*, electron-multiplying CCD camera, *RTC*, real-time controller. **(B)** Cells in 2i recapitulate a slightly earlier and more naïve stage of development than SL cells. Figure adapted from Morgani *et al.* (2013) [86]. **(C)** Confocal images of F-actin (red) and nucleus (blue) in colonies of cells in either SL (left) or 2i (right). Scale bars are 50 microns.

optical reconstruction microscopy (STORM), using the stochastic blinking of fluorophores linked to actin and individually detect and localize these molecules with nanoscale precision, thus mapping out the structure actin network [40–42, 47].

The viscoelastic properties of live cells are probed by optical trapping setup to record the (sub)-diffusive movements of individual lipid granules residing at the periphery of cells. This method allows very precise measurements of the mean-squared displacement (MSD) and quantification of the exponent of anomalous diffusion α, which allows us to characterize the local viscoelastic properties of the macro-molecular cellular cortext.

4.3 Methods

Cell culture

E14Ju mouse embryonic stem cells (ESC) were provided by the Brickman lab at the Novo Nordisk Foundation Center for Stem Cell Biology (DanStem, University of Copenhagen, Copenhagen, Denmark). ESC were maintained in either full stem cell media with serum and LIF (SL) or serum-free 2i-LIF media (2i), for no more than 25 passages, and were regularly tested for mycoplasma. SL medium consisted of Glasgow's Minimum Essential Media (GMEM, Sigma-Aldrich) supplemented with 10% Fetal Bovine Serum (FBS, Gibco), 2 mM L-Glutamine (Gibco), 1 mM Sodium Pyruvate (Gibco), 0.1 mM 2-mercaptoethanol (Sigma-Aldrich), 0.1 mM Non-Essential Amino Acids (Gibco), and 1000 units/mL LIF (DanStem, Copenhagen, Denmark).

2i medium consisted of a 1:1 mix of DMEM F/12 (Gibco) and Neurobasal Medium (Gibco), supplemented with 1X N2 (Gibco), 0.5X B27 (Gibco), 0.1 mM 2-mercaptoethanol (Sigma), 1000 units/mL LIF (DanStem), 3 µM Chir99021 (Sigma), and 1 µM PD0325901 (Sigma). Cells were grown on flasks (Corning) coated with 0.1% gelatin in PBS (Sigma) in 37 °C incubators containing 5% CO2, and were passaged with DPBS (Sigma) and Accutase (Sigma).

Cell fixation and staining

For confocal and STORM imaging, cells were plated on fibronectin coated #1.5H thickness 8-well glass bottom slides (Ibidi) Fibronectin coating was performed by covering the glass slides in a solution of DPBS with 10 $\mu g/mL$ human fibronectin (EMD Millipore) for at least 2 hours. Cells were then plated at a density of 20,000 cells/cm^2, and left growing at culture conditions overnight. For fixation, each well was washed once with DPBS before adding 4% paraformaldehyde in PBS for 10 minutes. After fixation, cells were again washed and stored in DPBS at 5 $^{\circ}C$ until staining.

Before staining, cells were permeabilized using 0.1% Triton X-100 (Sigma-Aldrich) in PBS for 15-30 minutes. After washing in PBS, blocking was performed by incubating the sample with 1% BSA in PBS for 30 minutes at room temperature. Cells were then stained with Alexa Fluor 647 Phalloidin (Invitrogen) at a 40X dilution of the stock solution for 20 minutes at room temperature. The cells were again washed in PBS before finally staining chromatin with the DAPI stain NucBlue Fixed Cell ReadyProbes Reagent (Invitrogen).

Confocal microscopy

Fixed and stained cells were imaged on a Leica SP5 confocal microscope (Leica Microsystems, Wetzlar, Germany) using a 63X water-immersion objective (NA=1.20, Leica).

Optical trapping experiments were carried out on the same confocal setup which had a 1064 nm laser coupled in through the side port. The setup is described in Andersen *et al* [87]. The focused laser was positioned on single granules located at the cell periphery as detected by bright field imaging. The laser light was collected by a high numerical aperture oil immersion condenser and the local diffusion of a granule was measured from the scattered light detected by a quadrant photodiode (QPD) located at the back focal plane of the condenser.

STORM imaging and analysis

For STORM imaging we used a custom built total internal reflection (TIRF) setup based on the Olympus IX83 inverted microscope (Olympus), equipped with a 150X 1.45 NA oil-immersion objective (Olympus), EMCCD camera (Hamamatsu), Z drift compensation system (IX3-ZDC2, Olympus), and a motorized TIRF module (cellTIRF, Olympus). For STORM imaging we individually fiber coupled a 640 nm laser (Toptica Photonics) and a 405 nm laser (Olympus) to the TIRF module, and independently set both illumination angles to a TIRF penetration depth of 100-200 nm for imaging.

To induce stochastic blinking of the Alexa Fluor™ 647 fluorophore, imaging was performed in a PBS based imaging buffer based on the *d*STORM buffer by van de Linde and others [42] containing 100 mM MEA (Sigma-Aldrich), 0.6 mg/mL Glucose Oxidase (Sigma-Aldrich), 10% (w/v) Glucose (Sigma-Aldrich), and 60 µg/mL Catalase (Sigma-Aldrich) in PBS. The pH of the buffer was adjusted to between 7.5 and 8.5 by addition of KOH.

Timelapse imaging of blinking events was performed at 20 ms exposure time for 10000-30000 images. Individual single molecule blinking events were then localized in the timelapse images using the ThunderSTORM [53] plugin in Fiji [52], and drift was corrected using the cross-correlation function in ThunderSTORM.

3D STORM imaging

3D STORM imaging was done as described above, with the important difference being the insertion of a cylindrical lens in the optical pathway before the camera. This produces astigmatism in the detected emission from single molecule blinking events, elongating them in either the vertical or horizontal direction depending on their z-distance from the focus plane [88].

The elongation was first correlated to z-position using surface-fixed beads and scanning them by moving the focus through z. 3D localization of blinking events was determined using ThunderSTORM and the elongation-correlation dataset.

Image analysis

All quantification was done using custom written scripts in MATLAB (MathWorks) and ImageJ [51]. The contact angles of figure 4.2C were calculated as the angle between the substrate (horizontal) and the edge of the cell colonies. For each colony, the angle was measured along the entire circumference of the colony and averaged.

Contact area in figure 4.3E was measured as the total area covered by actin for individual cells, as detected within the TIRF volume. The aspect ratios of figure 4.3F was calculated as the ratio of the lengths of the minor axis and major axis of individual cells. For the pixel orientation map of figure 4.5, STORM images were filtered using an implementation of a Gabor filter to enhance the spatial resolution of linear structures in images [89, 90] (see figure 4.7). After image enhancement, the local orientation angles of actin filaments in every pixel were extracted based on the structure tensor method using the freely available ImageJ plugin OrientationJ [51]. Figure 4.5A and B show the color-coded angle distribution for two cells cultured in SL and 2i media. The histograms of actin orientation (figure 4.5 C and D) indicate that actin filaments mostly are orientated in the direction of the cell angle for a cultured cell in SL and distributed randomly for a cell in 2i. This observation can be quantified as an orientation order parameter, $S = \langle \cos(2\theta) \rangle$ where θ is the local angle deviation of filaments at each pixel from the major direction of the cell [91, 92].

Data from the laser tracking of granules recorded by the QPD were analyzed by custom made MATLAB code based on the software published in [93]. Briefly, the time series were Fourier transformed and the slope of the high frequency part of the power spectrum was extracted by fitting and was used to find the scaling exponent α in eq. 4.1.

4.4 Results

Morphological differences in actin structure of colonies of stem cells

To address the question of cytoskeletal differences, we first performed fixed cell confocal imaging of colonies of cells in three dimensions. Cells were labeled for F-actin and with a nuclear stain (figures 4.1C and 4.2A,B). Firstly, we observed that actin is mainly localized in the cell cortex in both conditions, however SL cells show a higher number of filipodia and actin protrusions along the surface. Secondly, colonies of ESC in 2i grow in rounded dome-like colonies as has previously been observed [24, 94], as opposed to the flat and more spreading colonies of cells in SL, which is especially apparent from the x-z projection (figure 4.2A,B). Interestingly, we noted that at the colony edges, the contact angle of cells adhering to the substrate, in the two conditions were drastically different. Whereas cells in SL seemed to maximize their contact with the surface by spreading out towards the substrate, cells in 2i narrow at the substrate interface. This is readily apparent when we quantify the contact angle, taken as the angle between substrate and colony edge, averaged over the circumference of the colony (figure 4.2C). Cells in 2i attach at an obtuse angle of $\theta = (115 \pm 3)^\circ$, while cells in SL attach at an acute angle of $\theta = (69 \pm 9)^\circ$.

Nanoscale actin structure in stem cells

With the aim of probing the nanoscale organization of F-actin, we implemented STORM imaging at a depth of a few hundred nanometers using TIRF illumination (figure 4.3). Complementing the observations from confocal imaging, STORM

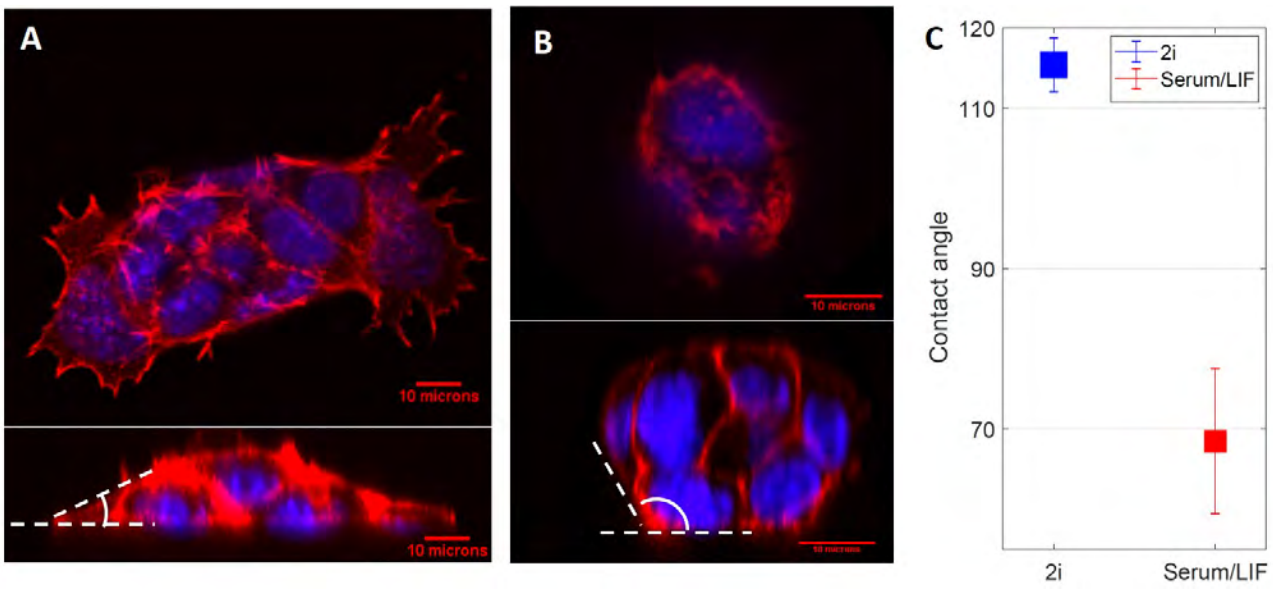

Figure 4.2: 3D confocal images of colonies of ESC grown in 2i or SL. **(A,B)** Top view (top) and side view (bottom) of typical cell colonies of cells grown in SL **(A)** or 2i **(B)**, with the contact angle of cells to the substrate marked in white. The scale bar is 10 µm in all images. **(C)** The contact angle of ESC to the surface in 2i and in SL, extracted from the side-view images as shown in **(A)** and **(B)**, n=5 colonies for each condition.

imaging reveals that cells in SL have concentrated actin filaments in dense stress fibers and bundles, mainly located on and along the periphery of the cell adhesion to the substrate, especially in filipodia. In the center of the SL cell adhesion area, we do not observe connected actin filaments (figure 4.3A,C), indicating either a lack of connected filaments, or that filaments are reaching the TIRF volume at an angle far from parallel to the surface. In comparison, actin in 2i cells is evenly distributed throughout the cell cortex at the substrate interface, and connected filaments are apparent even within the cell circumference, indicating filaments parallel to the surface in contrast to SL cells (figure 4.3B,D). In order to compare the spreading of the two stem cell stages, we quantified the contact area of cells with the substrate, which we defined as the total area of cells within the TIRF volume imaged by STORM (figure 4.3E). We found that ESC grown in SL media show significantly larger surface attachment than cells in 2i, corresponding well with our measurements of contact angle by confocal imaging. Both results indicate a tendency for SL cells to maximize their surface area by spreading along the substrate. We further quantified the aspect ratio of individual cells, finding that SL cells also tend to be more elongated than cells in 2i (figure 4.3F).

3D STORM imaging reveals dorsal stress fibers in Serum-LIF but not in 2i-LIF

From the imaging of colonies by STORM, we made the interesting observation that colonies in 2i typically revealed clear gaps between cells (figure 4.3B). This was not observed in our confocal imaging in the same conditions, indicating that the actin in individual cells in 2i are curving inward towards the center of the cell at the substrate. Since this curving was not detectable by confocal, we decided to implement 3D STORM to resolve it (figure 4.4). In 2i media, we did indeed see the actin curving away from the surface at the edges of the cell (figure 4.4B), contrasting cells in SL where actin bundles can clearly be seen pointing towards the edges and protrusions of the cell, coming from above and attaching near the substrate (figure 4.4A). This is consistent with dorsal stress fibers, attaching to the substrate at the leading edges of the cell [95, 96]. In 2i media on the other hand, we do not see clear evidence of dorsal stress fibers, instead observing an interconnected

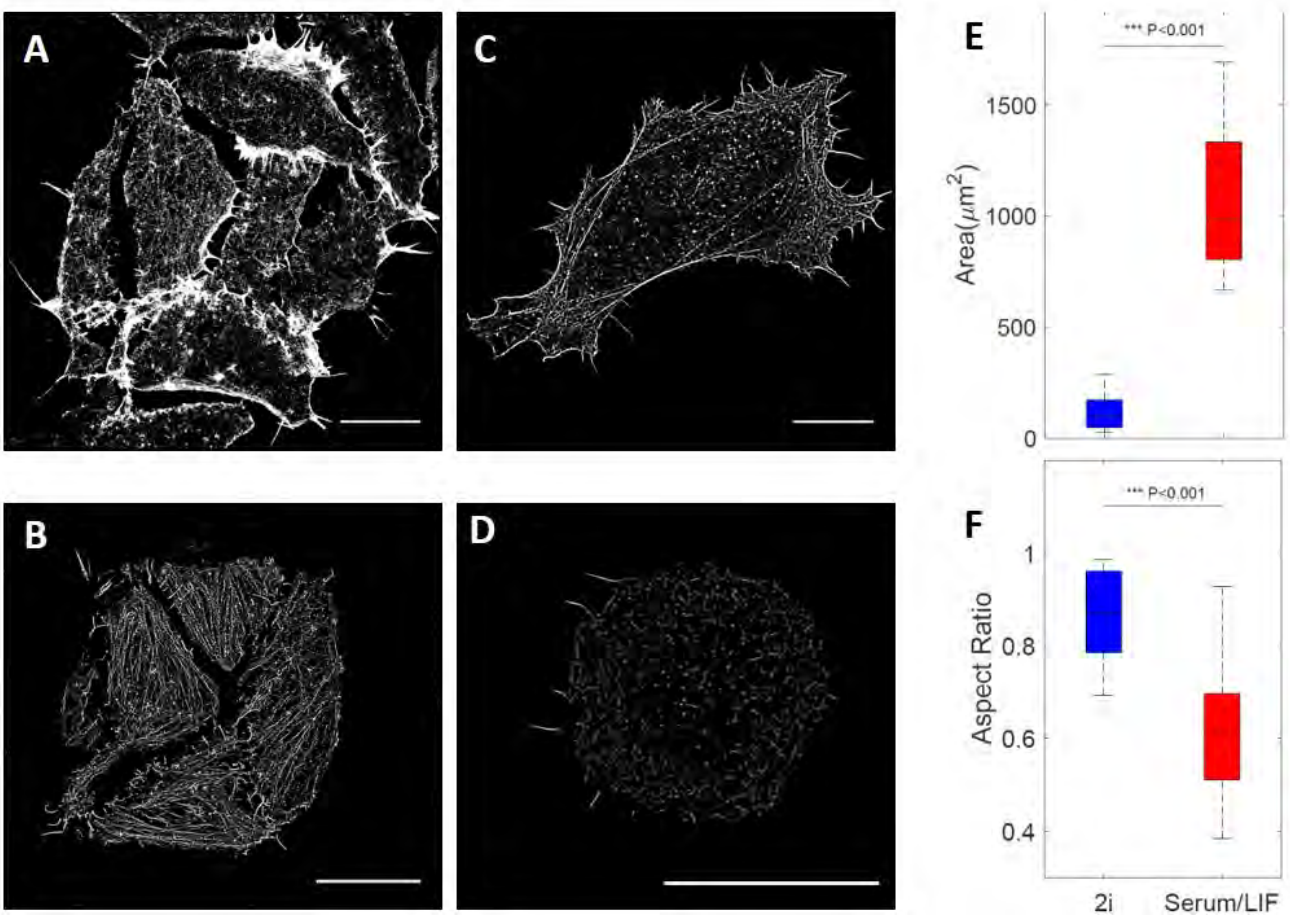

Figure 4.3: Super resolution images of the actin cytoskeleton organization in stem cells near the substrate. STORM images of the actin network in; **(A)** a colony of cells grown in SL media, **(B)** a colony of cells grown in 2i media, **(C)** a single cell grown in SL media, and **(D)** a single cell grown in 2i media. All scale bars are 10 μm. **(E)** Boxplot of the measured surface spreading area of stem cells grown in either 2i or SL media reveals significantly larger spreading in SL cells (n=16) (F) Boxplot of measured aspect ratios of stem cells grown in 2i or SL media (n=16).

.

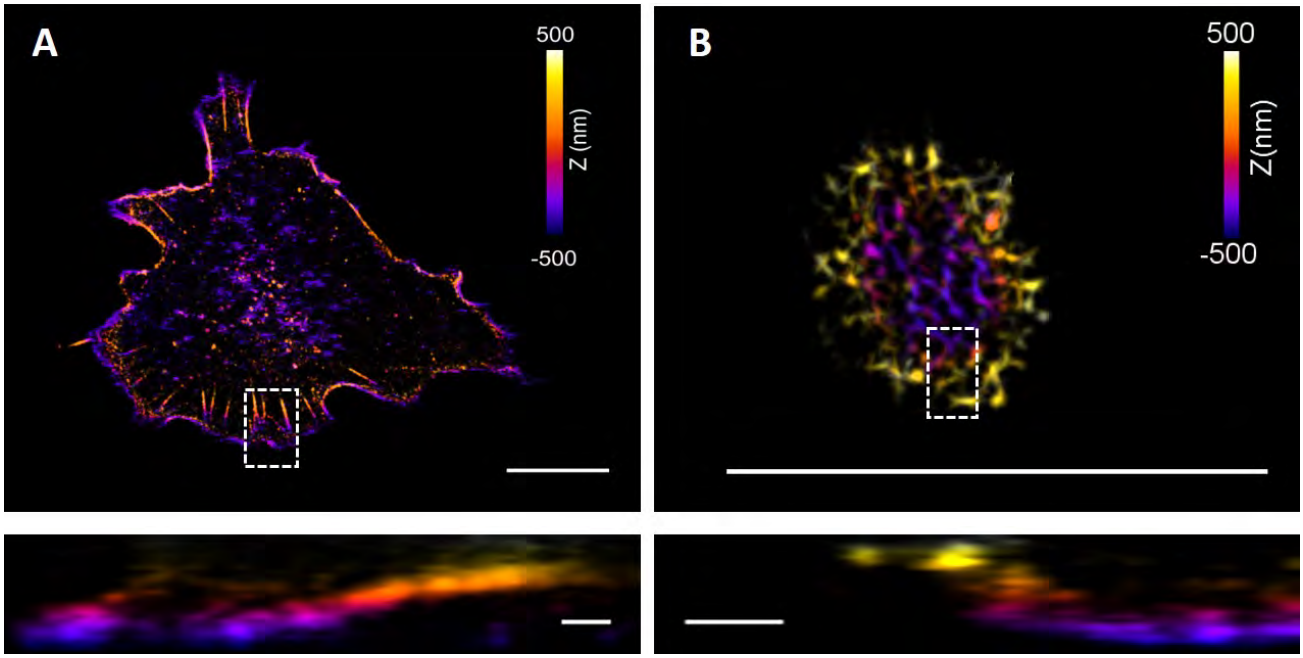

Figure 4.4: 3D STORM imaging for resolving actin and stress fiber organization in ESCs grown in either Serum-LIF or 2i-LIF media. 3D STORM images of actin filaments in stem cells grown in SL **(A)** or in 2i media **(B)**. The z-positions are color-coded (violet indicating the substrate). Lower panels show the side views of the box regions in the images above, indicating dorsal type of stress fibers in cells in Serum-LIF. Scale bars in upper panels 10 μm, in lower panels 500 nm.

mesh of actin without any clear directions or bundles (figure 4.4B).

Surface actin nanoscale organization is more disordered in 2i-LIF than in Serum-LIF

In order to quantify the seeming disorder in the orientations of filaments in 2i, we went back to 2D STORM, and used the Gabor filter in MATLAB to emphasize and connect the STORM point detections into filaments based on their common orientation with the Gabor filter kernels. The filaments were then color coded based on their orientation respective to the major axis of the cells (figure 4.5). While neither condition showed a clear overall alignment of actin filaments, we did however quantify the orientation order parameter $S = \langle \cos(2\theta) \rangle$, where θ is the angle between local fiber orientation and the overall cell orientation (figure 4.5E). Actin filaments in stem cells in 2i were found to be significantly more disordered than filaments in SL cells ($p < 0.01$).

Viscoelastic properties of stem cells in 2i compared to SL

Motivated by the significant differences in the actin structure we next set out to probe the viscoelastic properties of the two cell conditions. Specifically we measured the scaling exponent of diffusive movement for individual lipid granules near the cell periphery, where we see high concentrations of actin in both SL and 2i (figures 4.2 and 4.4). For anomalous diffusion, such as the thermal movements of a particle within a crowded environment of actin polymers, the mean squared displacement (MSD) follows a non-linear dependency on time (τ), given as $MSD(\tau) \sim \tau^{\alpha}$, where α is the scaling exponent. For 0<α< 1 we term the movement as sub-diffusion, meaning that the free diffusion of the particle is constricted by confinement or obstacles in the medium [97]. The lower the value of α, the more constricted the movement, and in the limit of α = 0, the particle would be completely confined in space. The value of α can be assessed from the power

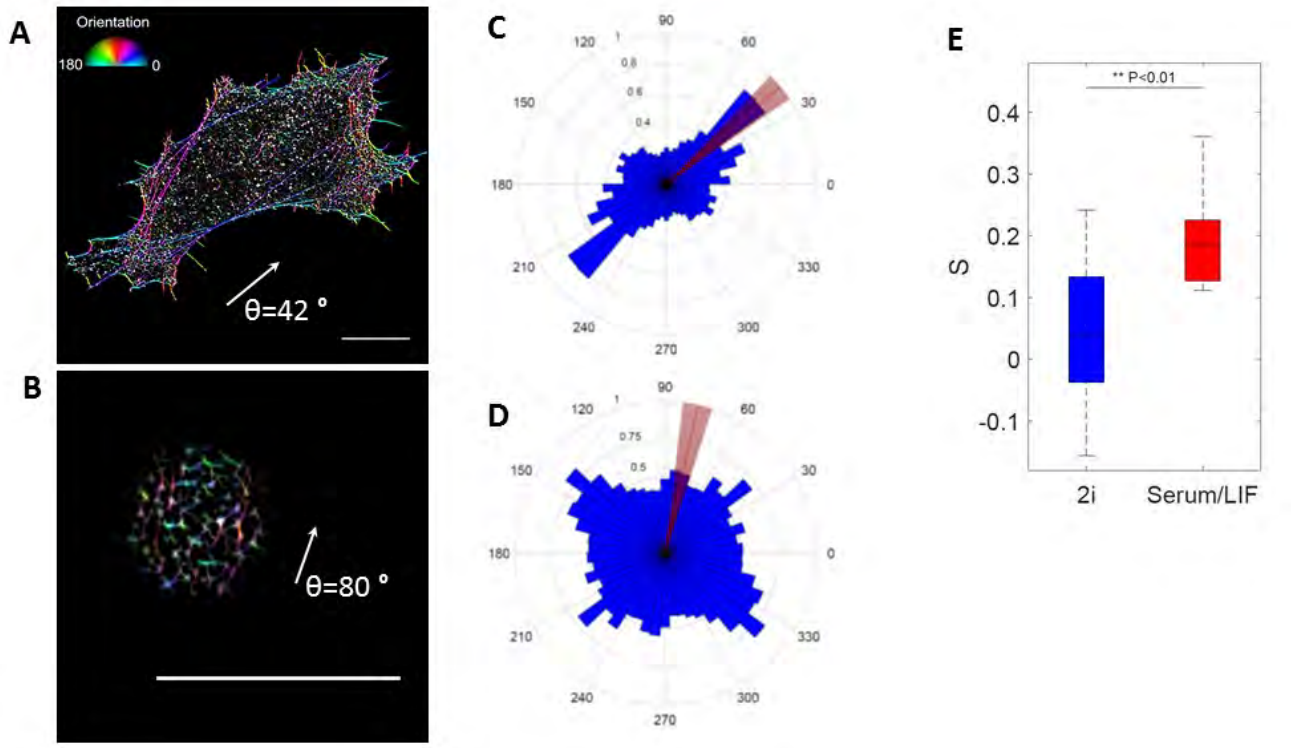

Figure 4.5: Characterization of the actin filament orientation in ESC near the surface. **(A)** Pixel orientation map of actin filaments in cells in SL, and **(B)** in cells in 2i media. Scale bars, 10 microns. **(C)** and **(D)** the angular plots corresponding to **(A)** and **(B)** respectively. Red line in the plots indicates the orientation of the cells. **E** Boxplot of the order parameter $S = \langle cos(2\theta) \rangle$ of the actin filament orientation, where θ is the difference between local fiber orientation and the overall cell orientation. Each boxplot represents 16 cells.

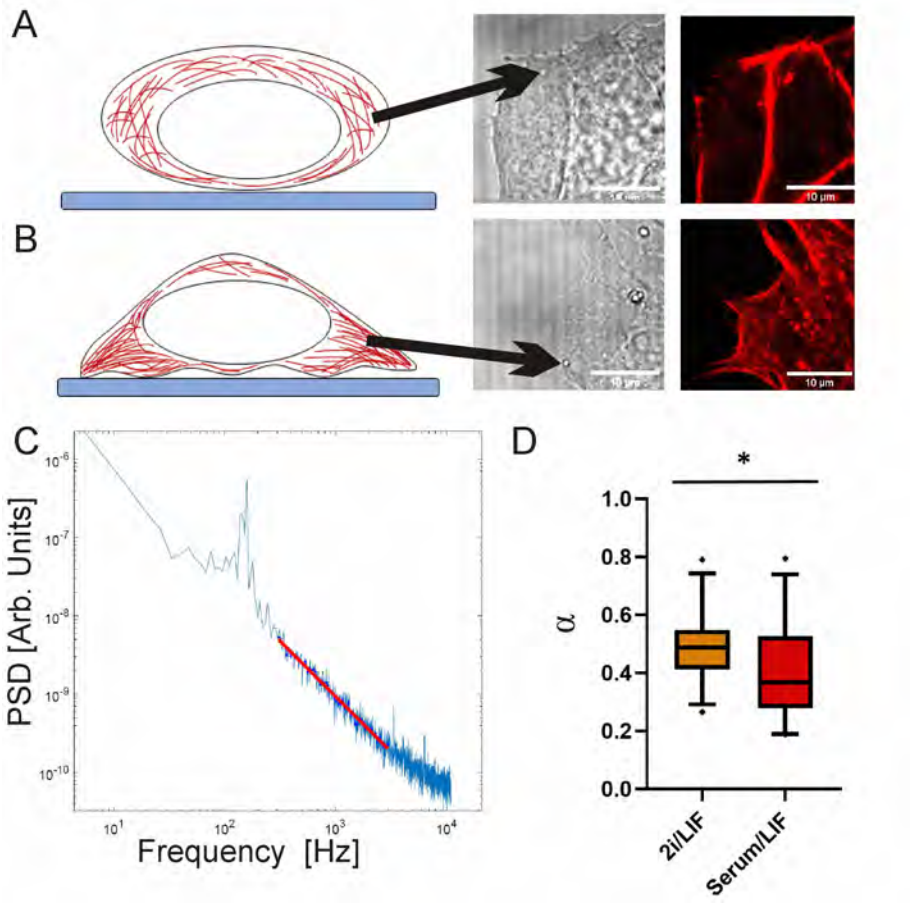

Figure 4.6: Measurement of viscoelastic properties of cells in 2i and SL. **(A)** From left to right, schematics, bright field and confocal images of trapping position in the actin cortex for cells in 2i, and **(B)** for cells in SL. F-actin is labeled by the SIR-actin probe. **(C)** Representative power spectrum of the movement of a lipid granule inside the optical trapping volume. Red line shows the linear range fitted for determination of the scaling exponent, $\alpha = 0.41$. **(D)** Box-plot of α-values from lipid granule trapping measurements in 2i and SL. * indicates p<0.05.

spectrum, $P_x(f)$. In the high frequency limit the power spectrum scales with frequency, f as:

$$P_x(f) \equiv \langle |x(f)|^2 \rangle \propto f^{-1+\alpha} \tag{4.1}$$

By focusing the optical trap on naturally occurring lipid granules, we acquired the lateral position of the granule at high spatio-temporal resolution. The position data were Fourier transformed in order to obtain the power spectrum of the movement. To determine the scaling exponent α, we fitted the range of 300-3000 Hz with a linear function in the double-logarithmic scale, as indicated in figure 4.6C. For each condition, 26 individual granules were trapped and their scaling exponent fitted and plotted in the boxplot of figure 4.6D. For cells in 2i, we found a scaling exponent of $\alpha = 0.48 \pm 0.11$, while in SL we found $\alpha = 0.40 \pm 0.15$. While the scaling exponent is similar in both conditions, we did however find a small but significant difference (p=0.03) between cells in SL and 2i. The movement of granules in the cell cortex of cells in 2i is thus less constrained than in cells grown in SL media, indicating a higher degree of actin compaction and crowding in the periphery of SL cells.

4.5 Discussion

Our confocal imaging of the two early stem cell stages revealed clear differences in the morphology of both single cells and colonies of cells. Cells grown in 2i-LIF had a more spherical shape and formed domed cell clusters whereas cells grown in Serum-LIF showed efficient spreading and formed monolayers when cultured on a surface. Closer inspection of the actin structured by STORM imaging showed characteristic differences in the nanoscale structure of the actin cytoskeleton between 2i-LIF and Serum-LIF cells. Generally, all our results point towards a tendency of cells in Serum-LIF to spread along the surface by flattening and extending actin protrusions out with dorsal stress fibers connecting to the substrate. Stem cells in 2i-LIF on the other hand, mainly organize their actin in a cortical mesh with very few radial actin structures protruding from the cell. Their low affinity for substrate attachment results in rounded and relatively symmetrical shape, unlike cells in Serum-LIF.

Conventionally, viscoelasticity measurements are performed on granules at random places in the cell, but here we attempted to manually select granules which were lying peripherally in the cell by using the actin imaging probe SIR-actin. The manual selection of granules should be emphasized, but these measurements anyhow give an indication that the viscoelastic environment at the cell periphery may be different under these two growth conditions and these preliminary results do merit further investigation of this phenomena.

The tension in membranes was recently found to regulate the fate of stem cells through tension regulation of endocytosis [24, 26]. Our results show that stem cells already in early stages of differentiation begin to show distinct viscoelastic properties in the cell cortex, although these differences are still much smaller than at later stages in development [79]. Stem cells grown in SL media show a higher degree of sub-diffusion in our experiments, indicating a more crowded and confining cytoskeleton at the periphery. This would intuitively correspond to a stiffer cell cortex at this later developmental stage. However, with inhibitor withdrawal from 2i media, cells were shown to decrease their membrane tension drastically as they transition towards differentiation [24, 26]. While at first glance this seems contrasting to our results, it is important to note two crucial differences between the experiments. Firstly, in this study we probe two separate stages of stem cell development; the slightly earlier and more naive early blastocyst stage, recapitulated by 2i-LIF media, and the later primed or differentiated epiblast stage, recapitulated by cells in Serum-LIF media [7, 22, 86]. These are both defined and stable stages in our media conditions, whereas the mentioned articles study the transition from the naive stage. Secondly, a higher degree of cytoskeletal confinement at the cell cortex does not necessarily exclude lower membrane tension. As described in Bergert and others [26], during the transition the membrane tension decreases from a reduction in membrane-cytoskeleton linkage, and the transition could indeed be inhibited by artificially linking the membrane to the cytoskeleton.

Here, we have shown that subtle differences in the developmental stages of ESCs are associated with dramatic differences in the actin nano-architecture of the cell cortex and also with changes in the physical properties. Future studies should focus on implications of these differences on cell fate and on the mechanisms behind such as how the dynamic actin structure just could regulate the fate of cells possibly by affecting processes in the much softer plasma cell membrane.

Author Contributions

K.G.H. wrote the first draft. Y.F.B and K.G.H. built the super-resolution setup and implemented STORM imaging. K.G.H. cultured and adjusted stem cells to the two conditions and prepared samples for imaging. Y.F.B performed super resolution and confocal experiments of the actin structure, I.I.P. and Y.F.B. performed the viscoelastic measurements. Y.F.B. performed the data analysis. L.B.O. and P.M.B. supervised the project. All authors commented and edited the manuscript

Acknowledgments

This work is financially supported by Danish National Research Foundation, grant DNRF116 (StemPhys), and the Danish Council for Independent Research, Natural Sciences, grant DFF-4181-00196.

4.6 Supplementary figures

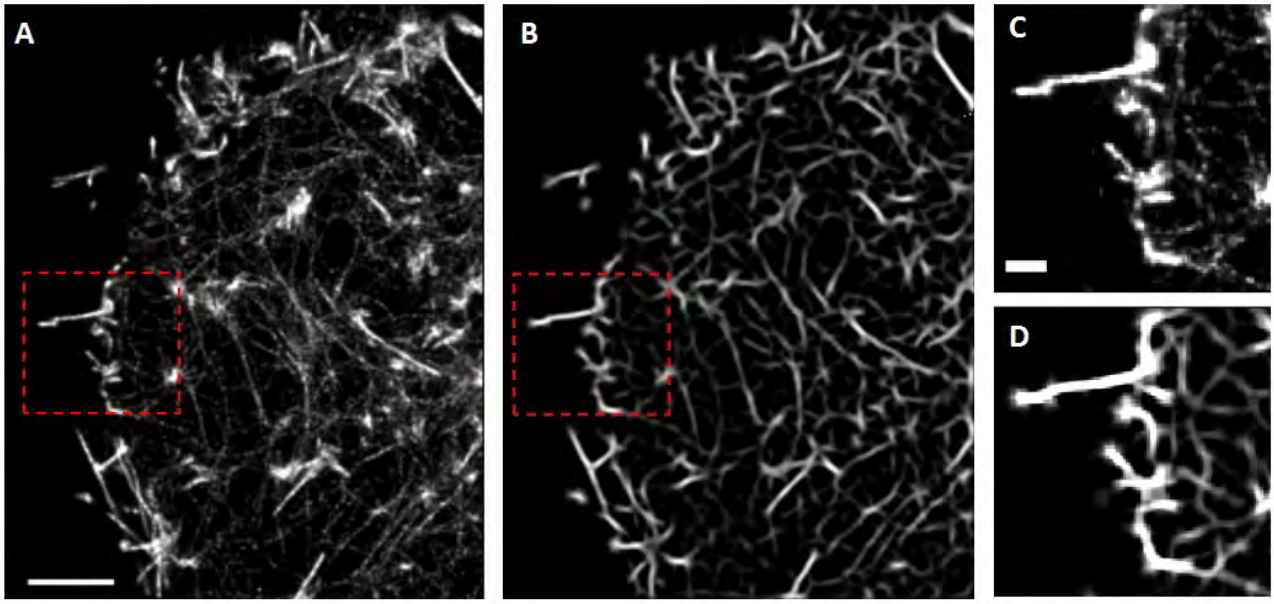

Figure 4.7: Spatial resolution enhancement. A) reconstructed STORM image of actin cytoskeleton in a stem cell. B) Corresponding image to A after a 2D Gobar filter. C and D) show magnified views of the red region in A and B, respectively. Scale bar in A and B is 2 um and 0.5 um in C and D.

Chapter 5

Imaging active transcription sites in fixed cells with mRNA FISH

5.1 Introduction

Single molecule tracking of Sox2 revealed a broad distribution of residence times, some lasting just half a second with a few events lasting hundreds of seconds. We furthermore saw that this distribution was relatively stable across transcription states, but notably, this distribution was measured by sampling a random subset of Sox2 molecules in the whole nucleus, and thus doesn't say much about how Sox2 binding is affected relative to specific genes. It would therefore be interesting to relate transcription factor binding with the transcription of a specific gene known to be regulated by the factor. Perhaps high levels of transcription are related to the highly stable binding events. A variety of methods exist with the aim of visualizing a specific gene or transcription site in cells. Live cell methods include labeling DNA sequences with a number of fluorescently labeled dCas9 [98], or by introducing a series of mRNA hairpin loop forming sequences from the bacteriophage MS2 to a gene of interest, which is then targeted by a fluorescently labeled protein [99]. However, both these methods require multiple rounds of cloning and transfection, and the generation of a stable reporter line is thus a lengthy process.

Instead, we aim to implement mRNA fluorescence in situ hybridization (mRNA FISH), which involves hybridizing sequences of mRNA transcript with complementary fluorescently labeled DNA oligos [99–101]. As a first step to test whether this method is feasible to visualize active transcription sites, we tested mRNA FISH on an ESC line containing an ERK activation construct [17], and targeting mRNA sequences we know are regulated by ERK activation.

5.2 Methods

Cell culture

E14Ju murine stem cells (ESC) containing an ERK-activation construct (ERBA, from W. Hamilton, DanStem) as described in Hamilton and Brickman (2014)[14], were cultured on gelatin covered flasks (Corning) at 37 degrees and 5% CO2. Cells were cultured in media consisting of Glasgow's Minimum Essential Media (GMEM, Sigma-Aldrich) supplemented with 10% Fetal Bovine Serum (FBS)(Gibco), 2 mM L-Glutamine (Gibco), 1 mM Sodium Pyruvate (Gibco), 0.1 mM Non-

Essential Amino Acids (Gibco), 1000 U/mL Leukemia Inhibitory Factor (LIF)(DanStem, University of Copenhagen), and 0.1 mM 2-mercaptoethanol (Sigma). For normal culturing, the cells were grown in the presence of the GSK3 inhibitor CHIR99021 (CHIR, Sigma), which was exchanged for the FGFR-1 and FGFR-3 inhibitor PD17 (Sigma) 24 hours prior to experiments with (and without) ERK activation. CHIR99021 was used at a concentration of 3μM and PD17 at a concentration of 250 nM. Since the construct contained puromycin resistance, the media was further supplemented with 1 $\mu g/mL$ puromycin.

Sample preparation and ERK activation

Cells were plated on round glass #1.5 coverslips coated with 10 µg/mL fibronectin (EMD Millipore) in DPBS, each placed at the bottom of a well in a 12 well plate at a density of 6000 cells/cm^2 in media supplemented with PD17 (PD173074, Sigma), and left to grow for 24 hours. After 24 hours (i.e. time 0 on figure 5.1), cells were either ERK-activated by addition of 250 nM tamoxifen (4-OHT) ((Z)-4-Hydroxytamoxifen, Sigma), or kept as a control by addition of an equal volume of ethanol (ETOH). Cells were then either fixed after 2 or 8 hours, or after 10 hours with addition of 1 μM MEK inhibitor PD03 (PD0325901, Sigma) for the final 2 hours of the experiment. Cells were fixed in freshly prepared 4% paraformaldehyde in PBS for 10 minutes, and stored at 5 $^\circ$C after washing with PBS.

Fluorescence in situ hybridization (FISH)

Samples were labeled with mRNA FISH probes targeting Nanog exons (Nanog-CAL Fluor Red 590, provided by S. Pozzi, Danstem) and Arc exons (Arc-Cy5, provided by T. Lionnet, New York University), according to a protocol developed and provided by W. Shuoshuo (New York University, currently Harvard University). Cells were first permeabilized with PBS containing 0.5% Triton X100 (Sigma), then prepared for hybridization by incubating with the pre-hybridization buffer consisting of 10% deionized formamide (invitrogen) in 2X saline-sodium citrate (SSC) buffer.

A hybridization solution was prepared based on the pre-hybridization buffer, with 10 ng of each probe per sample, 10% dextran sulfate, 2 mg/mL bovine serum albumin (BSA), in addition to 5 µg E. Coli tRNA and 5 µg salmon sperm ssDNA per sample. Each coverslip was covered with 45 µL of the hybridization solution and incubated in a humidified chamber at 37 $^\circ$ overnight.

Following hybridization, samples went through two rounds of 20 minute washes in the pre-hybridization buffer, and once in PBS. Finally samples were briefly stained with 500 ng/mL DAPI and resuspended in PBS.

Coverslips were mounted onto glass slides (ThermoFisher), using ProLong Gold antifade reagent (ThermoFisher) as a mounting medium. After curing overnight, the slides were sealed with nail polish.

FISH probe design

The Klf4 intron probe was designed using the Stellaris RNA FISH probe designer (Stellaris, [102]). To avoid unspecific binding outside the Klf4 intron sequences, masking level 3 was used based on mouse genome information. A probe-set of 20-nucleotide long oligos was generated based on the Klf4-201 splicing variant intron sequences from Ensembl (Mus Musculus (GRCm38.p6), [103]) allowing a 10 basepair (bp) overlap with exon sequences, and 48 oligos with GC content of 40-55% and >3 bp spacing were selected.

Fluorescence imaging

Imaging was performed on a Nikon Eclipse Ti inverted microscope using a 100X oil-immersion objective (Nikon, NA = 1.4) equipped with an iXon Ultra 897 EMCCD camera (Andor) resulting in a pixel width of 160 nm. The sample was excited by epifluorescence illumination from lasers of 405 nm (Vortran), 488 nm (Vortran), 561 nm (Cobolt), and 637 nm (Vortran), and emission light was filtered using bandpass filters of 470/24 (Chroma) for DAPI signal, 525/50 (Semrock) for Alexa Fluor 488, 593/40 (Semrock) for Cy3, and 679/41 (Semrock) for Cy5 signal. Images were acquired as z-stacks with a stepsize of 200 nm, encompassing the full cell colony volume. Each z-position was imaged using an exposure time of 50 ms for DAPI and 300 ms for the remaining channels. Nanog-CAL Fluor Red 590 was imaged using Cy3 settings.

Image analysis

Cell nuclei were manually identified and designated using the DAPI signal, and individual cell nuclei were analyzed using the AirLocalize software for MATLAB by Lionnet *et al.* [99], which uses a Gaussian kernel for single molecule detection and localization in 3D. Since detection parameters depend on the probe and wavelength, detection parameters were determined and selected individually for each channel. A fixed, high threshold for detection was set, in order to minimize detections of single mRNA molecules outside of the transcription sites, yet low enough to detect the high intensity clustering of mRNA FISH probes that would be observed at an active transcription site. The fixed detection parameters were then applied to images of all experimental conditions for each channel, and the total number of detections per cell nucleus was calculated.

5.3 Preliminary results

mRNA FISH probes are typically designed to target either exon or intron sequences of RNA (example figure 5.3). Introns are non-coding RNA, and upon splicing intron sequences are cleaved from the transcript and degraded, leaving behind the exon sequences which are translated into protein. It is therefore only up until splicing that introns are part of the transcript, whereas exons stay for the entire lifetime of the transcript.

By targeting FISH probes to the intron sequence of a transcript of interest, the fluorescent signal will primarily be located at the active transcription site. Since exons are also present at active transcription sites, however, they are less specific to this site, and are therefore less ideal for imaging active transcription sites. For the preliminary test however, we wished to visualize loci which we know are regulated by ERK signaling, for which we only had exon probes available. We chose two exon probes of Nanog and Arc, which after ERK activation are quickly down- and up-regulated respectively (figure 5.1).

Firstly observed that many cells revealed either one or two bright spots of clustered fluorescence in the nucleus (5.2) consistent with the presence of an active transcription site. We do also see weaker individual points of fluorescence both inside and outside the nucleus, as expected from exon probes. To assess whether the bright spots we detect are indeed active transcription sites, we activated ERK in cells and quantified the number of bright spots per nucleus for each probe (figure 5.2). In our ethanol control groups, we initially detect a very low frequency of Arc sites. Upon ERK activation we quickly see a rapid increase in number of detections per cell, going down after 8 hours of treatment, before reaching initial conditions after MEK inhibition with PD03. This is well correlated with the expected behavior of Arc expression as an immediate early response gene as identified by Hamilton and Brickman [14]. On the other hand, activation of ERK is associated with a drastic decrease in number of Nanog active transcription sites after both 2 and 8 hours, consistent with

previous studies. The expression of Nanog is furthermore rescued by MEK inhibition, which recapitulates the findings by Hamilton and Brickman [14].

We do detect some cells with more than 2 active transcription site, which could be explained by defects in the chromosomes of individual cells, or by false-positive detections arising from using an exon probe to identify transcription sites.

In conclusion, we have observed that detection of individual transcription sites is possible in our cell line with mRNA FISH, however future work must now be focused on both implementing intron probes (figure 5.3) and finally correlating it with single molecule tracking after fixation.

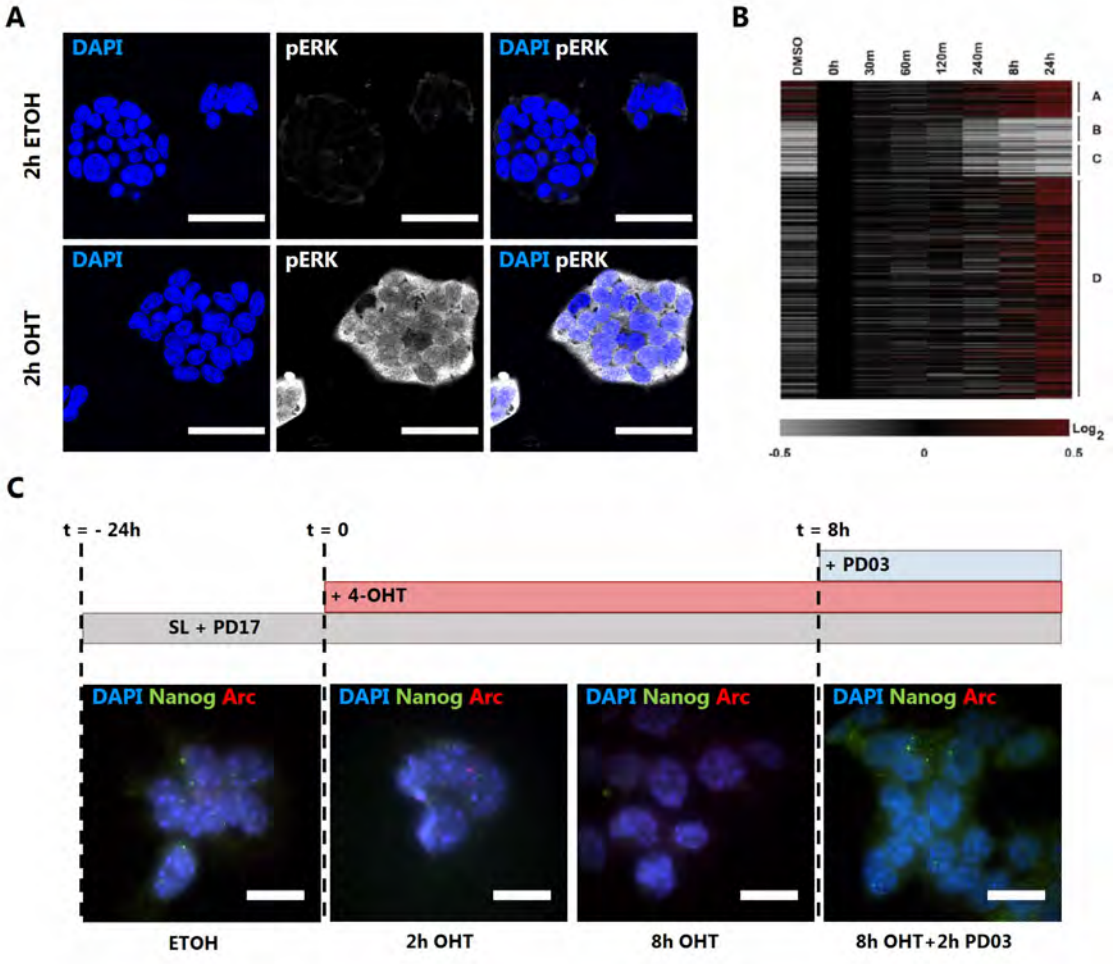

Figure 5.1: Overview of the ERK activation experiment and mRNA FISH imaging. A) Addition of 4-OHT to stem cells with the ERK activation construct results in a drastic and homogeneous increase in phosphorylated ERK already after 2 hours. B) Microarray data from Hamilton and Brickman (2014)[14] shown here, reveals up-regulation of primitive endoderm markers upon ERK activation, and down-regulation of blastocyst and pluripotency markers. Among up-regulated genes was Arc, while Nanog is down-regulated by ERK activation. C) ERK activation experiment timeline; 24 hours of FGFR inhibition by PD17 is followed by 4-OHT activation of constitutively active cRAF, and imaging after 2 and 8 hours. After 8 hours, MEK inhibitor PD03 is added, quickly blocking the ERK signalling pathway and reversing priming [14], and a final timepoint is taken after 8+2 hours. Ethanol was added to cells on the side as a control for each timepoint. Scalebars are 20 µm.

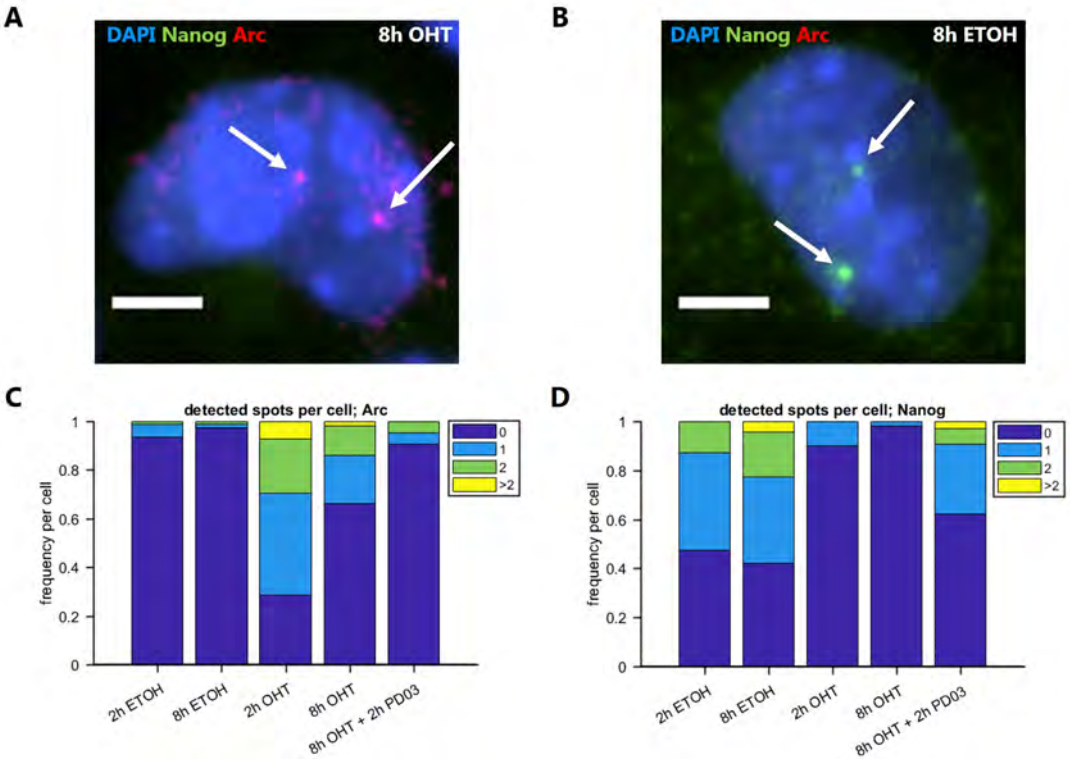

Figure 5.2: Results of mRNA FISH imaging quantification after ERK activation. A) Example of a stem cell nucleus (blue) after 8 hours of ERK activation. The image shown is the summed projection of 16 z-stack slices encompassing the entire cell nucleus. White arrows indicate bright spots assumed to be active transcription sites for preliminary analysis. B) Example of a stem cell nucleus from the ethanol control group after 8 hours. As in A), the image shown is the summed projection of 16 z-stack slices. Scale bars are 5 μm. C-D) Quantification of the number of detected spots per cell nucleus using a fixed, high intensity threshold that was evaluated

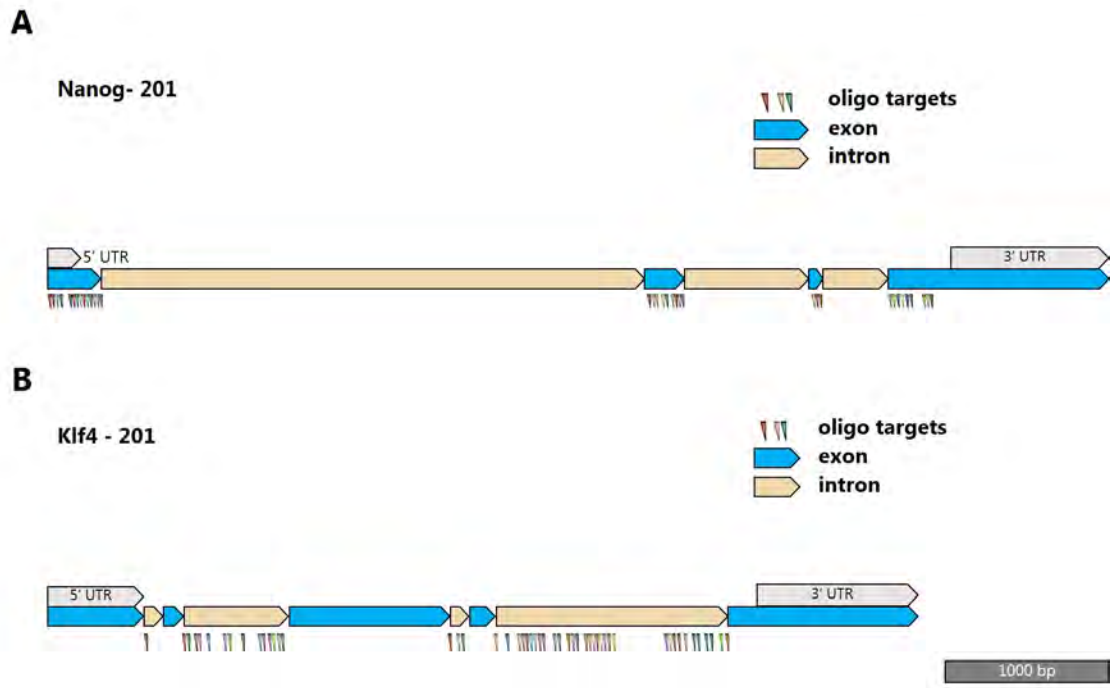

Figure 5.3: Schematic of exonic and intronic mRNA FISH probes. A) Nanog transcript divided into exons and introns, and the oligo targets of the exon probe set designed by S. Pozzi, which was used in this chapter. B) Klf4 transcript divided into exons and introns, and the oligo targets of the intron probe set that was designed for fixed cell imaging of active transcription sites in future and continued studies. The probe consists of 48 fluorescently labeled DNA oligos of 20 nucleotides length, complementary to intronic mRNA sequences separated by at least 3 nucleotides. Both Nanog and Klf4 are drawn to scale, and the scale bar of 1000 bp is shown in the bottom right corner.

Chapter 6

Conclusion

Our work on Sox2 single molecule tracking allowed quantification of the DNA residence time distribution, and revealed a small but consistent sub-population of highly stable binding events, which is not accounted for in previously published models. Interestingly we found this sub-population to be unaffected by primitive endoderm priming or p300 acetylation, but observed that it was reduced by the inhibition of the co-activator BRD4, implicated in highly stable liquid-phase condensates around super-enhancers [37].

Future work on Sox2 single-molecule tracking should aim to correlate binding with transcription co-factors and sites of active transcription. Co-localization with sites of transcription is already possible by fixation and mRNA FISH labeling with preferably intron probes, however more intricate methods such as MS2 or dCas9 labeling could also be developed for co-localization in live cells without fixation. Building on the model of highly stable liquid-phase condensates or aggregates of transcription factors and co-factors, it will be interesting to measure if stable Sox2 binding shows any correlation with these aggregates. This could be investigated by introducing a fluorescent label to BRD4, Mediator, or even RNAP, which we know aggregate in these clusters [35–37].

The nano-scale actin architecture in ESC revealed drastic differences between cells grown in the naive pluripotency promoting culturing media 2i-LIF compared to the more primed state of cells in serum-LIF media. Actin was shown to be organized in a disordered mesh running parallel to the cell cortex in naive 2i-LIF cells, with cells seeming to minimize their contact with the substrate. On the other hand, we showed distinct actin stress fibers attaching to the substrate at the periphery of cells, indicating a tendency to spread along the surface.

The two conditions investigated showed drastic differences between two stable states of cultured ESC, however it could be interesting in future work to focus on how the actin cytoskeleton is affected during the transition from a state of pluripotency to priming. The ERK activation construct used in the study of Sox2 residence time would provide a convenient system to quantitatively describe differences in actin nanoscale structure, as cells are primed and pushed towards differentiation.

Finally, this work demonstrates how single-molecule fluorescence microscopy applied to the study of ESC can yield new insights into the underlying mechanisms of both transcription and cytoskeleton architecture on the sub cellular level. Furthermore, the methods implemented and described in this study have broad applications, and this thesis lays a groundwork for future and continued studies of single molecules and nanoscale structures in cellular contexts.

Chapter 7

References

[1] Hvid, K. G. *Construction of a single-molecule localization microscope and Airy beam light-sheet microscopy for investigating stem cell development*. Master's thesis, University of Copenhagen (2017).

[2] Saiz, N. GATA6 and the power of single cells (2014). URL `https://thenode.biologists.com/gata6-and-the-power-of-single-cells/research/`, Accessed: 2019-12-22.

[3] Kang, M., Piliszek, A., Artus, J. & Hadjantonakis, A.-K. Fgf4 is required for lineage restriction and salt-and-pepper distribution of primitive endoderm factors but not their initial expression in the mouse. *Development* **140**, 267 (2013).

[4] Nissen, S. B. *et al.* Four simple rules that are sufficient to generate the mammalian blastocyst. *PLOS Biology* **15**, e2000737 (2017).

[5] Morgani, S. M. & Brickman, J. M. LIF supports primitive endoderm expansion during pre-implantation development. *Development* **142**, 3488 (2015).

[6] Nichols, J., Silva, J., Roode, M. & Smith, A. Suppression of Erk signalling promotes ground state pluripotency in the mouse embryo. *Development* **136**, 3215 (2009).

[7] Nichols, J. & Smith, A. Naive and primed pluripotent states. *Cell Stem Cell* **4**, 487–492 (2009).

[8] Rossant, J. Stem cells and early lineage development. *Cell* **132**, 527–531 (2008).

[9] Lanner, F. & Rossant, J. The role of FGF/Erk signaling in pluripotent cells. *Development* **137**, 3351 (2010).

[10] Dorey, K. & Amaya, E. FGF signalling: diverse roles during early vertebrate embryogenesis. *Development* **137**, 3731 (2010).

[11] Schrode, N., Saiz, N., Di Talia, S. & Hadjantonakis, A.-K. GATA6 levels modulate primitive endoderm cell fate choice and timing in the mouse blastocyst. *Developmental Cell* **29**, 454–467 (2014).

[12] Yamanaka, Y., Lanner, F. & Rossant, J. FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. *Development* **137**, 715 (2010).

[13] Zhang, W. & Liu, H. T. MAPK signal pathways in the regulation of cell proliferation in mammalian cells. *Cell Research* **12**, 9–18 (2002).

[14] Hamilton, W. B. & Brickman, J. M. Erk signaling suppresses embryonic stem cell self-renewal to specify endoderm. *Cell Rep* **9**, 2056–70 (2014).

[15] Grabarek, J. B. *et al.* Differential plasticity of epiblast and primitive endoderm precursors within the ICM of the early mouse embryo. *Development* **139**, 129 (2012).

[16] Ying, Q.-L. *et al.* The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519–523 (2008).

[17] Hamilton, W. B. *et al.* Dynamic lineage priming is driven via direct enhancer regulation by ERK. *Nature* **575**,

355–360 (2019).

[18] Czechanski, A. *et al.* Derivation and characterization of mouse embryonic stem cells from permissive and nonpermissive strains. *Nature Protocols* **9**, 559–574 (2014).

[19] Evans, M. J. & Kaufman, M. H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154–156 (1981).

[20] Martin, G. R. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences* **78**, 7634–7638 (1981).

[21] Canham, M. A., Sharov, A. A., Ko, M. S. H. & Brickman, J. M. Functional heterogeneity of embryonic stem cells revealed through translational amplification of an early endodermal transcript. *PLoS biology* **8**, e1000379–e1000379 (2010).

[22] Martin Gonzalez, J. *et al.* Embryonic stem cell culture conditions support distinct states associated with different developmental stages and potency. *Stem Cell Reports* **7**, 177–191 (2016).

[23] Mulas, C. *et al.* Defined conditions for propagation and manipulation of mouse embryonic stem cells. *Development* **146**, dev173146 (2019).

[24] De Belly, H., Jones, P. H., Paluch, E. K. & Chalut, K. J. Membrane tension mediated mechanotransduction drives fate choice in embryonic stem cells. *bioRxiv* 798959 (2019).

[25] Xia, S. *et al.* Nanoscale architecture of the cortical actin cytoskeleton in embryonic stem cells. *Cell Reports* **28**, 1251–1267.e7 (2019).

[26] Bergert, M. *et al.* Cell surface mechanics gate stem cell differentiation. *bioRxiv* 798918 (2019).

[27] Vernimmen, D. & Bickmore, W. A. The hierarchy of transcriptional activation: From enhancer to promoter. *Trends in Genetics* **31**, 696–708 (2015).

[28] Halfon, M. S. Studying transcriptional enhancers: The founder fallacy, validation creep, and other biases. *Trends in Genetics* **35**, 93–103 (2019).

[29] Andersson, R. & Sandelin, A. Determinants of enhancer and promoter activities of regulatory elements. *Nature Reviews Genetics* (2019).

[30] Fukaya, T., Lim, B. & Levine, M. Enhancer control of transcriptional bursting. *Cell* **166**, 358–368 (2016).

[31] Heist, T., Fukaya, T. & Levine, M. Large distances separate coregulated genes in living Drosophila embryos. *Proceedings of the National Academy of Sciences* **116**, 15062 (2019).

[32] Cho, W. K. *et al.* RNA Polymerase II cluster dynamics predict mrna output in living cells. *Elife* **5** (2016).

[33] Stadhouders, R., Filion, G. J. & Graf, T. Transcription factors and 3d genome conformation in cell-fate decisions. *Nature* **569**, 345–354 (2019).

[34] Hansen, A. S., Pustova, I., Cattoglio, C., Tjian, R. & Darzacq, X. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *eLife* **6**, e25776 (2017).

[35] Cho, W.-K. *et al.* Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* **361**, 412 (2018).

[36] Chong, S. *et al.* Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* **361**, eaar2555 (2018).

[37] Sabari, B. R. *et al.* Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 (2018).

[38] The Nobel Prize in Chemistry 2014 - Press Release. URL `nobelprize.org/nobel_prizes/chemistry/laureates/2014/press.html`, Accessed: 2017-06-23.

[39] Klar, T. A., Jakobs, S., Dyba, M., Egner, A. & Hell, S. W. Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission. *Proceedings of the National Academy of Sciences* **97**, 8206 (2000).

[40] Betzig, E. *et al.* Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**, 1642–5 (2006).

[41] Rust, M. J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat Methods* **3**, 793–5 (2006).

[42] van de Linde, S. *et al.* Direct stochastic optical reconstruction microscopy with standard fluorescent probes. *Nat Protoc* **6**, 991–1009 (2011).

[43] Gustafsson, M. G. L. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *Journal of Microscopy* **198**, 82–87 (2000).

[44] Vicidomini, G., Bianchini, P. & Diaspro, A. STED super-resolved microscopy. *Nature Methods* **15**, 173–182 (2018).

[45] Wäldchen, S., Lehmann, J., Klein, T., van de Linde, S. & Sauer, M. Light-induced cell damage in live-cell super-resolution microscopy. *Scientific Reports* **5**, 15348 (2015).

[46] Li, J. *et al.* Single-molecule nanoscopy elucidates RNA Polymerase II transcription at single genes in live cells. *Cell* **178**, 491–506.e28 (2019).

[47] Heilemann, M. *et al.* Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. *Angew Chem Int Ed Engl* **47**, 6172–6 (2008).

[48] Tokunaga, M., Imamoto, N. & Sakata-Sogawa, K. Highly inclined thin illumination enables clear single-molecule imaging in cells. *Nat Methods* **5**, 159–61 (2008).

[49] Axelrod, D. Total internal reflection fluorescence microscopy in cell biology. *Traffic* **2**, 764–774 (2001).

[50] Manley, S. *et al.* High-density mapping of single-molecule trajectories with photoactivated localization microscopy. *Nat. Methods* **5**, 155–157 (2008).

[51] Abràmoff, M. D., Magalhães, P. J. & Ram, S. J. Image processing with imageJ. *Biophotonics Int.* **11**, 36–41 (2004).

[52] Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676–82 (2012).

[53] Ovesny, M., Krizek, P., Borkovec, J., Svindrych, Z. & Hagen, G. M. ThunderSTORM: a comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging. *Bioinformatics* **30**, 2389–90 (2014).

[54] Jin, J. *et al.* Synergistic action of RNA polymerases in overcoming the nucleosomal barrier. *Nat. Struct. Mol. Biol* **17**, 745–752 (2010).

[55] Cisse, I. I. *et al.* Real-time dynamics of RNA Polymerase II clustering in live human cells. *Science* **341**, 664 (2013).

[56] Izeddin, I. *et al.* Single-molecule tracking in live cells reveals distinct target-search strategies of transcription factors in the nucleus. *eLife* **3**, e02230 (2014).

[57] Levine, M. & Tjian, R. Transcription regulation and animal diversity. *Nature* **424**, 147–151 (2003).

[58] Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).

[59] Huisken, J. & Stainier, D. Y. R. Selective plane illumination microscopy techniques in developmental biology. *Development* **136**, 1963 (2009).

[60] Chen, B.-C. *et al.* Lattice light-sheet microscopy: Imaging molecules to embryos at high spatiotemporal resolution. *Science* **346**, 1257998 (2014).

[61] Wang, S., Moffitt, J. R., Dempsey, G. T., Xie, X. S. & Zhuang, X. Characterization and development of photoactivatable fluorescent proteins for single-molecule–based superresolution imaging. *Proceedings of the National Academy of Sciences* **111**, 8452 (2014).

[62] Grimm, J. B. *et al.* A general method to improve fluorophores for live-cell and single-molecule microscopy. *Nat*

Methods **12**, 244–50, 3 p following 250 (2015).

[63] Grimm, J. B. *et al.* Bright photoactivatable fluorophores for single-molecule imaging. *Nat Methods* **13**, 985–988 (2016).

[64] Darzacq, X. *et al.* Imaging transcription in living cells. *Annual review of biophysics* **38**, 173–196 (2009).

[65] Chen, J. *et al.* Single-molecule dynamics of enhanceosome assembly in embryonic stem cells. *Cell* **156**, 1274–1285 (2014).

[66] Liu, Z. *et al.* 3D imaging of Sox2 enhancer clusters in embryonic stem cells. *Elife* **3**, e04236 (2014).

[67] Morisaki, T., Müller, W. G., Golob, N., Mazza, D. & McNally, J. G. Single-molecule analysis of transcription factor binding at transcription sites in live cells. *Nature Communications* **5**, 4456 (2014).

[68] Li, L. *et al.* Real-time imaging of Huntingtin aggregates diverting target search and gene transcription. *eLife* **5**, e17056 (2016).

[69] Tinevez, J.-Y. simpletracker (2019). URL `https://www.github.com/tinevez/simpletracker`, Accessed: 2019-12-19.

[70] Tarantino, N. *et al.* TNF and IL-1 exhibit distinct ubiquitin requirements for inducing NEMO–IKK supramolecular structures. *The Journal of Cell Biology* **204**, 231–245 (2014).

[71] Garcia, D. A. *et al.* A new model for single-molecule tracking analysis of transcription factor dynamics. *bioRxiv* 637355 (2019).

[72] Mazza, D., Abernathy, A., Golob, N., Morisaki, T. & McNally, J. G. A benchmark for chromatin binding measurements in live cells. *Nucleic Acids Research* **40**, e119–e119 (2012).

[73] Akaike, H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19**, 716–723 (1974).

[74] Hurvich, C. M. & Tsai, C.-L. Regression and time series model selection in small samples. *Biometrika* **76**, 297–307 (1989).

[75] Swinstead, E. *et al.* Steroid receptors reprogram foxa1 occupancy through dynamic chromatin transitions. *Cell* **165**, 593–605 (2016).

[76] Teves, S. S. *et al.* A dynamic mode of mitotic bookmarking by transcription factors. *eLife* **5**, e22280 (2016).

[77] Loffreda, A. *et al.* Live-cell p53 single-molecule binding is modulated by C-terminal acetylation and correlates with transcriptional activity. *Nature Communications* **8**, 313 (2017).

[78] Weinert, B. T. *et al.* Time-resolved analysis reveals rapid dynamics and broad scope of the CBP/p300 acetylome. *Cell* **174**, 231–244.e12 (2018).

[79] Urbanska, M. *et al.* Single-cell mechanical phenotype is an intrinsic marker of reprogramming and differentiation along the mouse neural lineage. *Development* **144**, 4313–4321 (2017).

[80] Engler, A. J., Sen, S., Sweeney, H. L. & Discher, D. E. Matrix elasticity directs stem cell lineage specification. *Cell* **126**, 677–689 (2006).

[81] Wullkopf, L. *et al.* Cancer cells' ability to mechanically adjust to extracellular matrix stiffness correlates with their invasive potential. *Molecular Biology of the Cell* **29**, 2378–2385 (2018).

[82] Han, Y. L. *et al.* Cell swelling, softening and invasion in a three-dimensional breast cancer model. *Nature Physics* (2019).

[83] Xu, K., Babcock, H. P. & Zhuang, X. Dual-objective STORM reveals three-dimensional filament organization in the actin cytoskeleton. *Nature methods* **9**, 185–188 (2012).

[84] Han, B., Zhou, R., Xia, C. & Zhuang, X. Structural organization of the actin-spectrin–based membrane skeleton in

dendrites and soma of neurons. *Proceedings of the National Academy of Sciences* **114**, E6678 (2017).

[85] Sim, Y.-J. *et al.* 2i maintains a naive ground state in ESCs through two distinct epigenetic mechanisms. *Stem Cell Reports* **8**, 1312–1328 (2017).

[86] Morgani, S. M. *et al.* Totipotent embryonic stem cells arise in ground-state culture conditions. *Cell Rep* **3**, 1945–57 (2013).

[87] Andersen, T. *et al.* Nanoscale phase behavior on flat and curved membranes. *Nanotechnology* **25**, 505101 (2014).

[88] Huang, B., Wang, W., Bates, M. & Zhuang, X. Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science* **319**, 810–813 (2008).

[89] Hendargo, H. C. *et al.* Automated non-rigid registration and mosaicing for robust imaging of distinct retinal capillary beds using speckle variance optical coherence tomography. *Biomedical optics express* **4**, 803–821 (2013).

[90] Dai, P. *et al.* A new approach to segment both main and peripheral retinal vessels based on gray-voting and Gaussian mixture model. *PLOS ONE* **10**, e0127748 (2015).

[91] Inoue, S. *et al.* Live cell tracking of symmetry break in actin cytoskeleton triggered by abrupt changes in micromechanical environments. *Biomaterials Science* **3**, 1539–1544 (2015).

[92] Elosegui-Artola, A. *et al.* Image analysis for the quantitative comparison of stress fibers and focal adhesions. *PLOS ONE* **9**, e107393 (2014).

[93] Hansen, P. M., Tolic-Nørrelykke, I. M., Flyvbjerg, H. & Berg-Sørensen, K. tweezercalib 2.1: Faster version of MatLab package for precise calibration of optical tweezers. *Computer Physics Communications* **175**, 572–573 (2006).

[94] Chalut, K. & Paluch, E. The actin cortex: A bridge between cell shape and function. *Developmental Cell* **38**, 571–573 (2016).

[95] Tojkander, S., Gateva, G. & Lappalainen, P. Actin stress fibers – assembly, dynamics and biological roles. *Journal of Cell Science* **125**, 1855–1864 (2012).

[96] Naumanen, P., Lappalainen, P. & Hotulainen, P. Mechanisms of actin stress fibre assembly. *Journal of Microscopy* **231**, 446–454 (2008).

[97] Norregaard, K., Metzler, R., Ritter, C. M., Berg-Sørensen, K. & Oddershede, L. B. Manipulation and motion of organelles and single molecules in living cells. *Chemical Reviews* **117**, 4342–4375 (2017).

[98] Chen, B. *et al.* Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* **155**, 1479–1491 (2013).

[99] Lionnet, T. *et al.* A transgenic mouse for in vivo detection of endogenous labeled mRNA. *Nature Methods* **8**, 165–170 (2011).

[100] Chen, L.-F. *et al.* Enhancer histone acetylation modulates transcriptional bursting dynamics of neuronal activity-inducible genes. *Cell Reports* **26**, 1174–1188.e5 (2019).

[101] Nicolas, D., Zoller, B., Suter, D. M. & Naef, F. Modulation of transcriptional burst frequency by histone acetylation. *Proceedings of the National Academy of Sciences* **115**, 7153 (2018).

[102] Stellaris Probe Designer version 4.2 (2019). URL `https://www.biosearchtech.com/stellaris-designer`, Accessed: 2019-12-19.

[103] Zerbino, D. R. *et al.* Ensembl 2018. *Nucleic Acids Research* **46**, D754–D761 (2017).