Imaging the binding and diffusive behaviours of epigenetic modifiers Tet1 and Tet2 in live embryonic stem cells

---

Dissertation zur Erlangung des Doktorgrades der Naturwissenschaften
Doctor rerum naturalium (Dr. rer. nat.) an der Fakultät für Biologie der
Ludwig-Maximilians-Universität München

Joel Ryan
09.03.2020

1. Gutachter: Prof. Dr. Heinrich Leonhardt
2. Gutachter: Prof. Dr. Andreas Klingl
Tag der Abgabe: 09.03.2020
Tag der mündlichen Prüfung: 15.05.2020

Eidesstattliche Versicherung

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbständig und ohne unerlaubte Hilfe angefertigt ist.

München, den 10.03.2020

Joel Ryan

Erklarung

Hiermit erkläre ich, dass die Dissertation nicht ganz oder in wesentlichen Teilen einer anderen Prüfungskommission vorgelegt worden ist und dass ich mich nicht anderweitig ohne Erfolg einer Doktorprüfung unterzogen habe.

München, den 10.03.2020

Joel Ryan

Acknowledgements

This work came about with the help of many scientists near and far. I wish to thank the Leonhardt lab for the environment in which to carry out these projects. In particular I wish to thank two brilliant scientists, Chris Mulholland and Miha Modic, who were fantastic team mates, travel mates and friends throughout this adventure. I thank Sebastian, Irina, and Don for their guidance, scientific input and wisdom.

I have been fortunate to spend time in different scientific environments for experiments, collaboration, training, and friendliness. I thank Philipp Messer, Nader Danaf, Ivo Glück and Don Lamb in Physical Chemistry; Georgia Squyres, Alex Bisson and Ethan Garner at Harvard; Joe Donovan and Robert Kasper at the MPI Neurobiologie; and the Advanced Imaging Center team at the Janelia Research Campus. I am grateful to all the students and faculty at MBL Physiology 2016 and 2017, for the inspiring summers.

Besides lab work, there are many people to thank for their good nature and the fantastic times I had here in Munich. I thank the constant friendship and support of Ryan Sherrard, shouldering me through difficult times, and making the good times happen. My beloved Fluffable: Drago, Ksenia, Joe, and Michi D, for the weekly rock sessions, support and shenanigans. Thanks to Ash, Addie & Niels, Chris, Dave & Louise, Evi, Geri, Martina, Michi B, Miha & Tjasa, Tajda, Verena & Burak, the Müsliforschungsgruppe, Philipp, Frank and everyone in AK Lamb and AK Tinnefeld. At various times during my PhD, and at various times of the day, you all offered me welcome, warmth, and friendship. Thank you all for the good times here in Munich and around Europe. And to Beth, for the times we shared, despite the distance.

Finally, I wish to thank my family, who have supported me through this overseas adventure. Their warm love and support has kept me going throughout the years.

À mes parents, qui m’ont tout donné.

Abstract

Chromatin modifications, such as DNA methylation (5mC) and histone modifications, are epigenetic pathways which can mediate the long-term suppression of gene expression and as such are essential for embryonic development. The Ten-eleven translocation (TET) family of dioxygenases modify DNA methylation by oxidizing methylcytosine to derivatives such as hydroxymethylcytosine (hmC). Tet1 has a dual role regulating gene expression in embryonic stem cells (ESCs), where it acts as both a transcriptional activator and repressor, and promotes both pluripotency and transitioning towards lineage commitment. Tet2, on the other hand, associates with enhancers, and appears to more globally regulate methylation levels.

The transcriptional and epigenetic functions depend on the mobility of TETs and their ability to find their targets in the nucleus. Here, we use orthogonal live-cell imaging methods, including fluorescence recovery after photobleaching (FRAP), single-molecule tracking (SMT) and raster image correlation spectroscopy (RICS) to study the spatial distribution of Tet1 and Tet2 in the nuclei of live embryonic stem cells (ESCs).

We show that Tet1 and Tet2 have markedly different kinetics which likely drive their contrasting functions. In naive and primed ESCs, Tet1 is mostly bound, and forms multivalent hubs which associate at euchromatin loci. In contrast, Tet2, while also localizing to euchromatin loci, is largely diffusive and less likely to form stable interactions in the nucleus.

Remarkably, these binding and diffusion behaviours appear to be independent of the methylation state of the genome. This finding reveals that the oxidative activity of Tet proteins consist of transient events which are tightly regulated by their N-terminal domains, and their interaction with features in the nucleus that have yet to be elucidated. Taken together, we propose that Tet1 is a compact explorer of the nucleus, enhancing local activity at proximal loci, whereas Tet2 is more globally diffuse, leading to sparse activity throughout the nucleus. We show for the first time physical evidence of distinct subnuclear kinetics of Tet1 and Tet2 in living ESCs, at the single-cell and single-molecule level, which underlie the distinct roles of Tet1 and Tet2.

Finally, we investigate the binding and diffusive properties of DPPA3, whose promoter is a target of TET protein catalytic activity. We show in living cells that DPPA3 interacts with UHRF1, leading to its release from chromatin, and which in turn impairs maintenance methylation.

Introduction

1.1 Nuclear organization and epigenetics

1.1.1 Genome organization

The nucleus is a crowded and heterogenous environment which houses the genome, as well as the processes which maintain and replicate it (Misteli review 2001, Science). The genome stores all the information required for the development, maintenance and propagation of living organisms. This information is encoded in DNA, an essential macromolecule which can be organized, read, modulated or replicated by a myriad of nuclear proteins such as transcription factors, chromatin modifiers and polymerases. Biological processes such as development and cell differentiation make use of different subsets of the information encoded in the genome, and thus rely on an intricate spatiotemporal regulation of the transcriptional network (Davidson et al 2010). In this regard, how nuclear proteins such as transcription factors, chromatin modifiers and polymerases find their targets in the genome, giving rise to timely transcriptional processes, is an actively studied question (Chen and Larson 2016).

Given the size of mammalian genomes (~1×10^9 base pairs), and the number of protein coding genes (~22 000) (BioNumbers), the genome must be organized and annotated at different levels to ensure the timely expression of different proteins throughout the life of an organism. As a first layer of organization, DNA wraps around a nucleosome core particle, which is an octamer composed of core histones H2A, H2B, H3, and H4 (Kornberg 1974, Luger et al, 1997) (FIGURE 1.1A). This "beads on a string" organization can then itself be organized into different packaging conformations, consistent with the transcriptional or cell cycle requirements of the cell. One such example of a higher order chromatin structure is heterochromatin, wherein repetitive genomic regions are densely packaged and transcriptionally silent (reviewed in Woodcock and Ghosh, 2010). In contrast, euchromatin is less densely packaged, and thought to be more accessible to nuclear proteins such as transcription factors and RNA polymerase II (FIGURE 1.1 B and C).

Finally, in addition to this highly organized 3D structure, the nucleus is visibly a heterogeneous environment. In addition to the different chromatin compartments, which occupy roughly 25% of the nuclear volume (Phillips et al, Cell Biology by the Numbers), the nucleus is a crowded environment (Bancaud et al 2009), which includes nuclear proteins, RNA, and several membraneless compartments, such as nucleoli and speckles (Boeynaems et al, 2018).

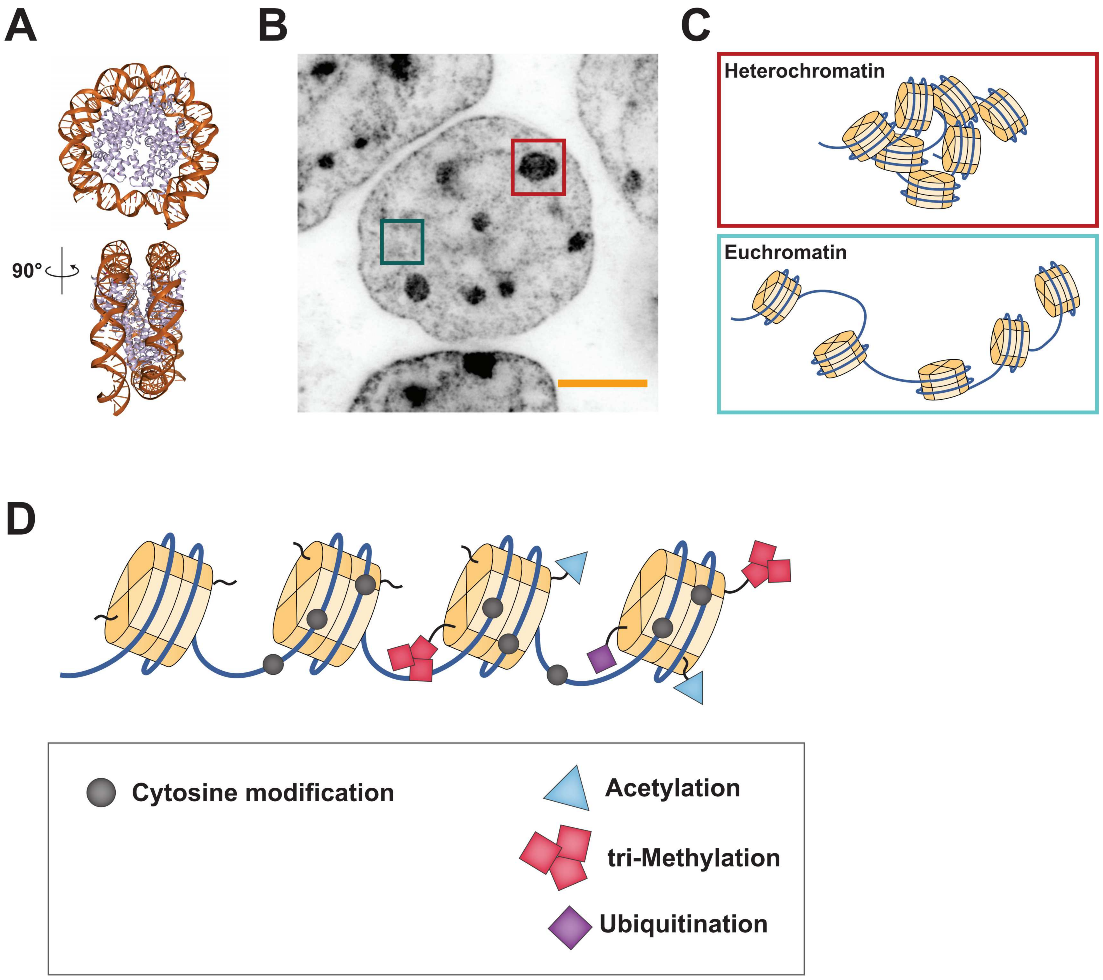


Figure 1.1 Overview of genome organization. A, 3D structure of a canonical nucleosome core particle consisting of histones (light purple) wrapped with roughly 1.8 turns of DNA (orange), PDB structure 2CV5. B, example image of embryonic stem cell nucleus, stained with SiR-DNA to fluorescently label DNA. Example regions of decompacted and dense chromatin are shown with cyan and red boxes. C. Cartoon of dense heterochromatin and decompacted euchromatin, with DNA shown as a solid blue line, wrapping around nucleosomes particles in light brown. D, In addition to DNA wrapping around nucleosome particles, DNA and histone tails can be enzymatically modified, which can alter the physico-chemical properties of chromatin. Cartoons by Michael Bartoschek.
Scale bar = 5 µm

In parallel to this architectural organization of the genome, including chemical modifications to DNA and histone tails, epigenetic marks influence gene expression at different timescales throughout development. The particular combinations of modifications, along with the expression of different subsets of transcription factors, shape cell identity, and underlies the myriad of functions accomplished by different cell types. Thus, it becomes evident that an intricate spatio-temporal regulation of the distributions of these modifications and nuclear proteins are crucial to development.

1.1.2 DNA methylation

Epigenetic marks consist broadly of post-translational modifications of histone tails, and covalent modifications of genomic DNA. DNA methylation is the first step in DNA modification, and consists of the addition of a methyl group to the fifth carbon atom of cytosines in CG dinucleotides, generating methylcytosine (mC) (reviewed in Bird 2002). Cytosine methylation is highly conserved throughout evolution, as it appears to have originated in bacteria, is thought to have existed in the first eukaryotic cells (Zemach et al, 2010), and is still conserved in plants (Feng et al 2010) and fungi (Antequera et al, 1984). Nonetheless, cytosine methylation has been lost in many organisms, including *Drosophila melanogaster*, *Caenorhabditis elegans*, *Schizosaccharomyces pombe*, *Saccharomyces cerevisiae* (reviewed in Greenberg and Bourch'is, 2019).

DNA methylation is among the most broadly characterized epigenetic modifications (Smith and Meissner, 2013). This modification is carried out by the catalytic activity of DNA methyltransferase enzymes DNMT1, DNMT3A and DNMT3B, each containing a a highly conserved catalytic domain (Bestor et al, 1988, Okano et al 1998).

DNMT1 is known to carry out “maintenance methylation”, in other words, DNMT1 propagates existing methylation marks on newly synthesized DNA strands, thereby insuring that daughter cells inherit faithfully replicated methylation marks. This maintenance methylation takes place concomitantly with DNA replication, (basic mechanism bit) as DNMT1 harbours a targeting sequence which mediates its interaction with the replication machinery (Leonhardt et al 1992). The association of DNMT1 with replication highlights it maintenance role in dividing cells, and has been shown using several imaging methods including FRAP and structured illumination microscopy (Schermelleh et al, 2007, Schneider et al, 2013). Maintenance methylation is dramatically enhanced by DNMT1’s association with UHRF1. UHRF1 is a multifunctional protein containing an E3 ligase domain, which binds hemi-methylated DNA after replication (Bostick et al 2007). An additional pathway for DNMT1 recruitment at replication sites involves PAF15, which is a ubiquitination target of UHRF1 (Nishiyama et al, 2020, *in press*).

DNMT3a and DNMT3b are known as "de novo" DNMTs, as they establish DNA methylation patterns during embryonic development (Okano 1999, Chen et al, 2003). Targets of DNMT3b include CpG islands in promoters, as well as repetitive elements. These newly methylated sequences are then largely maintained throughout rounds of cell division by DNMT1.

The role of DNA methylation as an epigenetic mark is not fully understood. In mammals, the majority of methylcytosine is found at repetitive elements, including long interspersed nuclear elements (LINES), In the context of promoters, DNA methylation is thought to inhibit transcription factor binding, which as a result, can lead to silencing (Keshet et al, 1986, Jaenisch and Jahner, 1984). However, many transcription factors, including pluripotency factor Klf4 specifically bind methylated promoters (Hu et al, 2013) in a transcription-activating manner (Wan et al, 2017). Moreover, long term silencing, while associated with DNA methylation, however, such as the silencing of pluripotency markers during lineage commitment and differentiation, appears to require H3K9 methylation (Ayyanathan et al, 2003, Feldman et al, 2006). In other genomic contexts, such as that of gene bodies, DNA methylation is typically associated with active transcription (Lister et al, 2009). Thus like many histone modifications, our understanding of the epigenetic code of DNA methylation is incomplete, and its role likely depends on the chromatin context in which they lie. Thus, DNA methylation offers a layer of fine-tuning of transcriptional regulation which is sensitive to its chromatin environment.

In mammals, methylation levels fluctuate throughout the life of the organism. Specifically, during pre-implantation, mammalian genomes undergo a wave of DNA demethylation (Monk et al 1987; Sanford et al 1987), which appears to be required for the establishment of pluripotency (Lee et al, 2014). Curiously, stem cells maintain their capacity for self-renewal in the absence of DNA methylation, however their potential for differentiation is almost completely abolished (Jackson et al 2004). Altogether many questions remain to be answered as to how changes in methylation, both globally and at gene-specific loci, conspire with other epigenetic marks to shape cell identity.

1.1.3 Tet proteins and DNA demethylation

For decades, DNA methylation was thought to be a stable epigenetic mark, unlikely to be modified throughout development. However, a growing interest in the search for methylcytosine-modifying proteins lead to a landmark discovery (Tahiliani et al, 2009). A bioinformatic search for mammalian protein homologs of two basemodifying enzymes found in Trypanosomes, JBP1 and JBP2 identified Tet1, Tet2, and Tet3 as potential candidates for methylcytosine modification. All three Tet proteins were predicted to contain a structure typically found in 2-oxoglutarate and Fe(II)-dependent oxygenases. The authors then showed that overexpression of TET1 could lead to a decrease of methylation, and that this was carried out by converting methylcytosine to a yet undescribed base. They then identified this base as 5-hydroxymethylcytosine (hmC) by mass spectrometry .

Shortly after, it was discovered that TET proteins catalysed the sequential oxidation of mC to 5-hydroxymethylcytosine, 5-formylcytosine (fC) and 5-carboxycytosine (caC), thereby revealing that TET proteins could generate a total of three oxidized cytosine derivatives (Ito et al, 2011) (FIGURE 1.2). These discoveries challenged the notion that methylcytosine was a stable epigenetic mark. Instead, TET protein activity revealed a pathway for modulating DNA methylation, and possibly DNA demethylation.

First, TET protein catalytic activity generates cytosine derivatives which can lead to demethylation through passive dilution (Saitou et al, 2012, Inoue and Zhang 2011), since maintenance methylation acts on hemi-methylated DNA between the template strand and newly synthesized strand after DNA replication. In this regard, the oxidation of methylcytosine at a locus on one strand would preclude maintenance methylation on the newly synthesized strand duration replication, thereby leading to a loss of methylation at that locus for one copy of the genome, which will be inherited by one daughter cell after cell division.

In parallel, an active pathway for DNA demethylation was discovered. It was found that TET-mediated oxidation of methylcytosine can lead to the excision of the modified cytosine by Thymine DNA Glycosylase (TDG) (He et al 2011, Maiti and Drohat 2011). Then, this excision would lead to base excision repair machinery activity, which would then insert an unmodified cytosine. This mechanism was shown in vitro and in vivo, where TDG knockout cells accumulated higher levels of caC. Thus, a multi-step active demethylation pathway was discovered, thereby adding another tangible layer of epigenetic regulation of differentiation and development in embryonic stem cells. This active demethylation pathway has been shown to re-activate silenced methylated DNA (Mueller et al, 2014), pointing towards a potential for gene expression regulation. However, it has yet to be shown that a dynamic methylation and demethylation system can directly and reversibly repress and activate genes at endogenous loci (Bestor et al, 2014).

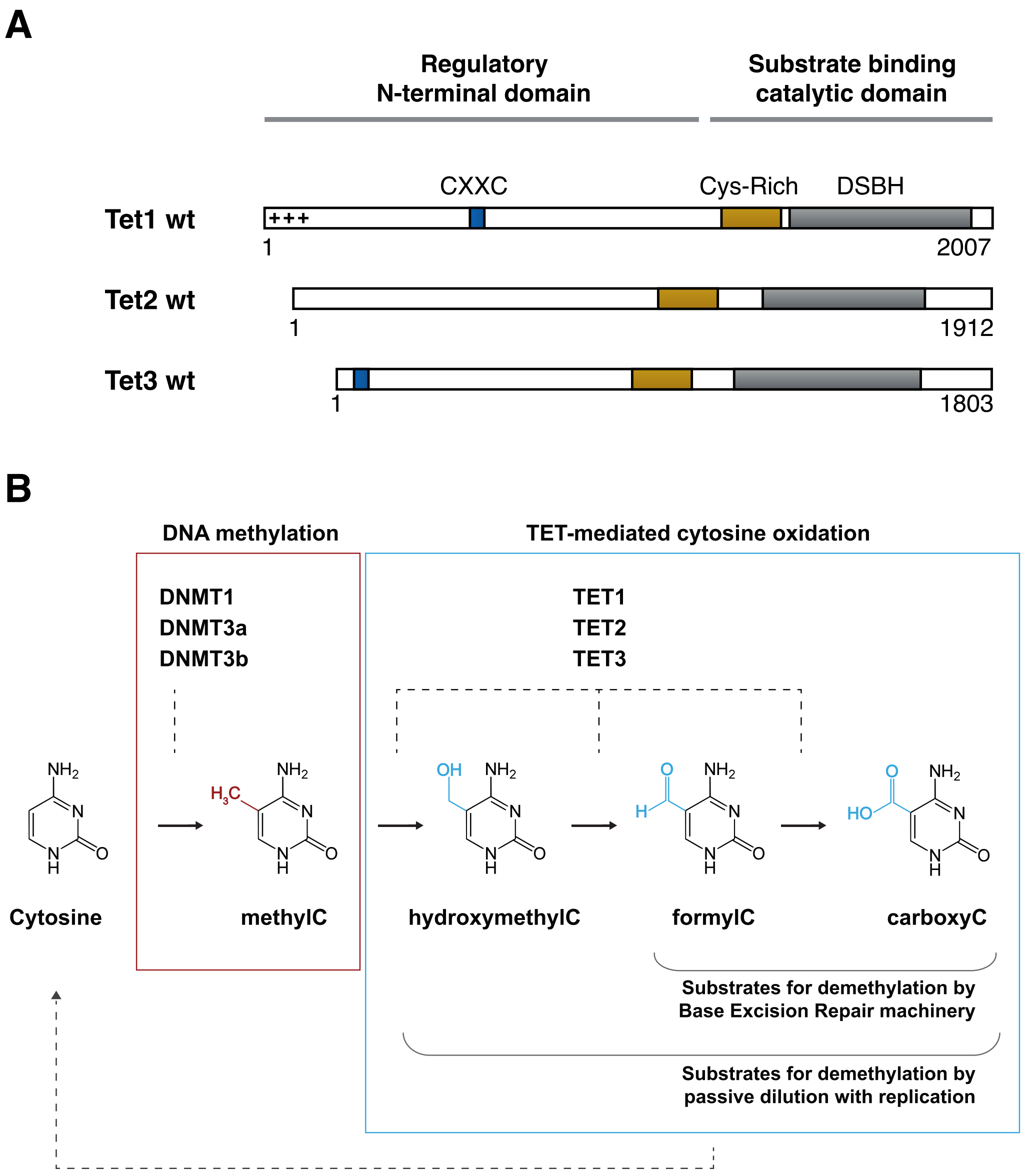


Figure 1.2 Overview of TET protein domains and cytosine modifications. A, Protein domain schematics of mouse Tet1, Tet2 and Tet3, showing positively-charged BC domain (+++), CXXC zinc-finger domains (CXXC, blue), cysteine-rich domains (Cys-Rich, yellow), and double-stranded beta-helix domains (DSBH, grey) and protein lengths for mouse TET proteins. B, Overview of cytosine methylation by DNMT proteins and further iterative cytosine oxidation TET proteins, along with canonical pathways for active and passive demethylation. Modifications relative to unmodified cytosine are indicated in red and blue.

The TET family of proteins consist of TET1, TET2, and TET3, and are found in mammals, fish, and amphibians (Bogdanović et al, 2016). TET proteins all harbour a similar global structure. The enzymatic oxidative activity of TET proteins is mediated by the catalytic domain, which is found at the C-terminal end of all TET proteins.

This catalytic domain consists of a double-stranded helix beta helix motif, as well as a spacer motif (Pastor et al, 2013). Together, this domain can bind methylated DNA or further oxidized methylated DNA (Hu et al, 2013) and modify the DNA to further oxidative derivatives, in a reaction requiring iron.

In terms of their activity in embryonic stem cells, several studies strongly hinted at a role in pluripotency maintenance, as well as lineage specification. It was shown that all three Tet proteins (TET1, TET2, and TET3) could catalyze the mC to hmC reaction in stem cells, and that this catalytic activity could play a role in pluripotency maintenance by modulating methylation levels of the Nanog promoter (Ito et al, 2010). Furthermore, Tet1 and Tet2 were then shown to contribute to establishing pluripotency via their physical association with the master transcription factor Nanog (Costa et all 2013). In parallel, it was found that Tet1 and Tet2 levels were regulated by Oct4 and that together, this network maintained hmC levels in mouse embryonic stem cells (Koh et al, 2011).

Genome-wide, two pioneering studies performing ChIP-Seq on Tet1 using different antibodies found Tet1 associated with over 35000 binding sites (Wu et al, 2011) and 6500 genes (Williams et al, 2011). The context of this association was largely the transcription start sites of CpG-rich promoters, which hinted at transcriptional regulation. In parallel, Tet1 colocalized with genes bound by Polycomb Repressive Complex 2 (PRC2) protein SUZ12, as well as promoters bound by the SIN3A repressor complex. Interestingly, the authors also mapped hmC in the same cell lines, and found, using a stringent cutoff, that only 35% of hmC positive transcription start sites are bound by Tet1, suggesting on the one hand, a strong role for Tet2 in generating hmC, but on the other hand raises the possibility that the oxidative activity of Tet proteins is a transient reaction that might not be directly captured by the cross-linking and washing steps required for immunoprecipitation. Finally, the authors also hinted at non-catalytic roles for Tet proteins, by looking at genes upregulated following Tet1 depletion in cells lacking all three DNA methyltransferases. Thus, a global picture of Tet1 and hmC localization was finally available and revealed that Tet1 could both promote transcription and repression, and a hint of non-catalytic activity was discussed.

All these studies indicated that hmC, fC and caC might be epigenetic marks modulating the chromatin environment on their own, in addition to being intermediate substrates for demethylation. In this regard, genome-wide mapping of hmC, fC and caC found that hmC mostly associated with euchromatin and active transcription (Ficz et al, 2011), and levels of mC and Tet-mediated oxidative derivatives fluctuated dramatically in TDG knockout stem cells (Shen et al, 2013). Indeed, the different chemical nature of Tet-mediated cytosine derivatives hinted that they might be chromatin modifications which could be “read” differently by different chromatin-binding proteins. In this regard, screening for hmC- and fC-binding proteins revealed a collection of transcriptional and chromatin regulators with a binding preference for fC compared to hmC. Thus, oxidized derivatives of methylcytosine appear to have distinct readers, and thereby may act as epigenetic marks (Spruijt et al, 2013, Iurlaro et al, 2013).

With some insight into the biochemistry of Tet proteins and hmC production, as well as their global localization in ESCs, it was then asked whether Tet knockout mice have a visible phenotype. This question lead to some controversy, as initially, it was suggested that Tet1 depletion lead to a decrease in hmC levels, but did not affect pluripotency or postnatal development (Dawlaty et al, 2011). In contrast, a depletion of both Tet1 and Tet2 lead to developmental defects which in turn lead a proportion of embyronic lethality, with a fraction of normal and viable mice (Dawlaty et al, 2013). In parallel, loss of Tet1 and Tet3, lead to a higher variability in the transcriptional profiles and methylation status of embryos, which highlights a fine-tuning role in transcriptional regulation for Tet proteins (Kang et al, 2015). Finally, loss of all three Tet enzymes lead to a near complete loss of hmC, as expected, and were not compatible with embryonic development (Dawlaty et al, 2014). These results highlighted the requirement of some Tet activity to support development, and suggested that there may be a partial overlap of function between Tet proteins. Furthermore, these studies hinted that Tet activity may play a fine-tuning role which can be in some cases partially dispensable, at least in the context of laboratory mouse strains.

Interestingly, the idea that Tet1 was dispensable for embryonic development was challenged a few years later where a striking lethal phenotype was reported for Tet1 knockout mice obtained from mixed-strain inter-crosses (Khoueiry et al, 2017). They then tested this by backcrossing within the same strain, which surprisingly yielded viable offspring. Then, these viable in in-bred mice were outbred again, which again lead to lethality. Thus, these experiments indicated a hidden mechanism where inbred Tet1 KO mice are viable, but outcrossed mice are lethal, blurring our understanding of a Tet1 KO phenotype, and hinting at yet unknown adaptation mechanisms.

Besides their roles in embryonic development, the oxidative activity of Tet proteins was found in the brain, as hmC was abundant in Purkinje neurons (Kriaucionis and Heintz, 2009), as well as in cells sorted from different structures of the brain (Muenzel et al, 2010). This catalytic activity is thought to play a role in gene regulation in the brain, and thus may play a role in aging and neurodegenerative diseases (Sun et al, 2014, Wen et al, 2014)

How Tet proteins find their genomic targets and bind chromatin in living cells is still unclear. A crystal structure of the catalytic domain of Tet2 in complex with DNA revealed recognition of the CpG dinucleotide by this domain as well as a cavity amenable to methylcytosine and further oxidative derivatives (Hu et al 2013). This structure was thought to have similar affinity for both methylated and unmethylated CpG dinucleotides, and no particular preference for flanking sequences. In light of this low-specificity of DNA binding, the authors suggest that genomic localization of TET proteins might instead be driven by their large N-terminal domain, as well as their interaction partners. This idea was supported by studies in which overexpression of fluorescently tagged Tet1 catalytic domain revealed stronger binding of this fragment at DNA-dense and methylation-rich heterochromatin loci (Zhang et al, 2017, Ludwig et al, 2017). Consistently, since ChIP-Seq binding profiles of endogenous Tet proteins suggest strong binding at euchromatin loci such as promoters, transcription start sites and enhancers (Williams et al, 2011, Wu et al, 2011), it is likely that the N-terminal domain of Tet proteins tightly guide their localization. In vitro, the catalytic domain of Tet2 was shown to act in a distributive fashion, wherein after enzymatic oxidation of mC, the Tet catalytic domain does not retain the modified cytosine, nor does it slide along the DNA fragment (Tamanaha et al, 2016).

Functional motifs of the N-terminal domain of Tet1 and Tet2 are poorly understood. In both cases, the majority of the N-terminal domain is unstructured, and shows little conservation with other proteins. Tet1 (and Tet3) contain a CXXC zinc finger DNA binding domain. In vitro work showing the crystal structure of the Tet1 CXXC domain in complex with DNA suggests a model in which this domain can bind DNA, preferentially at unmethylated CpGs islands (Xu et al, 2018), consistent with genome-wide binding profiles (Pastor et al, 2013). Curiously, however, overexpression of the isolated Tet1 CXXC domain in C2C12 myoblasts cells showed very little binding to chromatin, similar to unspecific binding of free GFP (Frauer et al, 2011), suggesting that in living cells Tet1 chromatin binding requires perhaps multiple domains, or a multistep process. Furthermore, Tet proteins are subject to a large series of post-translational modifications including phosphorylation, ubiquitinylation, and O-Glc-Nac-ylation, speculated to modulate the behaviour of Tet proteins in living cells (Bauer et al, 2015).

Recently, a novel isoform of Tet1 was found to be expressed in somatic cells and overexpressed in cancer (Zhang et al, 2016, Good et al, 2017). This isoform, termed

Tet1s (for somatic) lacks the first 653 amino acids, thereby missing the CXXC domain. They show that this isoform can still bind targets of full length Tet1, although less strongly. They also generate a series of deletion mutants which can surprisingly bind wild-type Tet1 targets, all less strongly than the full length protein. Interestingly, they identify a conserved positively charged domain, consisting of the first 131 amino acids, that seems to drive non-specific binding of Tet to chromatin in bulk assays.

Tet2 on the other hand, lacks any cognate DNA binding sites in its large N-terminal domain. Tet2's lack of DNA-binding domains, is thought to lead to looser binding of Tet2 to chromatin (Vella et al, 2013, Xiong et al), and a greater difficulty in generating genome-wide binding profiles. Nonetheless, Tet2 has been shown to associate with and drive hmC levels at enhancers (Xiong et al, 2016, Hon et al, 2014). Surprisingly, despite Tet2 appearing less bound to chromatin compared to Tet1, Tet2 appears to contribute to a greater fraction of hmC generation than Tet1 (Hon et al, 2014). However, depletion of Tet2 leads to much less severe global changes in expression levels compared to Tet1 (Mulholland et al, 2018). Tet2 has been found to associate with enhancers in hematopoeitic stem cells (Rasmussen et al, 2019), and mutations in Tet2 are largely associated with acute myeloid leukaemia (Weissmann et al, 2012). Interestingly, Tet2 is largely associated with naive pluripotency, and is largely down

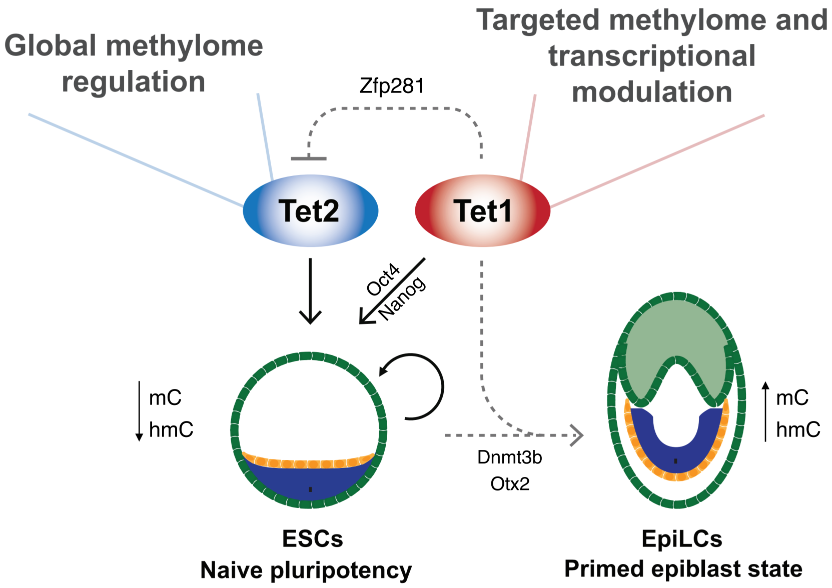


Figure 1.3 Overview of suspected roles of Tet1 and Tet2 in naive ESC self-renewal (black solid lines) and in the transition from naive to primed pluripotency (dashed lines). Tet1 and Tet2 contribute to naive pluripotency and maintaining a hypomethylated genome. When the naive to primed pluripotency is triggered, Tet1 contributes to this transition, increasing hmC production, and contributing to the downregulation of Tet2. Embryo cartoon by Sebastian Bultmann.

regulated as cells transition from naive to primed pluripotency (Hon et al, 2014), in a mechanism requiring Zfp281. Tet1, on the other hand, appears to contribute to both naive pluripotency as well further priming and lineage commitment (Fidalgo et al, 2016, Luo et al, 2020), (FIGURE 1.3).

These findings, along with the surprising finding of hmC detected beyond the binding loci of Tet detected, indicate (1) that TET catalytic activity is potentially transient and beyond what can be easily detected by cross-linking and ensemble averaging methods, (2) that Tet1 localization is more tightly targeted to chromatin loci, whereas (3) Tet2 has a more global activity.

1.1.4 DPPA3-mediated demethylation

During mammalian development, the genome undergoes large scale fluctuations of methylation levels. Shortly after fertilization, the genome first undergoes a wave to DNA demethylation. In mouse, this wave of demethylation takes place during the first three days of development, until the blastocyst stage of embryonic development (Reik et al, 2001). At this stage, cells of the inner cell mass of the blastocyst have low levels of DNA methylation, and are said to be in a naive pluripotent state. Then, a wave of *de novo* DNA methylation, driven by DNMT3a and DNMT3b, methylates the genome. This wave of *de novo* methylation takes roughly 48 hours. Methylation levels then reach their plateau as the embryo reaches the epiblast stage and a primer pluritpotency state, prior to lineage commitment.

TET proteins would be obvious candidate proteins to drive this wave of demethylation, given their enzymatic activity, which directly modulates methylation levels, and can lead to DNA demethylation. However, there is mounting evidence to suggest that TET-mediated methylcytosine oxidation is not the main driver of this demethylation wave (Amouroux et al, 2016).

These curious findings suggest that another mechanism, in parallel to Tet-mediated oxidation, may drive global DNA demethylation. In this regard, an alternative mechanism shown to drive DNA demethylation involves Developmental Pluripotency-Associated 3 (DPPA3; also known as Stella or PGC7). DPPA3 has been shown to contribute to DNA demethylation by inhibiting maintenance methylation in cancer cell lines (Funaki et al, 2014) as well as in HEK cells (Du et al, 2019), and in the early embryo (Li et al, 2018). Such inhibition is thought to be driven by DPPA3 binding and displacing UHRF1. UHRF1 is a multifunctional protein which mediates loading of DNMT1 on newly synthesized DNA (Sharif et al, 2007). Thus, by impairing maintenance methylation machinery, methylation levels are thought to be progressively diluted through sequential rounds of DNA replication and cell division.

While an interaction between UHRF1 and DPPA3 has been shown *in vitro*, whether this interaction occurs in cells is unclear. What's more, DPPA3 has been also been reported to drive the opposite function: in early embryogenesis, DPPA3 expression would protect the imprinted loci against DNA demethylation (Nakamura et al, 2007). Thus, the cellular and developmental context of DPPA3 expression appears to be critical in determining its function on methylation levels. In this regard, to understand the role of DPPA3 in demethylation in ESCs, it is critical to investigate its activity not only in directly in this cell type, but also use methods which probe its activity and interactions in living cells.

1.2 Target search mechanisms in the nucleus

1.2.1 Subdiffusion in the nucleus

Proteins such as transcription factors move in the nucleus by diffusion (Phair and Misteli). Diffusion allows a means of displacement that doesn't require the cell to input energy (ATP) (Mueller et al 2013). In the simplest model, proteins would simply undergo brownian motion until they happen to collide with their target, where they could bind and exert their function. However, such a random walk governed by brownian motion alone doesn't fully explain the observed behaviour of nuclear proteins. Notably, proteins in the nucleus appear to undergo subdiffusion, a broad term used to describe molecules that explore a smaller volume than what would be expected from a purely random walk .

Such subdiffusion appears to arise first from a tendency of many DNA binding proteins to transiently bind non-target sites along a DNA molecule. At first glance, such non-specific binding would appear to be detrimental to the output of a nuclear protein. However, surprising in vitro measurements suggested that transient binding could, in some cases, facilitate target search of DNA binding proteins. For example, the reaction rates of the lac repressor binding to the lac operon were found to be an order of magnitude higher than expected for a diffusion controlled reaction (Riggs et al, 1970). The authors hypothesized that the lac repressor may bind transiently to DNA at unspecific loci, increasing its local enrichment and on-rate, or enabling one-dimensional diffusion along the DNA molecule. Thus, it was thought that target searching could be facilitated by transient binding to unspecific loci on the DNA molecule, however there was no formal mathematical description for such a mode of diffusion.

Since these results were reproducible in different conditions, it was hypothesized that such a mechanism could exist in cells, since in a simple model of a bacterial cell, the concentration of target DNA sites is usually very low, roughly one copy of lac per cell, and this site is “hidden" in a long DNA molecule full of structurally similar non-specific non-target sites.

A mathematical model was established describing how proteins could slide, hop or jump along DNA molecules, and under what conditions facilitated diffusion would accelerate target finding. Conceptually, such a mechanism would involve a protein encountering a DNA molecule and binding transiently at a non-specific site. Then, displacement along the DNA or multiple rounds of the molecule binding and release while remaining proximity of the target would reduce the dimensionality of the search process. This reduction in dimensionality, in certain cases, would accelerate the target finding (Berg et al, 1981).

In a follow-up paper, the authors show that such a model accurately predicts the “faster-than-diffusion” kinetics of the lac repressor binding to the lac operon, specifically that the in vitro association rate scales with the length of the DNA fragment, rather than target concentration (Winter et al, 1981). Thus, they provide the first experimental evidence of a quantitatively predictable target search mechanism facilitated by transient binding to non-specific DNA sites.

Whether these mechanisms actually took place in living cells was still unknown since it was difficult to conceive and perform experiments which would give a clear indication as to the existence of this mechanism in living cells. Facilitated diffusion was shown to be a mechanism used by living cells in 2012 (Hammar et al 2012). In an elegant series of single-molecule imaging experiments, the association rate of the lac repressor with the lac operon was measured based on the timing of appearance of a visible diffraction limited spot following, and the “sliding distance” of the lac repressor was determined by adjusting the distance between two lac operon sites. For the lac repressor, they found this sliding distance to be approximately 45 bp, which they estimate increases target search efficiency ~40 fold. Thus, transient DNA binding appeared to be a mechanism used by DNA-binding proteins which increased their target search efficiency in living cells.

In mammalian nuclei, however, the number of targets for a given nuclear protein is typically much larger than in a bacterial cell, owing to a greater size and diversity of encoded functions in the mammalian genome. Thus whether facilitated diffusion was used by transcription factors in mammalian cells was a technically challenging question to address. Nonetheless, as a model, the Tetracycline repressor (TetR) showed facilitated diffusion when expressed in mammalian cells harbouring an ectopic Tet array (Normanno et al. 2015).

1.2.2 Transient binding of proteins in the nucleus

Multiple reports have shown that transcription factors and nuclear proteins undergo transient binding in mammalian nuclei both in vivo (Darzacq et al, 2007, Schmidt et al, 2016) and in vitro (Chen et al, 2014, Tafvizi et al 2001). Whether these transient binding events actually facilitated or impaired target searching wasn't fully clear (Chen and Larson 2016). For example, Pol II appears to undergo ~ 90 non-specific non-productive binding events before engaging in active transcription (Darzacq et al, 2007). Such a failure rate could be viewed not as enhancing target search, but rather as a regulatory mechanism to prevent spurious transcription. In any case, a general view emerged that nuclear proteins "scan" the genome by transiently binding to different loci, until they bind more stably to their cognate targets (Mueller et al, 2013).

This view of "genome scanning" suggested that on one hand DNA-binding domains might intrinsically exhibit a broad distribution of binding probabilities based on the binding energy potential, which is given by the underlying DNA sequence (Slutsky and Mirny, 2004). In these cases, even if a nuclear protein has a "preferred" target DNA sequence, the structural similarity of nearly any other sequence would likely lead to distribution of binding probabilities on different sequences leading to non-specific transient binding, sliding and/or hopping, rather than an all-or-none binding scenario. In parallel, DNA-binding proteins could also harbour multiple DNA-binding domains which could respectively drive "general/unspecific" DNA binding, as well as targeted "specific" binding. Such a distinction between global and targeting domains was shown for Cbx7 (Zhen et al, 2016), Suz12 (Youmans et al, 2018), Sox2 (Chen et al, 2014), CTCF (Hansen et al, 2017), by single-molecule imaging and was even suggested for Tet1 based on bulk assays (Zhang et al, 2016).

Given that transient binding events are so prevalent in DNA-binding proteins, it follows that the geometry and architecture of chromatin is intrinsically linked to target search mechanisms. The organization of chromatin, and the geometry in which nuclear proteins would diffuse was appreciated with a series of experiments involving quantifying the mobility of inert fluorescent tracers (GFP monomers, oligomers, dextrans, etc) using Fluorescence Correlation Spectroscopy and FRAP (Bancaud et al. 2009). Here it was postulated that chromatin acted as a primary crowding agent in the nucleus, and that the fractal organization of chromatin would leads to size-independent confinement of nuclear proteins, consistent with reports of subdiffusive behaviour of many nuclear proteins. In another study, single-molecule trajectories of fluorescently labeled streptavidin suggested that chromatin density was not a barrier to diffusion, but simply restricted the accessible volume. This in turn would contribute to different local concentrations of reactants, which could locally govern reaction rates (Grünwald et al, 2008). Thus, using inert fluorescent tracers, the geometry of chromatin itself was probed, which lead to modeling how the geometry of chromatin could influence target search mechanisms which indicated that the spatial arrangement of target sites, would influence target search kinetics (Benichou et al, 2010).

Such contrasting search mechanisms were later shown experimentally. By performing single-molecule tracking and analysis of trajectories of two transcription factors, c-Myc and PTEFb, it was shown that c-Myc seemed to globally explore the nucleus, by 3D diffusion with short bouts of transient binding, whereas PTEFb was more likely to be confined, and resample a local volume (Izeddin et al, 2014). Thus, PTEFb was seen as a compact explorer, whereas c-Myc was viewed as a global "non-compact" explorer. As a result, the (mean first passage time) probability of a non-compact explorer to find a target is much less dependent on the initial distance between the molecule and its target than a compact explorer (Benichou et al, 2010).

Thus, target search mechanisms of nuclear proteins are dependent on the geometry of the nucleus. As such, target search mechanisms can be broadly divided into two categories. On the one hand, global searching or non-compact exploration implies that binding events would be interspersed with longer bouts of 3D diffusion which would leave potential target sites unvisited. In contrast, compact explorers are more likely to have frequent binding events within a local volume, thereby sampling and re-sampling the same genomic loci (FIGURE 1.4).

As a result, the probability of a non-compact explorer to find a target is much less dependent on the initial distance between the molecule and its target than a compact explorer. Thus a non-compact explorer can exert its function more globally, while, in contrast, a local explorer is more likely to exert its activity on nearby targets (Izeddin et al, 2014). This resampling local activity, along with potential clustering of co-regulated sequences (Sproul et al, 2005), and is more prone to burst-like activity due to transient local enrichments (Meyer et al, 2012).

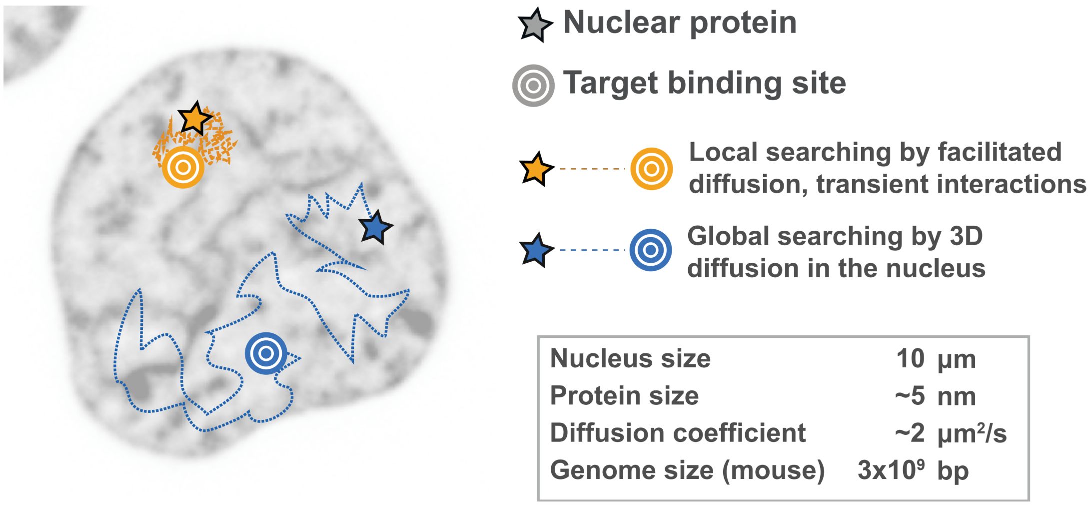


Figure 1.4 Left, schematic overview of trajectories representing global (blue) and local (orange) target search mechanisms of nuclear proteins, overlaid on a representative image of the of a nucleus (embryonic stem cell labeled with HaloH2B). Local compact explorers will sample and resample a smaller volume, whereas global non-compact explorers diffuse through a larger fraction of the nucleus before binding to a target. Box, with order-of-magnitude scales of sizes that key properties relevant to how such nuclear proteins must navigate the nucleus.

Thus, when considered in parallel with regulation of proteins through post-translational modifications, this model indicates that how nuclear proteins exert their activity and carry information spatially depends on their target search mechanism, which is exquisitely linked to chromatin architecture.

1.3 Live-cell imaging techniques to study protein dynamics

1.3.1 Fluorescence recovery after photobleaching

Fluorescence recovery after photobleaching (FRAP) is a commonly used imaging method to probe the mobility of fluorescent molecules in living cells (Lippincott-Schwartz et al, 2018). FRAP was originally described experimentally and theoretically in the pioneering work of Axelrod et al (1976). There, the authors showed that by briefly using a focused a high intensity laser to irreversibly photobleach fluorescent molecules within small area on a cell membrane, a readout of protein mobility could be derived from the apparent recovery of fluorescence at that spot over time. In this scenario, the fluorescence recovery would simply be the exchange of bleached molecules in the irradiated area with unbleached molecules from the surrounding area.

As with many live-cell fluorescent imaging methods, the development of GFP (Chalfie et al 1994) and GFP fusion proteins made these accessible to cell biology. The first studies to use FRAP with GFP-labeled proteins revealed the diffusion of proteins in the membranes of the Golgi apparatus (Cole et al 1996), and at the nuclear membrane (Ellenberg et al 1997). Given the simplicity of producing GFP-fusion proteins and performing FRAP on commercial laser-scanning confocal microscopy systems, the method gained broad appeal at the interface of cell biology and biophysics (Lippincott-Schwartz et al 2018).

The analysis of the fluorescence recovery yields a ensemble averaged view of protein mobility (Mueller et al, 2013). On the one hand, rough comparison of recovery curves from different fluorescent species acquired with equivalent settings can reveal quantitative differences in mobility of proteins, or different mutants of a given protein. On the other hand, fitting recovery curves to different mobility models can yield physical parameters of protein mobility, including average diffusion coefficient, on-rate, and off-rate. Such fitting models include, for example, pure diffusion and reaction-dominant models. Importantly, the selection of the fitting models is absolutely critical to deriving proper estimates of the physical parameters underlying the proteins mobility and recovery (Sprague et al 2004, Mueller et al, 2013), and selecting the wrong model can lead to estimates that are orders of magnitude different compared to those obtained with more sensitive methods.

Altogether, FRAP has been used to investigate the dynamics of fluorescently labeled proteins in nearly every compartment of the cell, and of course in the nucleus (Schneider et al, 2013, Darzacq et al, 2007, Phair and Misteli, 2001). It is still routinely used today, typically with with laser-scanning confocal microscope systems, or with camera-based widefield and spinning-disk confocal systems equipped with a galvo-scanning system to irradiate user-specified regions of interest.

1.3.2 Single-molecule tracking

Single-molecule tracking (SMT) is a microscopy tool commonly used to investigate the diffusive properties of molecules in the nucleus, cytosol, and membranes, and has been used in a wide array of cell typed and organisms. It is derived from single-particle tracking (SPT), which entails following the movement of a densely fluorescent particle over multiple sequential frames in a time-lapse image series.

In imaging, the signal from a single, stationary molecule has a characteristic shape defined by the point spread function of the microscope. Due to the diffraction limit, light emitted from a point source, can only be focused into a spot of roughly ~250 nm, even with the finest optics. The intensity profile of this spot can be approximated by a Gaussian curve. As a result, the centroid of this spot yields an approximation of the localization of the molecule, and the accuracy of this localization depends on the signal to noise.

Thus, even though the point spread function or the spot detected on a camera is roughly two orders of magnitude larger than the size of a protein, it is possible to estimate the localization of that protein, provided there are no other fluorescent molecules within roughly 300 nm in xy. This simple principle was key to the development of single molecule localization microscopy (SMLM) techniques, such as Photoactivation Localization Microscopy (PALM; Betzig et al 2006), and Stochastic optical reconstruction microscopy (STORM; Rust et al 2006). In these systems, only a sparse subset of proteins are fluorescent at a given time, meaning that individual molecules can be localized at the centroid of their diffraction-limited signal. By acquiring hundreds or even thousands of images with only a stochastically activated subset of fluorescent molecules, a localization map can be generated, providing spatial information beyond the resolution limit.

Single-molecule tracking, as commonly used today, was developed following the establishment of SMLM and SPT (Manley et al, 2008). Essentially, when only a sparse subset of fluorescently-labeled proteins are visible at a given time, and a rapid time-lapse image series can be acquired, and the movement of individual molecules can be traced from one frame to the next, until the leave the focal plane, or irreversibly photobleach.

Several methods are used to achieve sparse labeling. First, similar to PALM, photoactivatable or photoswitchable fluorescent protein such as mEos can be genetically fused to proteins of interest. In this context, low intensity activation light is used to stochastically activate a sparse subset of fluorescent proteins, thereby generating spatially separate fluorescence molecules. Alternatively, proteins of interest can be fused to self-labeling enzymes, such as Halo, SNAP or CLIP. These labels are inherently non-fluorescent, and bind an ectopically added ligand conjugated to bright, photostable organic dyes (Grimm et al 2015). Thus, sparse labeling can be achieved by incubating cells expressing Halo/SNAP-labelled proteins with picomolar concentrations of fluorescent ligands, yielding a subset of fluorescently-labeled, spatially-separate proteins. Finally, more recently, bright photoactivatable organic dyes conjugated to Halo-ligands were developed, and which permits the use of photoactivation with self-labeling enzymes, and engineered bright fluorophores (Grimm et al, 2016).

A wide array of different fluorescence microscopy systems can be used to visualize single-molecules in live cells, provided a reasonable signal-to-noise can be generated within an appropriate timeframe. Typically, a simple widefield setup with TIRF illumination provides high signal-to-noise and low background enabling single-molecule imaging of proteins in or near the basal membrane of living cells in culture. Such a system can also be used to image proteins in the nucleus by simply changing the angle of the excitation laser to HiLo or oblique illumination, thereby illuminating a portion of the cell roughly 1-2 μm from the surface of the coverslip. A spinning disk confocal can typically perform single-molecule tracking, however longer exposure times are required to collect enough light to detect the individual fluorophores (Martin et al, 2013). A lattice light sheet microscope (Mir et al, 2018) has also been used to image transcription factor dynamics in live *Drosophila* embryos.

In order to obtain 3D information on the molecule's localization, systems such as the 3D aberration-corrected multifocus microscope (Abrahamson et al, 2013) have been used to quantify the diffusive dynamics of Cas9 (Knight et al, 2015) and Sox2 (Chen et al, 2014). PSF engineering has also been used to acquire 3D information from single-molecule trajectories. In these systems, a phase mask is placed in the emission pathway which modifies the shape of the point spread function, in such a way that the xy-properties of the PSF convey information about the z-position of the molecule (Pavani et al, 2009). Finally, the lattice light sheet microscope has also been used for 3D imaging of single-molecules (Liu et al, 2014).

In terms of analysis, single-molecule tracking analysis typically involves, first, detecting single-molecule spots in each image of the time-lapse, then connecting those detected spots from one frame to the next in order the generate trajectories. When short exposure times and fast acquisition rates are used (over 25 frames per second) quantitative information about the diffusive behaviour of labeled proteins can be extracted from the trajectories. Essentially, the apparent diffusion coefficient of diffusive molecules can be calculated from the distribution of step sizes from a set of trajectories as per Einstein's theorem on Brownian motion:

$$D = \frac{1}{6\tau}\langle r^2 \rangle$$

where τ is the interval between exposures, and $<r^2>$ is the average of squared step sizes. However, in the presence of trajectories corresponding to bound particles, the average diffusion coefficient calculated here would be influenced by the short steps of bound molecules, thereby yielding inaccurate diffusion coefficients (Zhen et al, 2016). To solve this, different fitting (Hansen et al, 2017) and sorting (Persson et al 2014, Monnier et al, 2015) methods have been developed to determine populations of “bound” and “diffuse” trajectories. These fitting and sorting methods are critical to describing the displacements of molecules. First, such a fitting procedure yields information about the fraction of bound and diffuse trajectories, and second, this leads to a more accurate measurement of the diffusion coefficient, thereby allowing a more accurate assessment of protein diffusive behaviour.

In parallel, residence time information can be extracted from the trajectory of bound molecules. In this scenario, the duration of a track can be used as a readout for residence time, when the mean squared displacement of the track is within the confinement of the accuracy of the imaging system (Gebhardt et al, 2013). Since track duration is influenced by photobleaching, residence time measurements are typically performed with longer exposure times and lower laser intensity settings. First, longer exposure times lead to "motion blur" of diffusive molecules, thus enriching bound molecules in the analysis pipeline. Second, lower laser intensities lead to diminished photobleaching, enabling more accurate quantification of residence times, although photobleach correction is required to accurately estimate binding times.

In addition to studies using SMT measurements to elucidate general properties of DNA binding proteins in eukaryotic nuclei (Izeddin et al, 2014, Normanno et al, 2015), SMT measurements have contributed to elucidating recruitment/activity mechanisms of several nuclear proteins. For example, SMT measurements on Sox2 and Oct4 revealed a cooperative binding mechanism between these pluripotency factors, wherein Sox2 binding facilitates the binding of Oct4 at genomic loci (Chen et al 2014). SMT measurements on telomerase maintenance proteins TRF and TERT have shown that transient binding at telomeres maintains an enrichment of these proteins at the chromosome end. Thus, the chromosome end acts as a sink for telomere elongation proteins, enabling telomere maintenance at low global concentrations of telomere proteins (Schmidt et al, 2016). Single-molecule imaging of Pol II revealed that transient binding of roughly 80 molecules of Pol II drive transcription initiation and elongation (Cho et al, 2014).

In terms of epigenetic proteins, SMT was used to investigate to dynamics of Suz12 and EZH2, two proteins of PRC2, involved in the deposition of the repressive H3K27me3 mark (Youmans et al, 2018). Here the authors show that, on the one hand, the chromatin-binding activity of Suz12 is unaffected by the absence of its substrate H3K27 methylation, indicating that chromatin-binding is intrinsically encoded in the protein. Furthermore, their results suggest that freely diffusing (or very transiently binding) PRC2 is capable of generating non-specific H3K27 methylation marks. Thus, in these cases transient binding may be sufficient for its enzymatic activity, while too rapid to detect via cross-linking methods such ChIP-Seq.

Finally, a recent study using SMT revealed the nanoscale organization of Fyn kinase, a signalling hub protein involved in memory formation. In dendritic spines of live neurons, they found that the diffusive states, and nanoscale organisation of Fyn are broadly regulated by Tau, and perturbed in the presence of a pathological mutant of Tau involved in Alzheimer's disease (Padmanabhan et al, 2019). These results directly demonstrate that SMT can further our understanding of pathological states

Thus, SMT is a powerful method to measure the dynamics of proteins in live-cells. In addition to revealing the molecular determinants of chromatin binding for several nuclear proteins, SMT measurements have contributed to establishing different novel models for how proteins move inside the nucleus (Saxton 2019, Hansen et al, 2019), which steadily unveil the spatiotemporal regulation of nuclear proteins.

1.3.3 Fluorescence Correlation Spectroscopy

Fluorescence Correlation Spectroscopy methods make use of the temporal information and fluorescence intensity fluctuations typically acquired on a laser-scanning confocal microscope to measure the dynamics of fluorescently labeled molecules.

The first conceptualization and experimental implementation of FCS came about 50 years ago (Elson and Magde, 1974, Magde and Elson, 1974). The authors showed that, assuming that the fluorescence intensity detected in a single recording is proportional to the number of fluorescent molecules, and that the sample volume is essentially the excitation volume, how fluorescence intensity fluctuates over time is a result of the concentration of the fluorescent molecules and their diffusion coefficient. Thus, with a large enough data set of fluctuations of fluorescence intensity, concentration an diffusion coefficients can be derived from the temporal decay of these fluctuations by calculating the autocorrelation function (ACF), which indicates how a signal correlates with itself at different time intervals. The general equation for autocorrelation (G) of a signal at different time intervals is:

$$G(\tau) = \frac{\langle F(t)F(t+\tau)\rangle - \langle F(t)\rangle^2}{\langle F(t)\rangle^2}$$

where, $F(t)$ is the fluorescence intensity at a given time, $F(t + \tau)$ is the fluorescence intensity after a time interval τ, and square brackets denote the average of these values, typically acquired from a large dataset. Then, a model is fit to the ACF, taking into account the size of the excitation volume and the timing of the acquisition, in order to determine the concentration and diffusion rate which best describe the data. Since this initial description, FCS has been expanded and the development of sensitive detectors and intracellular fluorescent proteins paved the way for measuring diffusive behaviours using FCS in living cells (Schwille et al 1999).

An extension of FCS for measurements in living cells is Raster Image Correlation Spectroscopy (RICS), wherein spatio-temporal information of fluctuations are derived from a raster scan image on a laser scanning confocal microscope (Digman et al, 2005, Brown et al, 2008). From this image, spatio-temporal information about the labeled molecules is derived from the two-dimensional autocorrelation function. The x- and y-dimensions of the ACF respectively yield information at two different timescales: the nature of the raster scan results in neighbouring pixels on the same line having time intervals on the timescale of the pixel dwell time (typically ~10 μs), whereas neighbouring pixels on different lines have time interval on the timescale of a whole line, usually on the order of ~5 ms. Thus, the shape of the ACF yields information about the concentration, diffusion coefficient, and bound fraction of the fluorescently labeled molecules (Digman et al, 2005).

In the context of two or three proteins of interest labeled with spectrally distinct fluorophores, their interaction or complexing can be detected by calculating the cross-correlation between these different fluorescent species. To this end, Pulsed Interleaved Excitation, wherein pulsed lasers are synchronized then slightly delayed from one another, enables quasi-simultaneous excitation of two distinct fluorophores, while using microtime information to discard signal acquired from excitation cross-talk (Mueller et al, 2005, Hendrix et al, 2013).

These methods have been used in live-cells to measure the mobility, stoichiometry, and interactions of HIV Gag protein (Hendrix et al, 2015), which revealed nucleation steps prior to viral capsid assembly. In parallel, RICS and pair-correlation spectroscopy was used to dissect the mobility of transcription factors (Clark et al, 2016), and how they are transported between neighbouring cells in plants. Using photoactivatable FCS, it was found that the distribution of residence times of Sox2 were predictive of cell fate between different cells of the four-cell embryo. Cells with longer Sox2 residence times gave rise to more pluripotent progeny than cells with shorter Sox2 residence times (White et al, 2016). Recently, RICS measurements revealed that HP1a is less mobile near boundaries heterochromatin, compared to both the inside of heterochromatin and to nucleoplasm. Such a finding underscored the liquid-like phase separated properties of heterochromatin (Strom et al 2017), and thus further our understanding of how such a nuclear compartment is formed, and how phase separated droplets influence protein dynamics in the nucleus.

Importantly, FCS methods using Pulsed Interleave Excitation as well as time-correlated single-photon counting units yields information about the lifetime of the fluorescently labeled molecules under investigation (Hendrix et al, 2015). Fluorescence lifetime is influenced by the chemical environment of the fluorescent molecule, and by the proximity of neighbouring fluorophores. Thus lifetime information can be used to measure Forster Resonance Energy Transfer (FRET) between fluorescent molecules, allowing to accurately measure distances between proteins on the order of <10 nm.

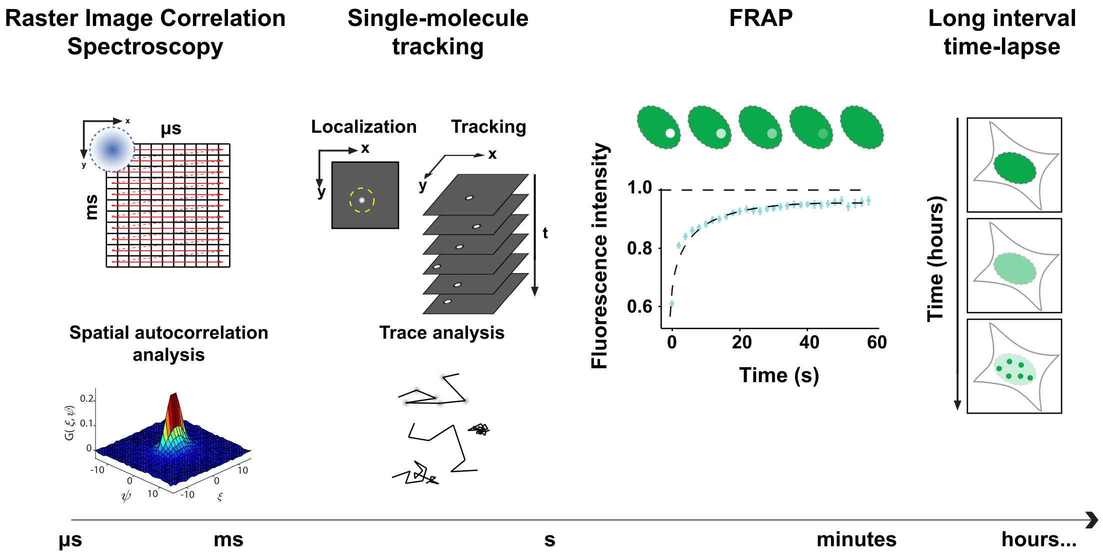


Figure 1.5 Overview of live-cell fluorescence-based imaging methods used in the work to study protein dynamics in live-cells at different time scales. Acquisition or analysis schemes are shown over the range of timescales at which they provide relevant information of protein distribution. From left to right, fast to slow timescales, Raster Image Correlation Spectroscopy, consisting of confocal scanning and analysis of the spatial autocorrelation function; single-molecule tracking consisting of the localization of individual molecules within a time-lapse image series, followed by determining their trajectories from one frame to the next, and extracting dynamics from trajectory analysis; FRAP wherein the exchange over time of bleached and unbleached fluorescently-tagged molecules give an indication of global protein mobility; and finally long-interval time lapses, for example an image taken every hour, revealing large-scale changes in protein distribution as a result longer time-scale biological processes, such as cell cycle progression and pluripotency state transitioning.

Thus, FCS-derived methods allow the measurement of physical behaviour of fluorescently-labeled proteins in living cells, which yields average measurements at the single-cell level.

1.3.4 Orthogonal methods to study the behaviour of nuclear proteins

These imaging techniques have different assumptions, leading to different advantages and limitations (Mazza et al, 2012, Liu and Tijan 2018), and yield information over different length and timescales (FIGURE 1.5). As a result, multiple studies have used more than one of these methods to dissect the binding and diffusive dynamics of nuclear proteins. For example, the kinetics of CTCF were studied using single-molecule tracking at two different timescales and FRAP in order to measure to diffusive properties and residence times (Hansen et al 2017). SMT and FRAP were also used in combination to determine hopping behaviour of cohesin protein Scc2, and how the behaviour of individual proteins (measured by SMT) translated to the collective behaviour of bulk Scc2 in living cells (measured by FRAP) (Rhodes et al, 2017). FRAP and FCS were used to measure the behaviour of p75 (Hendrix et al, 2011). In parallel, biophysical behaviour properties of Sox2 was elucidated using 3D-SMT along with FCS as an orthogonal method to measure diffusion coefficients and concentration, and integrate this data into a framework of Sox2 binding and diffusive behaviour (Chen et al, 2014). Thus, the use of orthogonal methods imaging methods leading to converging results increases the robustness of resulting models, and likely to their reproducibility.

1.4 Aims of this study

Our understanding of protein function relies on the integration of data obtained using different methodologies. Microscopy along with quantitative image analysis offer a powerful tools to investigate the spatial and temporal distribution of proteins within cells, and the inherently visual nature of imaging data greatly influences our thinking and the hypotheses we generate.

TET proteins have been shown to have distinct functions in pluripotency maintenance and transitioning towards lineage specification. Specifically, Tet1 is associated with both pluripotency maintenance as well as promoting as transition out of pluripotency towards lineage specification. In contrast, Tet2 is associated largely with pluripotency maintenance (Fidalgo et al, 2016), and its ectopic expression can facilitate reprogramming (Costa et al, 2013). How do two homolog proteins, with globally similar structure and catalytic activity drive different biological activity?

To date, our understanding of TET protein behaviour stem from experiments which used cross-linking, bulk lysis, and/or consecutive washing steps, which obscure the underlying distribution of processes.

In parallel, it is increasingly appreciated that the activity of nuclear proteins is tightly linked to their binding and diffusive properties, and the geometry of the volume in which they diffuse. Therefore, an understanding of their diffusive properties is critical to better understanding their activity. Live-cell imaging methods are best suited to observing these biophysical behaviours of proteins, and allow robust quantification of the physical parameters describing these behaviours at the single cell and single molecule level.

Here I aim to determine how the underlying diffusive and binding properties of TET proteins underlie genomic and biochemical activity inferred from large-scale ensemble averaging methods. To this end, we measured biophysical behaviours of TET proteins in live-cells as they bind and diffuse through the nucleus, with an emphasis on obtaining measurements of physical quantities in standard units. Using orthogonal live-cell imaging methods at different timescales, we captured the first order kinetics of TET proteins. We find that Tet1 and Tet2 show different physical behaviours in living ESCs, underscored by increased chromatin binding of Tet1 compared to Tet2. These behaviours indicate distinct target search mechanisms for Tet1 and Tet2. Tet1 appears to undergo compact exploration of the genome by resampling and rebinding the proximal areas. In contrast, whereas Tet2 appears to undergo non compact exploration, where binding events are separated by longer bouts of 3D diffusion. Surprisingly, these search mechanisms appear to be methylation-independent, and are rather encoded within the proteins themselves. In this regard, a short positively-charged stretch of amino-acids at N-terminal end of Tet1 drives transient binding events which enhances compact exploration and mitotic chromosome binding. We thus provide direct quantitative evidence that the N-terminal domain of Tet1 guides its localization and search kinetics. Importantly, our results provide a biophysical framework which underlies the biological activity of Tet1 and Tet2 in ESCs.

Next, we investigate the biophysical behaviour of DPPA3, a protein whose expression is regulated by TET proteins. DPPA3 was recently reported to influence methylation levels by inhibiting UHRF1 (Du et al, 2019), a protein critical to maintenance methylation. Using RICS, we demonstrate a clear interaction between DPPA3 and UHRF1 in living cells, where DPPA3 displaces UHRF1 off chromatin, and sequesters it in the cytosol. This displacement of UHRF1 leads to a decrease in maintenance DNA methylation, thereby showing that TET proteins can regulate methylation levels indirectly, by stimulating the expression of a methylation inhibiting protein.

Finally, I conclude by discussing open questions and follow-up experiments with regards to Tet proteins and high resolution live-cell imaging.

Table 1.1 Diffusion rates and bound fractions of nuclear proteins, as measured in living cells using SMT or FCS.

Protein	D_{free} (um^2/s)	F_{bound}	Length (a.a.)	Reference
EZH2	**2.09**	**0.2**	**746**	Youmans et al 2018
SUZ12	**1.99**	**0.25**	**739**	Youmans et al 2018
Sox2	**2.7**	**0.24**	**317**	Liu et al 2014
CTCF	**2.5**	**0.6**	**727**	Hansen et al 2017
Rad21	**1.5**	**0.55**	**631**	Hansen et al 2017
TERT	**1.3**	**0.14**	**1132**	Schmidt et al 2016
TRF2	**1.3**	**0.23**	**542**	Schmidt et al 2016
Zelda	**1.55**	**0.49**	**1596**	Mir et al 2018
Bicoid	**1.69**	**0.5**	**494**	Mir et al 2018
RNA PolII	**3**	**0.35**	**1930**	McSwiggen et al 2019
SHORTROOT	**2.4**	**N/A**	**531**	Clark et al 2016

Methods

Cell culture

J1 mouse ESCs were cultured as previously described (Mulholland et al, 2015). Briefly, all ESC lines used in this study were J1 embryonic stem cells from mouse. They were routinely maintained in Serum 2i LIF medium, containing of DMEM high glucose (Sigma-Aldrich) supplemented with 50U/ml Penicillin-Streptomycin (Sigma-Aldrich), 2 mM L-Glutamine (Sigma-Aldrich), 1X non-essential amino acids (Sigma-Aldrich), 50mM beta-mercaptoethanol (Sigma-Aldrich), 20% Fetal Bovine Serum (Sigma-Aldrich), 1 μM PD032591 and 3 μM CHIR99021 (Axon Medchem, Netherlands), and Leukemia Inhibitory Factor (LIF, produced in lab). Cells were passaged every 2 - 3 days by washing with PBS (Sigma-Aldrich), then detached using minimal coating volumes of Accutase (Gibco) (for example, 300 μl for one well of a standard 6-well dish), and incubated at 37°C for 5 minutes. Cells were then briefly triturated with a P-200 pipette, and one tenth of the Accutase volume was then transferred to a fresh, gelatin-coated dish.

For cells transitioning between from naive state to an Epiblast-like state, cells were cultured in N2B27 media, consisting of 50% Neurobasal (Life Technologies), 50% DMEM/F-12 Glutamax (Life Technologies), 2 mM L-glutamine (Life Technologies), 0.1 mM β-mercaptoethanol (Life Technologies), N2 supplement (Life Technologies), B27 serum-free supplement (Life Technologies), 50U/ml Penicillin-Streptomycin (Sigma-Aldrich). Naive cells were cultured on dishes coated with 0.2% gelatin, in "N2B27 2i LIF" media, consisting of N2B27 supplemented with 1 μM PD032591 and 3 μM CHIR99021 (Axon Medchem, Netherlands), 1000 U/mL recombinant Leukemia Inhibitory Factor (LIF, Millipore), and 0.3% Bovine Serum Albumin (Gibco) for at least 4 passages. To induce the transition to the epiblast-like state, cells were passaged onto plates coated with Geltrex (Life technologies), in "N2B27 FGF ActA" media consisting of N2B27 media supplemented with 10 ng/ml FGF2, 20 ng/ml Activin A, and 0.1X Knockout Serum Replacement for 48 hours.

Cloning

PCRs were carried out using Phusion polymerase (Thermo) as per manufacturers instructions, 35 cycles consisting of denaturing for 15 s at 95°C, annealing 15 s temperatures based on oligos, and extension for 20 s/kb at 72°C.

Donor plasmids for direct insertion of the coding sequences of Halo or SNAP were generated from pUC57-Tet1-mNeonGreen-Tet1 and pUC57-Tet2-mNeonGreen-Tet2 donor plasmids previously generated for endogenous tagging of TET proteins with mNeonGreen (AG Bultmann Labguru numbers). In these plasmids, the mNeonGreen coding sequence was replaced with HaloTag coding sequence. The HaloTag was amplified by PCR from plasmid pET302-6His-dCas9-Halo (Addgene number 72269) using ATGGCACGAAGTGGGCTCTGAAATCGGTAC and reverse primer CTCAGGCCGCTACCGGAAATCTCCAG generating overhangs homologous to the donor plasmid sequences upstream and downstream of mNeonGreen. The backbone of the mNeonGreen donor plasmids were amplified by PCR with CACTTCGTGCCATTCGATTTTCTG on the 5′ side of mNeonGreen and AGCGGCCTGAGGAGCAGAGCC on the 3′ side on mNeonGreen. The HaloTag coding sequence and Tet1 or Tet2 donor backbone were ligated using Gibson assembly, (described further), thereby generating pUC57-Tet1-HALO-Tet1 and pUC57-Tet2-HALO-Tet2 (AGL plasmid database pc4356 and pc4357).

The SNAP-tag was amplified from attb-SNAP-PuroR (AGL plasmid database pc3656) using forward primer ATGGCACGAAGTGGACAAAGACTGCGAAATG and reverse primer CTCAGGCCGCTCTCGAGGGATCCTGGCGC, in order to generate overhangs compatible with the donor plasmids. This PCR product was then ligated with the donor backbones of Tet1 and Tet2, amplified with the same primer pair as described for the insertion of the HaloTag.

The donor plasmid to generate cells expressing HaloTet1Δ1-131 was generated from pUC57-Tet1-HALO-Tet1, wherein the coding sequence of the first 131 amino acids was removed. To this end, pUC57-Tet1-HALO-Tet1 was amplified by PCR with AAGCGATCGCATGAAGCAGTGTACACATAATATC and AAGCGATCGCCCATGGTGGCGTCCACGGCG to generate a single a fragment containing the plasmid without the coding sequence of the first 131 amino acids, along with an overhang compatible with self-ligation.

In all above cases, PCR amplified products were first treated with DpnI to get rid of template plasmids, and cleaned up using a column-based PCR cleanup kit (Machery-Nagel). Ligations using the Gibson Assembly kit or HiFi cloning kit (NEB) were carried out at 50°C for an hour, as per the manufacturer's protocol. Ligations were transformed into JM109 chemically competent cells with 1 minute incubation at 37°C, and plated on gentamycin-containing LB/Agar plates.

Cloning for MIN-tag cell lines

A Bxb1-compatible plasmid containing mNeonGreen was generated by swapping eGFP from the attb-GFP plasmid with mNeonGreen. To this end, mNeonGreen was amplified from a template obtained from Marc Bramkamp's group, using primers GTGAGCAAGGGCGAGGAGGATAAC and GCTCTTGTACAGCTCGTCCATGCC to amplify mNeonGreen, and GTTATCCTCCTCGCCCTTGCTCAC and

GGCATGGACGAGCTGTACAAGAGC to amplify the backbone of attb-GFP, without the GFP coding sequence. Fragments were ligated by Gibson assembly, and verified by sequencing. Then, the mNeonGreen fragment was cut out of attb-mNeonGreen with SalI and AsiSI restriction enzymes (Fermentas), and inserted with T4 DNA ligase (Thermo) into vectors containing selectable resistance markers (attb-GFP-Neo, and attb-GFP-Puro) digested with the same restriction enzymes, in order to generate mNeonGreen vector containing resistance markers. Ligation products were transformed into JM109 chemically competent cells (see above), and plated on kanamycin-containing LB/Agar plates.

Cloning for transient transfection

A plasmid suitable for transient transfection of a GFP fusion of he short positively charged fragment consisting of the first 131 amino acids of Tet1 was generated. To this end, the fragment was amplified from a template plasmid containing the entire coding sequence of Tet1 (AG Leonhardt pc3164) using forward primer ATGGGCGATCGCATGTCTCGGTCCCGCC and reverse primer AAGCGGCCGCTTACTTTGAAGGTGGTAC. This amplification added AsiSI and NotI restriction enzymes sites on either side of the fragment. The fragment and destination vector pCAG-GFP were then digested with AsiSI and NotI FastDigest enzymes (Fermentas), and ligated using T4 ligase (Thermo).

Cell lines

All transfections were carried out using Lipofectamine 3000, following the manufacturer's protocol, with 1-2 µg of DNA used for transfecting each well of a 6-well plate.

Direct insertion of Halo- or SNAP-tags in J1 mouse ESCs was carried out by co-transfecting cells with pUC57 donor plasmids (described above) along plasmids driving expression of Cas9 from *S. pyogenes* as well as a Tet1 or Tet2 specific gRNA (AG Leonhardt plasmid database pc3834 and pc3835), in cells plated in one well of a standard cell culture 6-well plate. Two days after transfections, cells were plated at clonal density in p150 plates and maintained in puromycin selection media for 5 days. Individual clones were picked from the p150 plate and individually transferred to one well of a 96-well plate. After 3-4 days of growth, the 96-well plate was duplicated into two 96-well plates (one standard and one glass bottom). Clones were screened by fluorescence microscopy of the Nikon Spinning disk system (further described in the microscopy section) with a 40x 0.8 NA air objective, using the JOBS module of NIS Elements software to acquire 4-8 images in each well of the 96-well plate. Positive clones were then grown in larger dishes (6-well plates), and HaloTet was verified for proper localization using higher-magnification fluorescence.

We verified the integration of Halo tags by in-gel fluorescence. Live-cells were labeled with 50 μM Halo-TMR. Cells were then treated with Accutase to detach them, pelleted, washed with PBS, and pelleted again, then stored at -80°C. Pellets were then lysed in 100 μl IP buffer on ice (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM EGTA, 2 mM $MgCl_2$, 0.5% NP-40, 1X protease inhibitors (Serva), 0.1 mM PMSF), and boiled in Laemmli buffer, prior to running on an 8% polyacrylamide gel containing 10% SDS, run at 100 V for ~3 hours, then imaged on an Amersham fluorescence images (GE) using 534 nm/Cy3 and 640 nm/Cy5 filter cubes.

Halo-H2B was randomly inserted in the genome of J1 ESCs using the piggyBAC transposase system by co-transfecting LZ10 PBREBAC-H2BHalo (Addgene plasmid number 91564) along with the PB-transposase, and selecting positively transfected cells with G418.

MIN-tag lines were generated by co-transfecting a plasmid containing a Tet1 cDNA construct fused to mNeonGreen and antibiotic selection cassette flanked with L1 R1 sites, and another plasmid driving the expression of Bxb1 recombinase to insert the cDNA construct and selection marker. These plasmids were transfected into cells which harboured an attb insertion site inserted via CRISPR/Cas9 (Mulholland et al 2015).

Live-cell microscopy sample preparation

Unless otherwise indicated, all live-cell imaging was carried out by plating ESCs one day before imaging on gelatin- or Geltrex-coated 2- or 4-well glass bottom μ-slides (Ibidi), in Serum 2i LIF, N2B27 2i LIF, N2B27 FGF ActA, described above, substituting standard DMEM High glucose, Neurobasal and DMEM-F12 media for their Phenol red free equivalents (Gibco).

2D single-molecule imaging

Both fast and slow single-molecule tracking in 2D was performed on a Nikon Ti-Eclipse equipped with a 100X 1.49 NA TIRF objective and an Andor 888 EMCCD camera, ALC 400 laser combiner, a TIRF illuminator, a CSU-W spinning disk unit, and a FRAPPA photoactivation device and an environmental chamber (Okolab).

For all live-cell imaging, the environmental chamber was always pre-heated to 37°C and maintained at 5% CO2 for at least two hours prior to imaging.

For single-molecule tracking, HiLO illlumination set using the TIRF illuminator, with the emission light bypassing the spinning disk unit. Briefly, the focal plane of the surface of the coverslip was found, and the focus locked using the Perfect Focus System, and the laser was collimated. Then the critical angle was found by adjusting the angle of the laser until total internal reflection lead to extinction, and then adjusted back just until nuclei were visible.

Slow single-molecule tracking was performed with exposure times ranging from 0.45 to 0.8 seconds, acquired at a rate of 0.5 - 2 frames per second.

Fast single-molecule tracking was performed with exposure times of 8, 18 or 30 ms. Due to camera dead time to transfer data (~4 ms), this lead to acquisition rates of 80, 45, or 30 frames per second. Cells sparsely labelled with Halo-JF549 were illuminated with the 561 nm laser with the AOTF set to 80%, corresponding to ~9 mW behind to objective (for 30 ms exposure). Cells sparsely labeled with Halo-JF646 were illuminated with the AOTF set to 100%, 80% or 50% laser power (for 8, 18, and 30 ms exposure times, respectively), measured to be ~ 18, 16 and 10 mW behind the objective.

Images datasets were processed in Fiji, with rolling background subtraction and Super-redundancy denoising (unpublished, Ricardo Henriques group). Regions of interest corresponding to nuclei were selected using reference images of bulk Tet1, Tet2 or H2B bulk labeled with a fluorophore spectrally distinct from the fluorophore used for sparse labeling. Tracking was performed with the Trackmate plugin, limiting detection to nucleus ROIs. Spots were detected by manually setting a diameter of 0.6 µm, roughly twice the size of the full-width half-max of the point spread function. A threshold was manually set for each image keeping clearly visible spots. No filtering was performed on spots. Then, a LAP tracker was used, allowing a maximum search radius of 1 µm and a maximum time gap of 3 frames. Tracks with a minimum of 4 frames were kept for downstream analysis.

Single-molecule trajectory analysis was carried using SpotON software (Hansen et al, 2017), running locally in MATLAB (versions 2016 and 2019). Both two- and three-state kinetic models were fit to the histograms of displacements at different time intervals, considering only the first 4 displacements of each track, in order to avoid bias due to long-duration tracks of bound molecules.

Photobleach step counting was performed by imaging cells fixed in 2% paraformaldehyde-fixed for 10 minutes then washed and maintained in PBS. Spots were detected using Trackmate as described above, but with a search radius of 0.15 μm to keep immobile spots and to account for localization error and drift. Using the xy coordinates of these tracks, intensities at these spots were extracted for each frame of the time-lapse series. To count photobleach steps, vbFRET was used in MATLAB to count number of intensity states.

3D single-molecule imaging

3D single-molecule imaging was performed on a custom built aberration-corrected 3D multifocus microscope at the Janelia Research Campus, on a Nikon Ti-Eclipse stand, equipped with a 100X 1.4 NA objective (Nikon), and a Tokai Hit stage heater. Cells plated in 8-well chambered Lab-tek coverglasses (Nunc) were illuminated with a 561 nm laser set to 50 mW in the collimated space behind the objective.

Images were collected on the full camera chip of a 512x512 pixel EMCCD camera (Andor) at the maximum acquisition rate of 30 frames per second, using NIS elements (Nikon).

Calibration was performed at the start of each acquisition setting, by acquiring z-stacks of fluorescent 0.1 μm Tetraspeck beads (Thermo), and computation of registration settings between focal planes using an affine transformation was performed using a custom written script in MATLAB.

Image stacks were deconvolved using a Richardson-Lucy algorithm, and single-molecule tracking was performed either in 3D on full multifocus stacks, or in 2D on average projections, using the Trackmate plugin in Fiji. Single-molecule trajectories were then fit to a 2- or 3-state kinetic model using Spot-ON software (Hansen et al 2017b), running locally in Matlab (versions 2016a and 2019a).

Spinning Disk stills and time-lapse images

Still images for interphase and mitotic cells were performed on the same Nikon system described above, but with the spinning disk unit in the excitation and emission light path for confocal sectioning. Cells expressing HaloTet1 and mNeonGreen-Tet2, as well as cells expressing HaloTet1, HaloTet Δ1-131, or HaloTet2, were plated in two-well glass bottom coverslips (Ibidi) and labeled with 50 μM HaloTMR, and 100 nM SiR-DNA to label DNA an hour prior to imaging. Samples were excited with 561 and 640 nm lasers (and 488 if mNeonGreen was expressed).

Long-term time-lapse images were obtained by acquiring an image at a single-focal plane every 15 minutes at 10-20 fields of view for 16-24 hours.

Stimulated Emission Depletion microscopy

Super-resolution imaging was on living cells expressing HaloTet1 labeled with Halo-atto594 (a kind gift from Robert Kasper, MPI-Neurobiologie) and SiR-DNA. A STED microscope system (Abberior) equipped with a 100x 1.4 NA oil-immersion objective (Olympus), pulsed 594 and 640 excitation lasers, a 775-nm depletion laser shaped with a spatial light modulator (Hammamatsu), and avalanche photodiode detectors (Excelitas), controlled with IMspector software.

Immunofluorescence and widefield imaging

For hmC stainings, wt J1 cells were co-cultured with cells expressing mNeonGreen-Tet1 and HaloTet2. Immunostainings were performed on cells plated the night before on geltrex-coated 1.5 thickness (170 ± 5 µm) High Precision glass borosilicate coverslips (Carl Roth Gmbh) in a 6-well cell culture plate. Cells were washed twice in PBS, then fixed in 2% paraformaldehyde at room temperature for 10 minutes. After step-wise washing in PBS, cells were washed in PBS with 0.02% Tween-20 (Carl Roth, "PBST"), then permeabilized with PBS containing 0.05% Triton X detergent (Sigma-Aldrich). Cells were washed, then blocked for one hour in Block buffer (PBS with 2% Bovine Serum Albumin). Then cells were incubated for one to two hours with primary antibody diluted in Block buffer at room temperature. The primary antibody was then washed out with PBST, and then cells were incubated with the secondary antibody for one to two hours at room temperature. Cells were then washed with PBST, and incubated with 1 µg/ml of DAPI (Sigma-Aldrich) in PBST for 7 minutes. Cells were washed, then coverslips were mounted onto glass slides by quickly dipping in water to remove excess salts and placing cell-side down on a 5 µL drop of Vectashield mounting media (Vector Labs) on a glass slide . After blotting excess water and mounting media, coverslips were sealed with nail polish.

Cells stained for hmC were imaged on a Delta Vision elite equipped with an LED illumination system (GE Healthcare Life Sciences, formerly Applied Precision), a CoolSnap HQ2 camera (Photometrics), mounted on an inverted IX-71 stand with a 60x 1.42 NA Plan-Apo objective (Olympus), operated with SoftWorX acquisition software.

Raster image correlation spectroscopy

Raster Image Correlation Spectroscopy (RICS) was performed in the AK Lamb group on a custom-built system as described in Hendrix et al, 2015. The system consists laser scanning confocal microscopy setup equipped with a pulsed interleaved excitation (PIE) system consisting of pulsed lasers at 470, 640 (Picoquant), and/or 561 nm (Toptica Photonics) wavelengths. These lasers are synchronized to a master clock, and then delayed ~20 ns relative to one another to achieve PIE. Two galvo-scanning mirrors steer the excitation light, filling the back aperture of a 100x 1.45 NA oil-immersion objective (Nikon) with a correction collar set for 37°C. Emitted photons are descanned and spectrally separated from the excitation light with a Di01-R405/488/561/635 polychroic mirror (Semrock). Then, emitted photons from eGFP/ mNeonGreen, mScarlet/JF549 fluorescence emission were separated by a 565 DCXR dichroic mirror (AHF Analysentechnik) and collected on avalanche photodiodes, a Count Blue (Laser Components) and a SPCM-AQR-14 (Perkin-Elmer) with 520/40 and a 630/75 emission filters (Chroma, AHF Analysentechnik). Time-correlated single-photon counting units were used to record photon arrival times on the APDs.

At the start of each measurement session, the alignment of the system was verified and the dimensions of the excitatory point spread function were determined by performing FCS with PIE of a mixture of Atto-488 and Atto565 dyes, excited with pulsed 470, 561 and 640 nm lasers focused roughly 1 µm above the surface of the coverslip, with power set to 10 µW. The laser intensity was measured at a fixed area in the collimated space behind the galvo-scanning mirror system. FCS data was processed using PAM and FCSfit, which are part of the freely Pulsed Interleaved Excitation Analysis with Matlab software package (Schrimpf et al, 2018).

For measurements of cells expressing UHRF1-GFP and DPPA3-mScarlet (section 2.3.3) cells were plated on Ibidi two-well glass bottom slides the day before measuring, and 1 µg/ml doxycycline was added overnight to induce expression of mScarlet constructs. For all measurements in cells, a stage top incubator maintained cells at 37˚C. Unidirectional raster scanning was performed on a total field of view of 12 x 12 µm, acquiring photon information from 300 pixels x 300 lines, a pixel spacing size of 40 nm. The pixel dwell time was 11 µs, thereby leading to a line time of 3.33 ms. To help mitigate photobleaching, the acquisition rate was of one frame per second, and 100-200 frames were collected to obtain sufficient information for fitting. Pulsed 470 and 561 nm lasers were adjusted to 4 and 5 µW respectively. No line- or frame-averaging was used for acquisition.

Processing and quantitative analysis of scanned images was performed entirely in the freely available Pulsed Interleaved Excitation Analysis with Matlab (PAM) software package (Schrimpf et al, 2018). First, time gating of the raw photon stream was performed by selecting only photons emitted on the appropriate detector after excitation pulse in PAM, thereby allowing to generate cross-talk free images for each channel. Microtime Image Analysis (MIA) was then used to remove slow fluctuations by subtracting a moving average of 3 frames. Finally, to only perform the analysis on nuclear proteins diffusing in proximity to chromatin, a region of interest corresponding to the nucleus was selected, excluding nucleoli and dense aggregates. The spatial autocorrelation and cross-correlation functions (SACF and SCCF) were calculated on these selected ROIs as done previously using arbitrary region RICS (Hendrix et al, 2016):

$$G(\xi,\psi) = \frac{\langle I_{RICS,1}(x,y) I_{RICS,2}(x+\xi, y+\psi)\rangle_{XY}}{\langle I_{RICS,1}\rangle_{XY}\langle I_{RICS,2}\rangle_{XY}}$$

where I_{RICS} is the intensity in photons at a given coordinate in pixels, ξ and ψ are the shifts along the x- and y-axis scan directions, respectively, in pixel units, and brackets denote the average over the whole image. Then, the correlation function was fitted to a two-component model diffusive model, consisting of one mobile fraction and one immobile fraction, in MIAfit:

$$G_{fit}(\xi,\psi) = A_{mob} G_{fit,\,mob}(\xi,\psi) + A_{imm}\ \exp\left(-\delta r^2 \omega_{imm}{}^{-2}\left(\xi^2+\psi^2\right)\right) + y_0$$

where

$$G_{fit,\,mob}(\xi,\psi) = \left(1+\frac{4D(\tau_p\xi+\tau_l\psi)}{\omega^2{}_r}\right)^{-1} \cdot \left(1+\frac{4D(\tau_p\xi+\tau_l\psi)}{\omega^2{}_z}\right)^{-1/2} \cdot \exp\left(-\frac{\delta r^2(\xi^2+\psi^2)}{\omega^2{}_r + 4D(\tau_p\xi+\tau_l\psi)}\right)$$

where ω_r and ω_z are the radial and axial widths of the excitatory volume and τ_p and τ_l are the time between pixels and the time between lines, respectively. This fitting then yields parameters such as the diffusion coefficient (D) and the amplitudes of the mobile and immobile fractions (A_{mob} and A_{imm}). The average number of mobile molecules per excitation volume on the RICS timescale was determined by:

$$N_{mob} = \frac{\gamma}{A_{mob}} \cdot \frac{2\Delta F}{2\Delta F + 1}$$

where γ is a correction factor which is used with the 3D Gaussian shape of the excitatory volume, and $2\Delta F/(2\Delta F+1)$ is used to correct for apparent intensity when images are processed with a moving average subtraction prior to calculating the SACF. The bound fraction is the fraction of particles which remain visible during the acquisition of 5-10 lines of the raster scan, corresponding to ~20-40 ms, depending on the excitatory volume. The cross-correlation model was fitted with a cross-correlation function, and the extent of cross-correlation was calculated based on the amplitude of the mobile fraction of the cross-correlation fit divided by the amplitude of the mobile fraction of the autocorrelation fit of DPPA3-mScarlet, as used previously (Hendrix et al, 2015).

Fluorescence recovery after photobleaching

FRAP on Halo-labeled Tet cell lines, as well as DPPA3 cell lines were performed on the Nikon system describe above using the FRAPPA photoactivation device. A 4x4 pixel box was selected as a bleach ROI. Prior to bleach, 8-10 images were acquired, then the selected ROI was bleached with dwell-times of 500 - 1000 us per pixel, using the same laser as that used for acquisition (488 for mNeonGreen fusions, 561 for mScarlet fusions and Halo fusions labeled with TMR). Then a recovery image time-lapse was acquired at 2-4 frames per second, for 1-2 minutes.

Analysis was performed in Fiji, where the bleached area, nuclear area, and background were found automatically, and FRAP recovery was calculated using double-normalization (Phair et al, 2004; Miura, 2005) to account for acquisition photobleaching. Since the bleached spot was maintained at the same coordinates in the center of the image, an automated analysis script to extract intensities at all timepoints was written in Fiji to generate an ROI for the nucleus (using Otsu thresholding on an average projection), and using an ROI for the bleached area at the same coordinates for an entire dataset. Intensities were extracted for these ROIs at all timepoints. A global background intensity was measured for each condition by getting the average intensity in ROIs without cells. ROIs were then visually inspected, and cells were discarded if nuclear ROIs did not properly find the nucleus, or in cells where drift would clearly lead to erroneous measurements.

FRAP on mNeonGreen-labeled Tet1 cDNA constructs inserted via the MIN tag (section 2.2.1) was carried out on a PerkinElmer Ultraview Vox system set on Axiovert stand with a 63x 1.4 NA oil-immersion objective (Zeiss), a spinning disk confocal head (Yokogawa) and an EMCCD camera (Hammamatsu), controlled by Volocity software. Three pre-bleach images were acquired, and spot ROIs were set on 4-5 cells per field of view. Bleaching was carried out on those spots for 50 ms per spot using the 488 nm laser, and post-bleach recovery images were acquired at a rate of one image every two seconds for two minutes.

For these, an automated script was written in Fiji to extract intensity data. First, ROIs of nuclei were found with Otsu thresholding on an average projection of prebleach images. Bleached nuclei were selected using binary images mapping bleach locations automatically generated by Volocity acquisition software. Since the actual bleached spot differed from the recorded target by ~1-2 µm, the actual bleached spot was found by selecting the maximum intensity pixel in each ROI of bleached nuclei, from an image consisting of a subtraction of the last pre-bleach image and the first post-bleach image. Automated ROI selection were then visually inspected, and filtered as described earlier.

Data processing, analysis, and visualization software

Image data was processed and extracted using Fiji (Schindelin et al 2012), along with plugins Trackmate (Tinevez et al 2016), ThunderSTORM (Ovezny et al 2014), NanoJ (Laine et al 2019) and FFT/ICS plugins by Jay Unruh at Stowers Institute for Medical Research in Kansas City, MO.

MATLAB was used for data analysis. Specifically, SpotON (Hansen et al, 2017) was used to analyze single-molecule trajectory data, Pulsed interleaved Analysis in MATLAB (PAM) was used for RICS data (Schrimpf et al, 2018). A custom-written script was used for reconstruction of 3D acMFM images (written by Jesse Aaron, Janelia Research Campus). Amira and Imaris were used for rendering 3D acMFM images.

R-studio was used as a scripting environment for R, for data processing and plotting with the tidyr, dplyr, ggplot2, viridis, and ggsci packages.

Python was used with numpy, scipy, pandas and matplotlib, to calculate mean squared displacement and plot 3D trajectories.

Dye-conjugated ligands

An overview of fluorescent proteins and dyes used in this work is given in Figure 2.1.

For single-molecule tracking with HaloJF549 (for imaging with a 560 nm excitation laser), cells were incubated with 1-2 pM of Halo-JF549 (for sparse labeling) and 5-10 nM Halo-JF646 (for bulk labeling) for 30 minutes at 37C, then washed twice with PBS, incubated for 5-10 minutes in regular media at 37, then washed again with PBS. For single-molecule tracking with Halo-JF646 (with a 640 nm excitation laser), cells were incubated with 50-100 pM Halo-JF646 (for sparse labeling) and 50 nM of Halo-TMR (for bulk labeling) for 30 minutes at 37C, and washed as indicated earlier.

For RICS experiments, Halo-JF549 or Halo-JF646 were used at bulk labeling concentrations. After washing, cells were allowed to incubate for 1-2 hours, and washed again, since RICS measurements are exquisitely sensitive to freely diffusing dyes.

For STED imaging, cells were incubated with 100 ng/ml Halo-atto594 and 100 nM SiR-DNA for one hour at 37 and washed as indicated earlier. Alternatively, 5 nM of Halo-JF646 was successfully used on the STED to label Tet1 and Tet2.

For FRAP, still images, time-lapse images, in-gel fluorescence, cells were incubated in 50-100 nM Halo-TMR for one hour, and washed as indicated above. For DNA staining, cells were incubated with 100 nM SiR-DNA for an hour, and washed once prior to imaging. Finally, SNAP-TMR and SNAP-SiR were used at 1:100 of the manufacturer's recommended concentration to verify insertion of SNAP-tag.

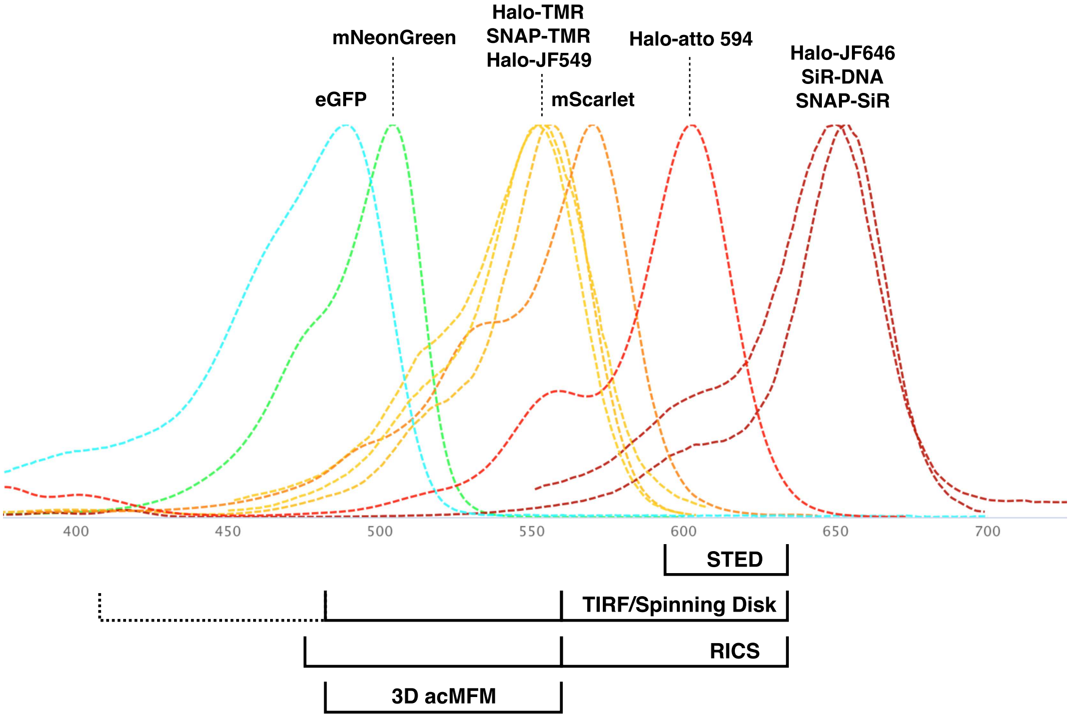


Figure 2.1 Overview of fluorescent proteins and dye-conjugated ligands used in this work, showcasing their excitation spectra (top), along with microscopes laser lines available of the imaging systems used to measure the spatiotemporal activities of proteins (bottom). Spectra generated with FPViewer.

Table 4.1 Cell lines used in this study

Number	Cell line name	Background cell line	Plasmids
ccJR17.001	HaloTet1 mNeonGreenTet2	mNeonGreen-Tet2 J1 ESCs	pUC57-Tet1-Halo-Tet1; pSpCas9-gRNA-Tet1
ccJR17.002	HaloTet2 mNeonGreenTet1	mNeonGreen-Tet1 J1 ESCs	pUC57-Tet1-Halo-Tet2; pSpCas9-gRNA-Tet2
ccJR17.003	SNAP-Tet1 mNeonGreenTet2	mNeonGreen-Tet2 J1 ESCs	pUC57-Tet1-SNAP-Tet1; pSpCas9-gRNA-Tet1
ccJR18.001	HaloTet1 *DTKO*	DTKO J1 ESCs	pUC57-Tet1-Halo-Tet1; pSpCas9-gRNA-Tet1
ccJR18.002	HaloTet2 *DTKO*	DTKO J1 ESCs	pUC57-Tet1-Halo-Tet2; pSpCas9-gRNA-Tet2
ccJR18.003	HaloTet1 *Tet1CM*	Tet1 catalytic mutant J1 ESCs	pUC57-Tet1-Halo-Tet1; pSpCas9-gRNA-Tet1
ccJR18.004	HaloTet2 *Tet2CM*	Tet2 catalytic mutant J1 ESCs	pUC57-Tet1-Halo-Tet2; pSpCas9-gRNA-Tet2
ccJR18.005	HaloH2B	WT J1 ESCs	PREBAC-H2B-Halo; piggyBac transposase
ccJR19.001	HaloTet1	WT J1 ESCs	pUC57-Tet1-Halo-Tet1; pSpCas9-gRNA-Tet1
ccJR19.002	HaloTet2	WT J1 ESCs	pUC57-Tet1-Halo-Tet2; pSpCas9-gRNA-Tet2
ccJR19.003	HaloTet1Δ1-131	WT J1 ESCs	pUC57-Tet1-Halo-Tet1Δ1-131; pSpCas9-gRNA-Tet1
unavailable	UHRF1-GFP, DPPA3KO, dox:mScarlet-DPPA3 wt	UHRF1-GFP DPPA3-KO J1 ESCs	Generated by Christopher Mulholland
unavailable	UHRF1-GFP, DPPA3KO, dox:mScarlet-DPPA3 KRR	UHRF1-GFP DPPA3-KO J1 ESCs	Generated by Christopher Mulholland
unavailable	UHRF1-KO (eGFP), DPPA3KO, dox:mScarlet-DPPA3 wt	UHRF1-KO-eGFP DPPA3 KO J1 ESCs	Generated by Christopher Mulholland
unavailable	UHRF1-KO (eGFP), DPPA3KO, dox:mScarlet	UHRF1-KO-eGFP DPPA3 KO J1 ESCs	Generated by Christopher Mulholland
unavailable	UHRF1-GFP, DPPA3KO, dox:mScarlet	UHRF1-GFP DPPA3-KO J1 ESCs	Generated by Christopher Mulholland

Table 4.1 (continued)

Number	Cell line name	Background cell line	Plasmids
ccJR16.001	mNeonGreen-Tet1-FL (cDNA)	Tet1 attp/attp J1 ESCs	attb-mNeonGreen-Tet1 Puro, pCAG-NLS-HA-Bxb
ccJR16.002	mNeonGreen-Tet1-CD (cDNA)	Tet1 attp/attp J1 ESCs	attb-mNeonGreen-Tet1-CD Puro, pCAG-NLS-HA-Bxb
ccJR16.003	mNeonGreen-Tet1-Δ1-833 (cDNA)	Tet1 attp/attp J1 ESCs	attb-mNeonGreen-Tet1-Δ1-833 Puro, pCAG-NLS-HA-Bxb
ccJR16.004	mNeonGreen-Tet1-Δ834-1363 (cDNA)	Tet1 attp/attp J1 ESCs	attb-mNeonGreen-Tet1-Δ834-1363 Puro, pCAG-NLS-HA-Bxb
unavailable	mNeonGreen-Tet1-N-term (cDNA)	Tet1 attp/attp J1 ESCs	Generated by Paul Stolz and Burak Ozan
unavailable	mNeonGreen-Tet1-CXXCmut (cDNA)	Tet1 attp/attp J1 ESCs	Generated by Paul Stolz and Burak Ozan

Results

3.1 Dynamics of Tet1 and Tet2 in live embryonic stem cells

3.1.1 Cas9-mediated genome-editing to label Tet proteins

The cell lines generated and discussed in this section were generated with the help of Christopher Mulholland and Daniel Nixdorf. Halo-JF549 and Halo-JF646 were kind gifts from Dr Luke Lavis (Janelia Research Campus) and Halo-atto 594 was a kind gift from Dr. Robert Kasper (MPI Neurobiologie).

In order to measure to measure the binding and diffusive behaviour of Tet proteins in living cells, we first needed to generate cell lines which express fluorescently labeled TETs at endogenous levels. To this end, we used CRISPR/Cas9 mediated genome engineering to insert the coding sequences of fluorescent tags in frame after the endogenous ATG start codons of Tet1 and Tet2. Cell lines expressing Tet1 or Tet2 endogenously labeled with mNeonGreen (Shaner et al, 2013) were designed, generated and validated by Chris Mulholland and Daniel Nixdorf (Mastersarbeit 2017). Briefly, a plasmid containing spCas9 and a TET1- or TET2-specific gRNA was co-transfected this with a plasmid containing a donor sequence consisting of the coding sequence of the tag, flanked with 1 kb of homology upstream and downstream of the Cas9 target, which in this case was the TET1 or TET2 ATG start codon (FIGURE 3.1 A). The insertion of mNeonGreen was verified by screening for fluorescence, and by PCR and Sanger sequencing.

By simply swapping out the coding sequence of the fluorescent tag, we generated plasmids cells with TET1 or TET2 labeled with the HaloTag, or SNAP tags. Both HaloTag and SNAP tags are non-fluorescent. However, they can be labeled by incubating cells for ~30 minutes with cell-permeable organic dye-conjugated ligands which specifically and covalently bind to their respective tags (FIGURE 3.1 B). Halo- and SNAP fluorescent ligands used in this study include Halo-TMR, Halo-JF549, Halo-JF646, SNAP-TMR, and SNAP-SiR.

Thus, we used the previously generated cell lines (mNeonGreen-Tet1 or mNeonGreen-Tet2) to generate dual-labeled cell lines, where the Tet2 or Tet1 was labeled with the Halo, thereby generating mNeonGreen-Tet1 / HaloTet2 cells, and HaloTet1 / mNeonGreen-Tet2 cells (FIGURE 3.1 C). We screened for positive clones by fluorescence microscopy, and selected clones where fluorescence was clearly visible in the nucleus, where TET1 proteins are expected to be found.

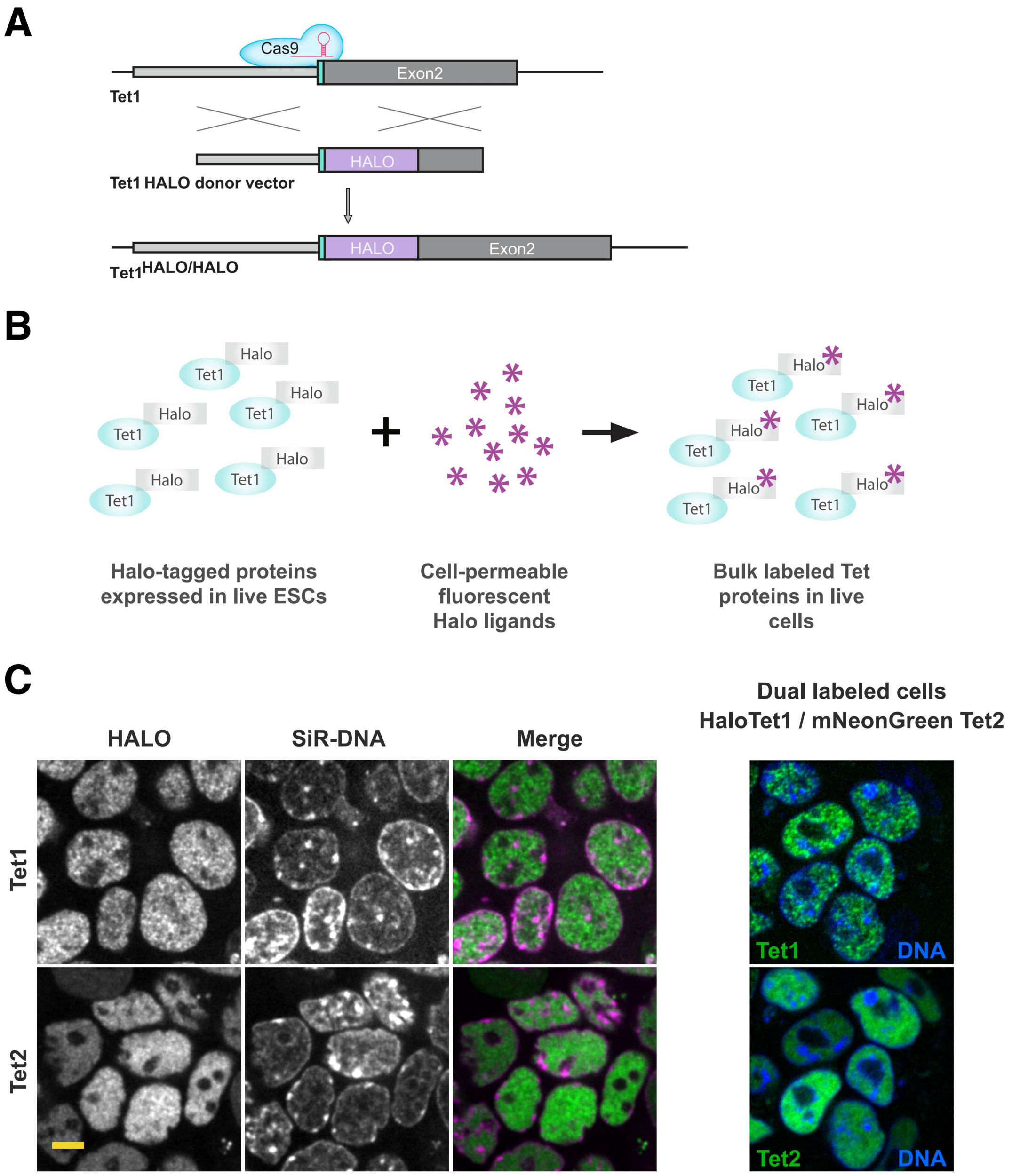


Figure 3.1 Generation of stable ESC lines expressing the HaloTag inserted via CRISPR-Cas9 in frame at the endogenous ATG start site of Tet1 or Tet2. A. General targeting strategy using a homologous donor harbouring the Halo-tag coding sequence in frame with Tet1 (or Tet2). B, labeling method for Halo-tagged constructs. C, Example images of cells expressing Tet1 (upper panels) and Tet2 (lower panels) endogenously labeled with the Halo-tag, and counterstained for DNA (with SiR-DNA). Right, dual-labeled cells were sequentially generated to insert the Halo-tag at the Tet1 start site in cells which already had mNeonGreen inserted at the Tet2 start site. Scale bar = 5 μm

We verified that Tet1 and Tet2 were labeled with the HaloTag by labeling cells with Halo-TMR, and running their lysate on a poly-acrylamide gel, which yielded fluorescent bands at their expected sizes ~250 kDa (FIGURE 3.2 A). We then sought to verify that the labeled TETs were still catalytically functional in genome-edited cell lines. To this end, we performed immunofluorescence to stain hmC, the catalytic product of TET enzymes, in cells where both Tet1 and Tet2 were endogenously labelled with either HaloTag or mNeonGreen. We found that hmC levels were visibly maintained in the nucleus (FIGURE 3.2 B), indicating that TET enzymes were still active in cells expressing HaloTet1/mNeonGreen-Tet2. We quantified the fluorescence intensity of the staining, and surprisingly found that hmC levels were slightly higher in labeled cells compared to WT (FIGURE 3.2 C). We thus conclude that our N-terminally labeled TET proteins are enzymatically functional in naive ESCs.

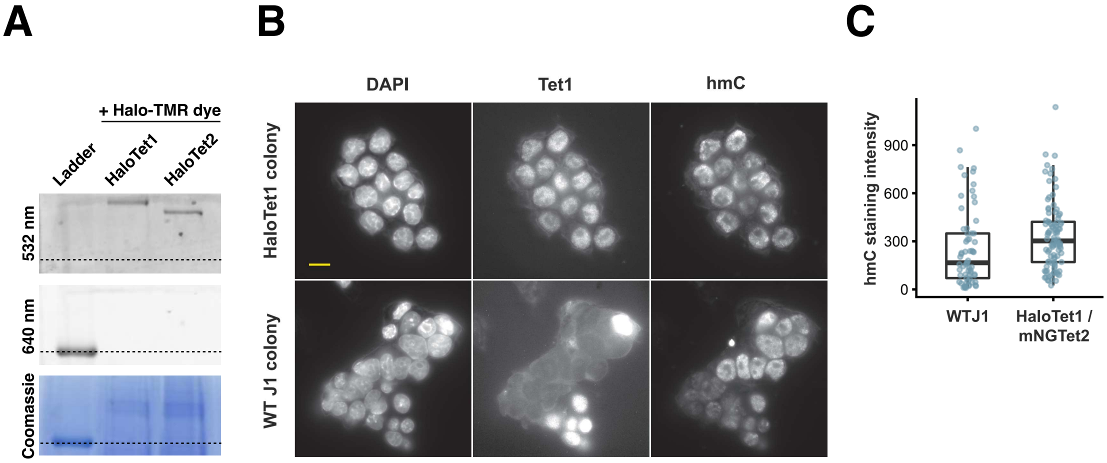


Figure 3.2 A, in-gel fluorescence assay of Halo-TMR labeled HaloTet1 and HaloTet2 cell lysates, excited with 532 nm (top panel) and 640 nm light (middle panel). Trans-illuminated coomassie-stained gel (bottom panel) as a sample loading control. Dotted line indicates 170 kDa band B, example images of wild-type ESCs co-cultured with HaloTet1 / mNeonGreen-Tet2 cells, and stained for hmC. C. Quantification of hmC in wt and HaloTet1 / mNeonGreen-Tet2 cells, which were sorted based on fluorescence intensity of Tet1. Scale bar = 10 µm

We then asked whether the insertion of the tag affected the temporal expression patterns of Tet proteins. It has been shown previously by RNA-Seq and by proteomic analysis that Tet2 is largely downregulated at the Epiblast-like state, whereas Tet1 undergoes a slight down regulation (Fidalgo et al, 2016; Mulholland et al, 2018). We tested whether our genome-edited cells were still capable of these temporal expression patterns by measuring the intensity of fluorescently labeled Tet1 and Tet2 as a readout of expression levels before and after the transition from a naive pluripotent state to a primed Epiblast-like state. Consistent with published data derived from other methods, we found that Tet1 is indeed somewhat downregulated at the epiblast state, and Tet2 levels are reduced to near background levels, indicative of a massive downregulation at the protein level (FIGURE 3.3). Together, these data suggest that the temporal regulation of expression levels of Tet1 and Tet2 is mostly unaffected by the genomic insertion of tags.

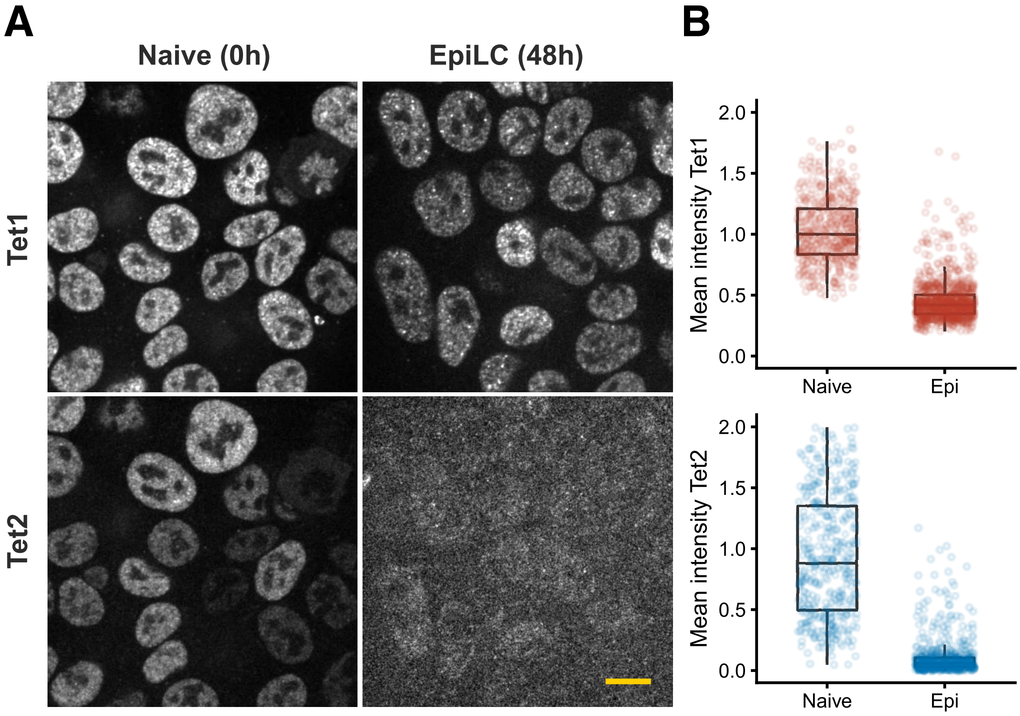


Figure 3.3 Temporal expression patterns are maintained in Halo-labeled ESCs. A. Example images of cells expressing Halo-labeled Tet1 (top) and Tet2 (bottom) maintained in naive stem cell media (right panels) or pushed to transition to the primed pluripotency epiblast state (right panels). B, Mean intensity of Tet1 (top) and Tet2 (bottom) in stem cell nuclei (each data point corresponds to a cell), normalized to the mean intensity at the naive state. Scale bar = 10 μm

Taken together, these results suggest that we have successfully generated cell lines wherein Tet proteins are fluorescently labeled at their endogenous genomic loci, that these labeled proteins are enzymatically functional, and that the insertion of these tags do not visibly affect their expression patterns in ESCs.

With these characterized cell lines, we investigated the subnuclear distribution of TET1 and TET2 at the single-cell level in the nucleus. Fluorescently labeled TET proteins were imaged with DAPI or SiR-DNA to label the DNA. Both Tet1 and Tet2 appeared to predominantly localize to areas of low DNA compaction, typically corresponding to euchromatin, while being partially depleted from areas of dense DNA compaction generally corresponding to heterochromatin (chromocenters, pericentric and peripheral heterochromatin) (FIGURE 3.1 C). These localization patterns are consistent with ChIP-Seq profiles, which indicate that Tet proteins associate with euchromatin loci, such as promoters and enhancers of actively transcribed genes. Thus, we show in live-cells and at the single-cell level, a spatial distribution TET proteins predominantly associated with euchromatin. We verified the localization of TET1 at higher resolution, by performing STED on living cells expressing HaloTet1 (FIGURE 3.4). At these higher resolutions, again, Tet1 was mostly localized to areas of low DNA compaction.

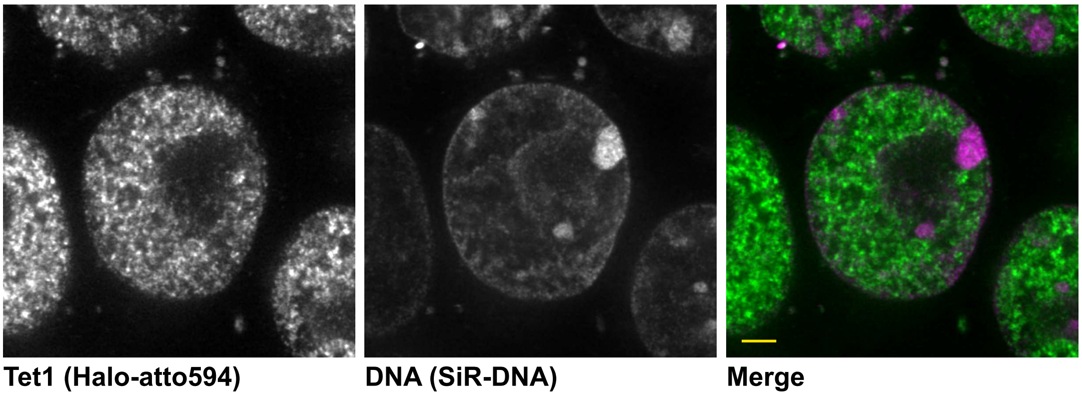


Figure 3.4 Super-resolution microscopy of Tet1 in living cells. Example images of cells expressing Halo-labeled Tet1 counterstained with SiR-DNA, imaged with a super-resolution STED microscope. Scale bar = 2 µm

In living cells, the spatial distribution of endogenously labeled Tet1 and Tet2 were visibly different. Tet1 appeared to localize in small clusters, whereas the distribution of Tet2 appeared more uniform and diffuse (FIGURE 3.5 A). As a measure of this parameter, we looked at the variance in pixel intensity within nuclei, with the notion that evenly distributed protein would have a smaller variance than proteins which have a clustered distribution, as used elsewhere (Maddox et al, 2006). We calculated the coefficient of variation in still images of nuclei of live cells. We took the standard deviation of fluorescence intensity in all pixels of the nucleus, and normalized it to the mean intensity of the nucleus (Gassler et al, 2017), to account for an increase in noise with mean intensity (FIGURE 3.5 B). We found that while Tet1 and Tet2 fluorescence intensities are in a similar range, the distribution of Tet1 has a larger coefficient of variation than Tet2, suggesting an uneven spatial distribution (for example, dense clusters interspersed with sparse areas) (FIGURE 3.5 C). In contrast, Tet2 showed a more even distribution, suggesting that it is in a more diffusive state, or less likely to form clusters. Thus, we show that in living cells, Tet1 and Tet2 differ in their spatial distribution, which hinted at underlying differences in their binding and diffusive properties in living ESCs. Motivated by these results, we next set out use these cells expressing endogenously labeled Tet proteins to measure their binding and diffusive properties at the single-cell level living cells at different timescales, in order to establish a model TET protein mobility.

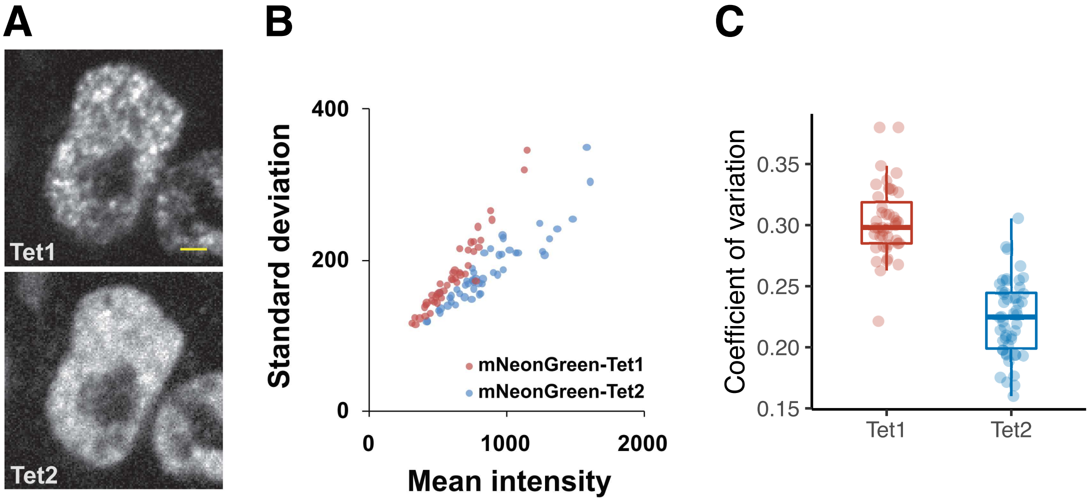


Figure 3.5 Contrasting spatial distribution of Tet1 and Tet2 in live ESCs. A. Example images of cells expressing HaloTet1 and mNeonGreen-Tet2. B. Scatter plot of the standard deviation of fluorescence intensity over the mean intensity in cells expressing either mNeonGreen-Tet1 or mNeonGreen-Tet2. Thick grey line indicates data obtained from diffuse fluorescent signals as a calibration for the non-linear dependence of camera noise on mean intensity. C. detector normalized standard deviation as a readout of intensity variance for cells expressing mNeonGreen Tet1 or mNeonGreen-Tet2. Scale bar = 2 μm

3.1.2 Tet1 and Tet2 show contrasting localization during mitosis

The maintenance of cell identity depends on the reliable and timely binding of nuclear proteins to specific genomic loci, in order to maintain or modulate gene expression programs. In most measured nuclear proteins, the binding and diffusive properties are such that proteins diffuse throughout the nucleus at rates of 2-5 μm^2/ s, and bind to both target and non-specific loci for in the range of 1-30 seconds. In parallel, the local nuclear architecture can also influence how likely a protein is to bind to a certain locus. In this regard, the global likelihood of a nuclear protein finding its target will depend on its binding and diffusive properties, its concentration and the availability of its binding sites.

In proliferating cells, large scale structural changes to the nucleus and to chromosomes during mitosis can greatly disrupt the steady-state binding rates of nuclear proteins, and thus present a challenge to maintenance of cell identity. Most visibly, chromosome condensation results in a large-scale reorganization and reshaping of chromatin, largely driven and maintained by condensin (Gibcus et al, 2019). Curiously, ATAC-seq data shows that the accessibility of chromatin loci is generally maintained (Teves et al, 2016). Thus, even if targets are still mostly accessible, their spatial distribution is largely perturbed.

Importantly, the effective concentration of diffusive nuclear proteins decreases dramatically during cell division. Shortly after the beginning of chromosome condensation, during prophase, the nuclear envelope breaks down, which allows unbound nuclear proteins to diffuse throughout the whole volume of the cell, thereby decreasing their probability of binding to genomic loci.

Recently, live-cell imaging of transcription factors in proliferating cells has showed that several pluripotency factors remain enriched on mitotic chromosomes throughout cell division (Teves et al, 2016). In some cases, this is viewed as "mitotic bookmarking", a means to label a genomic locus during cell division, in order to rapidly influence transcription at that locus in early G1. In other cases, nuclear protein enrichment on mitotic chromosomes is not particularly associated with a "bookmarking" function, but rather seen as a result of the electrostatic interactions of the nuclear protein with DNA. In essence, proteins could bind transiently with mitotic chromosomes, leading to visible enrichment, given a high enough on-rate.

Since Tet1 and Tet2 contribute to methylation homeostasis, and thus cell identity, we asked whether they could bind to mitotic chromosomes, despite their contrasting distribution in the nucleus. To this end, we used cells expressing halo-labeled Tet1 or Tet2, and incubated them with Halo-TMR for bulk labeling and SiR-DNA, to visualize DNA. As a first test, we identified cells in metaphase using SiR-DNA to see the aligned chromosomes, and acquired still images of HaloTet1 or HaloTet2 and SiR-DNA in live cells, on a spinning disk confocal. Strikingly, HaloTet1 appeared consistently enriched on mitotic chromosomes with a lesser signal in the cytosol, whereas HaloTet2 was consistently depleted from mitotic chromosomes, with a stronger signal in the cytosol (FIGURE 3.6 A).

We measured the enrichment of fluorescence intensity on the chromosomes relative to the cytosol, by tracing an ROI around the chromosomes in the SiR-DNA channel, and around the cell in the Halo-TMR channel (Teves et al, 2016, Teves et al, 2018). In keeping with the visually striking results, Tet1 was consistently enriched on mitotic chromosomes, whereas Tet2 was consistently depleted (FIGURE 3.6 B). To get a

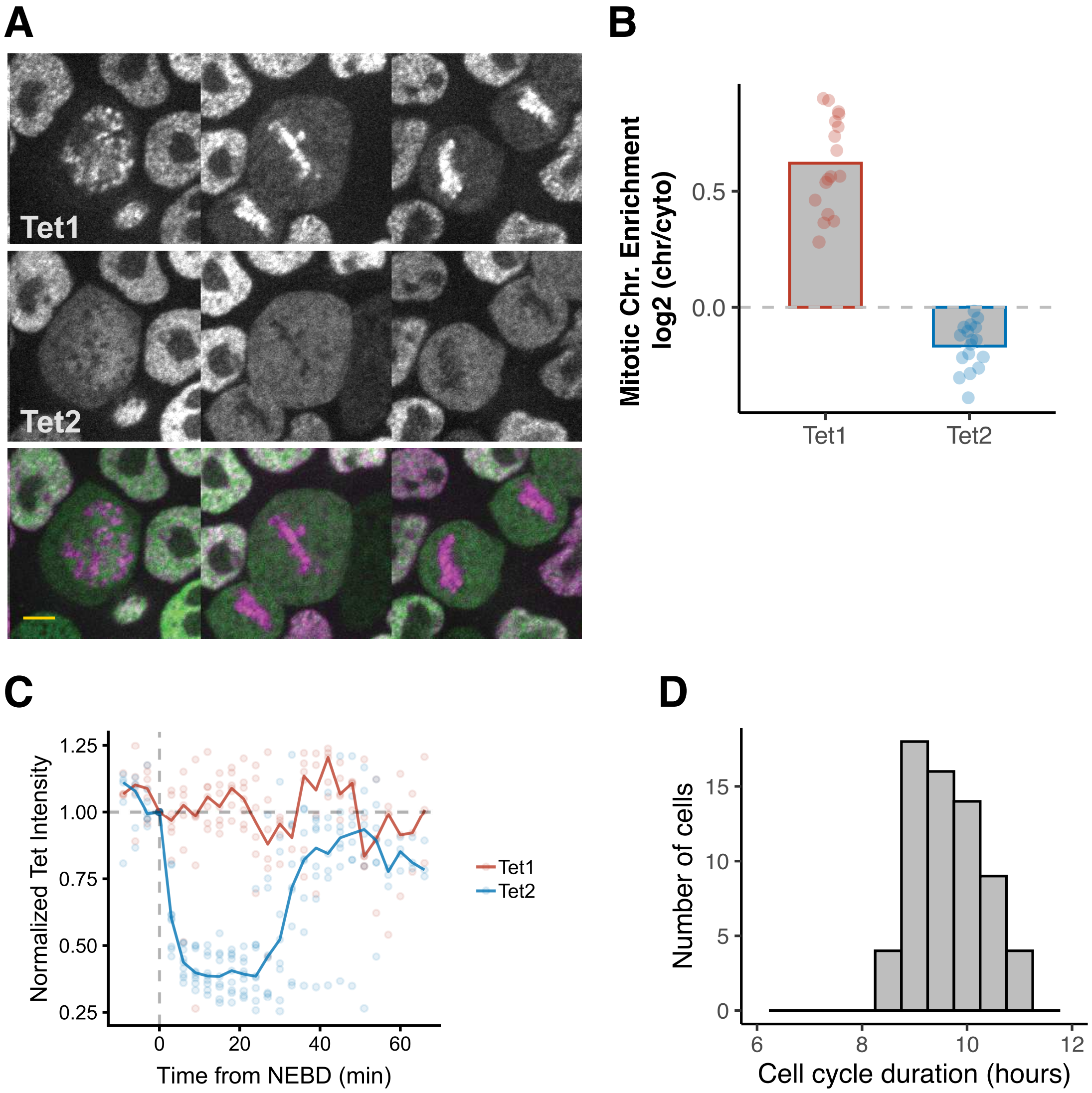


Figure 3.6 Mitotic chromosome localization of Tet1 and Tet2 in ESCs.. A. Example images of live cells expressing HaloTet1 and mNeonGreen Tet2 at different phases of mitosis. Scale bar = 5 μm. B, Chromosome to cytosol ratio of fluorescence signal for Tet1 and Tet2. C. Normalized fluorescence intensity of Tet1 (red) and Tet2 (blue) at chromosomes during mitosis, obtained from time-lapse image series, aligned at the first frame to show signs of nuclear envelope breakdown (NEBD). D, cell cycle duration of ESCs, measured as the duration from anaphase to anaphase from time-lapse images of ESCs labeled with SiR-DNA. Scale bar = 5 μm

better temporal view of these dramatic changes in chromosome binding by Tet proteins, we performed time-lapse imaging of SiR-DNA along with HaloTet1 or HaloTet2, with an image taken in both channels every 2 minutes. We observed that concomitantly with nuclear envelope breakdown, a large fraction of Tet2 rapidly diffuses freely throughout the volume of the cell, leaving a void of fluorescence in the volume occupied by chromosomes throughout the duration of mitosis. In contrast, a small fraction of Tet1 appears to diffuse into the cytosol following nuclear envelope breakdown, while a large fraction remains bound to mitotic chromosomes throughout the duration of mitosis.

We then measured the fluorescence intensity of Tet in chromosome ROIs throughout the time-lapse, and aligned temporal intensity profiles to the first timepoint where nuclear envelope breakdown was visible. We observed that while Tet1 levels drop modestly with nuclear envelope break down, Tet2 drop substantially, and remain low throughout the duration of prometaphase, metaphase and anaphase. Tet2 is then replenished in the nucleus after cell division, concurrent with chromosome decondensation (FIGURE 3.6 C). Intensity profiles show that this replenishment of Tet2 in the nucleus after cell division appears to occur at a slower rate than that of its depletion after nuclear envelope breakdown. This observation suggests that after cell division, Tet2 is transported into the nucleus by an active import mechanism.

To get an idea of the relative proportion of time during which Tet2 is depleted from chromatin, we measured the average duration of mitosis in ESCs, from nuclear envelope breakdown until anaphase, to be roughly 30 minutes. In parallel, we performed long interval time-lapse imaging, with an image acquired every 10 minutes, to measure the duration of the cell cycle. We measured the cell cycle to be roughly 10 hours by manually tracking cells from anaphase to anaphase (FIGURE 3.6 D). Thus, Tet2's depletion from chromatin spans roughly 5% of the cell cycle.

Taken together, these data show that Tet1 has an intrinsic capacity to remain enriched on mitotic chromosomes after nuclear envelope breakdown, suggesting a higher on-rate or lower off-rate of Tet1. Tet2, on the other hand is much more sensitive to nuclear topology and concentration, thus its activity is more dependent on concentration and on nuclear architecture. These results are indicative of Tet2 having a lower on-rate and a higher off-rate. Additionally, Tet2 would spend a greater fraction of time off chromatin, and in the cytosol, potentially enhancing a different subset of interactions with cytosolic proteins. These data are consistent with a large scale study, which showed that mitotic chromosome binding of transcription factors globally correlates with on-rate, but not with residence time (Raccaud et al 2019).

In terms of biological processes, Tet1's greater retention on mitotic chromosomes may reflect the more targeted aspect of its activity, favouring the maintenance of its activity through cell division. Tet2, on the other hand, is mostly depleted from chromosomes, and thus needs to be imported into the nucleus following cell division.

3.1.3 FRAP of Tet proteins

To visualize how the diffusive and binding behaviours come together in the context of the entire population of Tet proteins, at the scale of the nucleus, over the tens-of-seconds timescale, we performed FRAP on cells expressing fluorescently labeled Tet1 or Tet2. FRAP doesn't allow direct measurements of residence time, since the rate of fluorescence recovery depends on the diffusion rates as well as the association and dissociation rates (Mueller et al, 2013). Nonetheless, FRAP measurements serve as a semi-quantitative readout of global protein mobility.

We thus performed FRAP on cells expressing HaloTet1 or HaloTet2. Briefly, a small area of ~1 μm in diameter was illuminated with a high intensity laser, and a time-lapse image series was then acquired at low laser intensities (FIGURE 3.7 A), consisting of 4 images per second for a total duration of one minute. Fluorescence intensity profiles of the bleached area indicated visibly contrasting kinetics for Tet1 and Tet2. While Tet2 recovered to roughly ~95% of its initial intensity, Tet1 recovered to only ~80% of its initial intensity after one minute (FIGURE 3.7 B). These data indicate a globally lower mobility of Tet1 compared to that of Tet2, on the tens-of-seconds timescale.

In agreement with a molecule that has a larger bound fraction, the initial fluorescence intensity recorded after photobleaching (t ~ 250 ms) was consistently lower for Tet1 compared Tet2, (0.39 ± 0.015 for Tet1; 0.55 ± 0.011 for Tet2, normalized to the average prebleach intensity). Consistently, since freely diffusing unbleached molecules rapidly exchange with freely diffusing bleached molecules, the bleached area was more narrow for Tet1 (σ, 0.64 μm for Tet1 0.84 μm for Tet2).

As a rough estimate of recovery, the half-time of recovery for Tet1 and Tet2 were ~5 and ~2 seconds, respectively. Together, these data suggest that the Tet1 might have a higher on-rate and a lower off-rate than Tet2. Such a result would translate to Tet2 having a shorter residence time than Tet1, as well as a greater diffusive fraction, and possibly a greater diffusion coefficient than Tet1. To measure these properties more directly, we will turn to single-molecule tracking at different timescales as well as Raster Image Correlation spectroscopy.

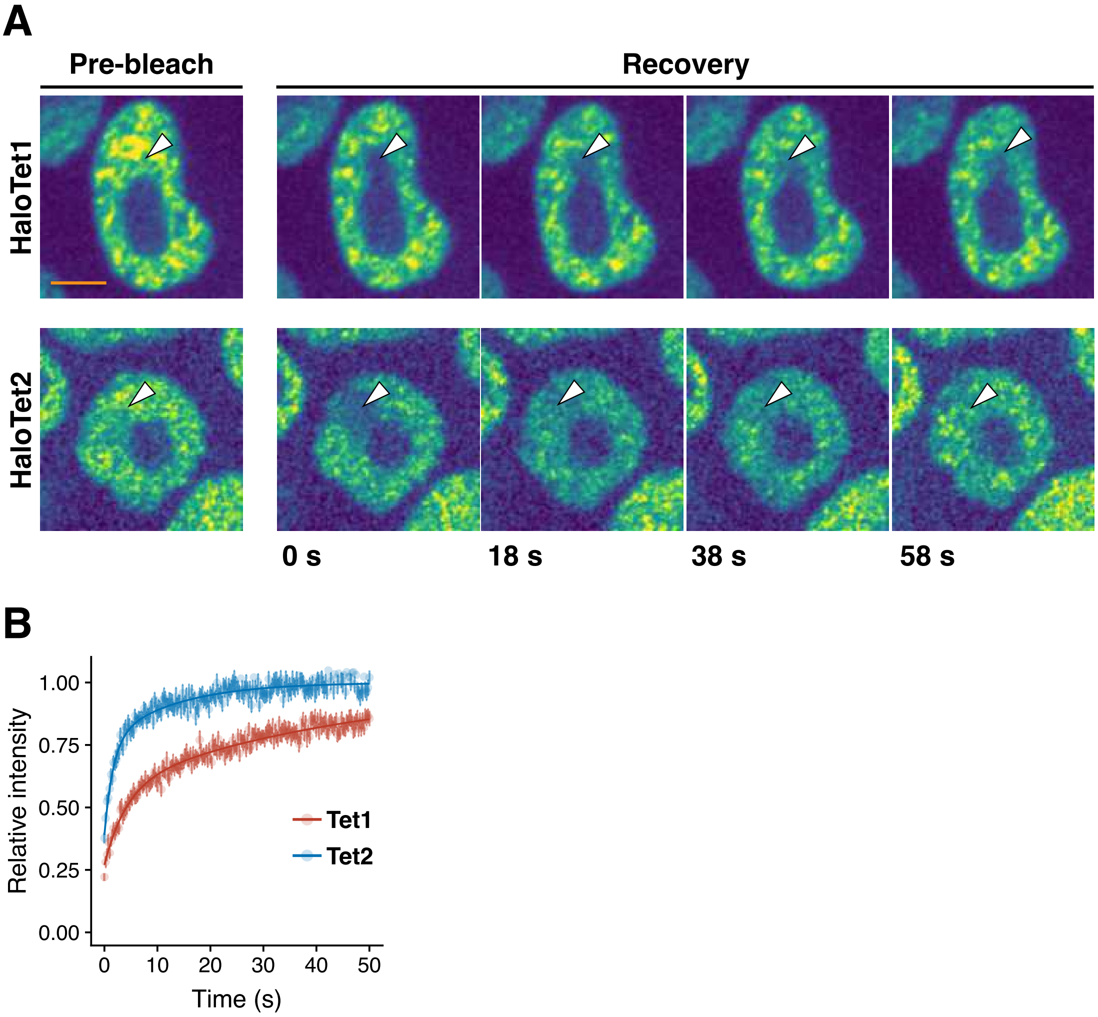


Figure 3.7 Fluorescence recovery after photobleaching of mNeonGreen-labeled Tet1 and Tet2 in embryonic stem cells. A, example images of mNeonGreen-Tet1 before bleaching (first panel on the left "Pre-bleach") and after bleaching ("Post-bleach"), with bleached spot indicated with a yellow box. B, averaged normalized fluorescence intensity traces over time for Tet1 (red) and Tet2 (blue), where timepoint 0 indicates the time of the first post-bleach image. Scale bar = 5 µm

3.1.4 Residence time measurements by single-molecule tracking

We then asked what are temporal aspects of Tet protein binding in the nucleus, in other words, what is the duration of Tet proteins binding events at chromatin loci - referred to as the mean residence time (reciprocal to the off-rate). To this end, I established 2D single-molecule tracking in the Leonhardt laboratory, to first extract residence time from the trajectories of single Tet proteins, and the diffusive properties of Tet proteins (discussed later in section 3.1.3).

For any single-molecule tracking experiment, the main requirement isn't necessarily the imaging system, but rather the labeling strategy. In order to distinguish individual molecules, sparse labeling needs to be achieved in the sample, wherein only a subset of the molecules of interest are fluorescent. To this end, different strategies have been employed, including micro-injection of labeled proteins, photobleaching autofluorescent proteins until low densities remain, photoactivation/ photoconversion of fluorescent proteins, and using low densities of ligands for self-labeling enzymes, such as Halo and SNAP tags (reviewed by Liu and Tijan, 2018).

To this end, we incubated cells expressing Halo-H2B, HaloTet1 or HaloTet2 with picomolar and nanomolar concentrations of Halo-ligands respectively conjugated with JF549 and JF646 dyes (FIGURE 3.8 A). In this labeling regime, only a small fraction of Tet proteins are fluorescently labeled, which allows to distinguish individual molecules in one channel. Meanwhile, JF646, a spectrally separate dye, allowed to visualize the bulk distribution of Tet in another imaging channel. We then imaged these cells on a spinning disk confocal microscope equipped with an EMCCD camera. With this labeling scheme, we managed to visualize in one channel a sparse constellation of spots, thought to correspond individual protein molecules, and in the other, the bulk of the protein labeled (FIGURE 3.8 B).

Due to the resolution limit, both single-molecules and aggregates of several molecules will appear as a diffraction-limited spots. To ensure that fluorescently labeled diffraction-limited objects are really in the range of single-molecules, we labeled and fixed cells expressing HaloTet1 (FIGURE 3.9 A), and performed time-lapse imaging to bleach these spots (FIGURE 3.9 B). We then plotted the intensity profiles over time (FIGURE 3.9 C). The majority of these profiles revealed a single photobleaching step, as opposed to multiple steps or exponential decay patterns, indicating that these spots do correspond to single molecules of Tet1. Note that this indicates that the fluorescent signal emitted from that object comes from a single fluorophore covalently bound to a single HaloTet molecule, but that it is unknown at this stage whether particular Tet molecule is bound to other proteins.

To measure the residence time of Tet proteins in living cells, we performed single-molecule tracking of Halo-tagged Tet proteins. We used a low laser power (~2 mW measured behind the objective) with a long exposure time (0.5 - 1 s) in order to minimize photobleaching, and to "blur out" quickly diffusing particles. With these acquisition parameters, stably bound molecules are clearly visible, whereas diffusive molecules fade into the background due to motion blur (FIGURE 3.10 A).

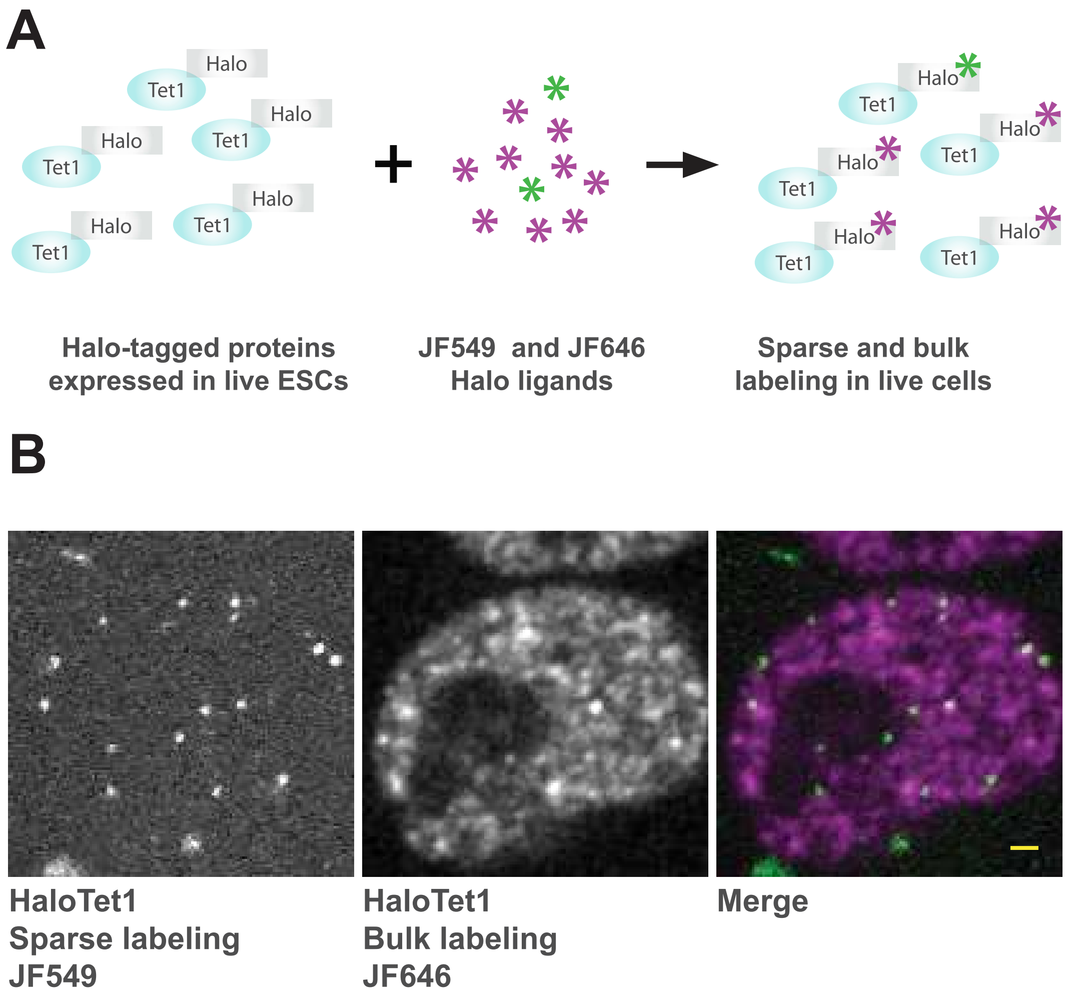


Figure 3.8 A, overview of Halo labeling method to obtain sparsely-labeled and bulk-labeled populations of a Halo-tagged protein, using two spectrally distinct Halo dyes at different concentrations. B, example images of cells expressing HaloTet11 labeled sparsely with pico-molar concentrations of Halo-JF549 and nano-molar concentrations of Halo-JF646 ligands. Scale bar = 1 μm

Stably bound molecules are clearly visible until they photobleach, or unbind and diffuse away (FIGURE 3.10 B). We then extracted the spatial trajectories of single-molecules from these time-lapse image series using the Trackmate plugin in Fiji (Tinevez et al, 2017). Briefly, individual spots of a given radius (~300 nm) are found in each time frame of the time series and their centroid is determined by a gaussian fit. Then, spots are linked from one frame to the next based on a modified nearest neighbour algorithm, with a limited search radius of 250 nm (to account for drift and localization error). Trajectories started when a molecule appeared in focus in the field of view, and were detected by Trackmate, and trajectories ended when the spot was no longer detected, either due to diffusing out of the focal plane, or due to photobleaching. We then used the duration of the trajectory as a readout for the residence time of the Tet proteins at chromatin loci.

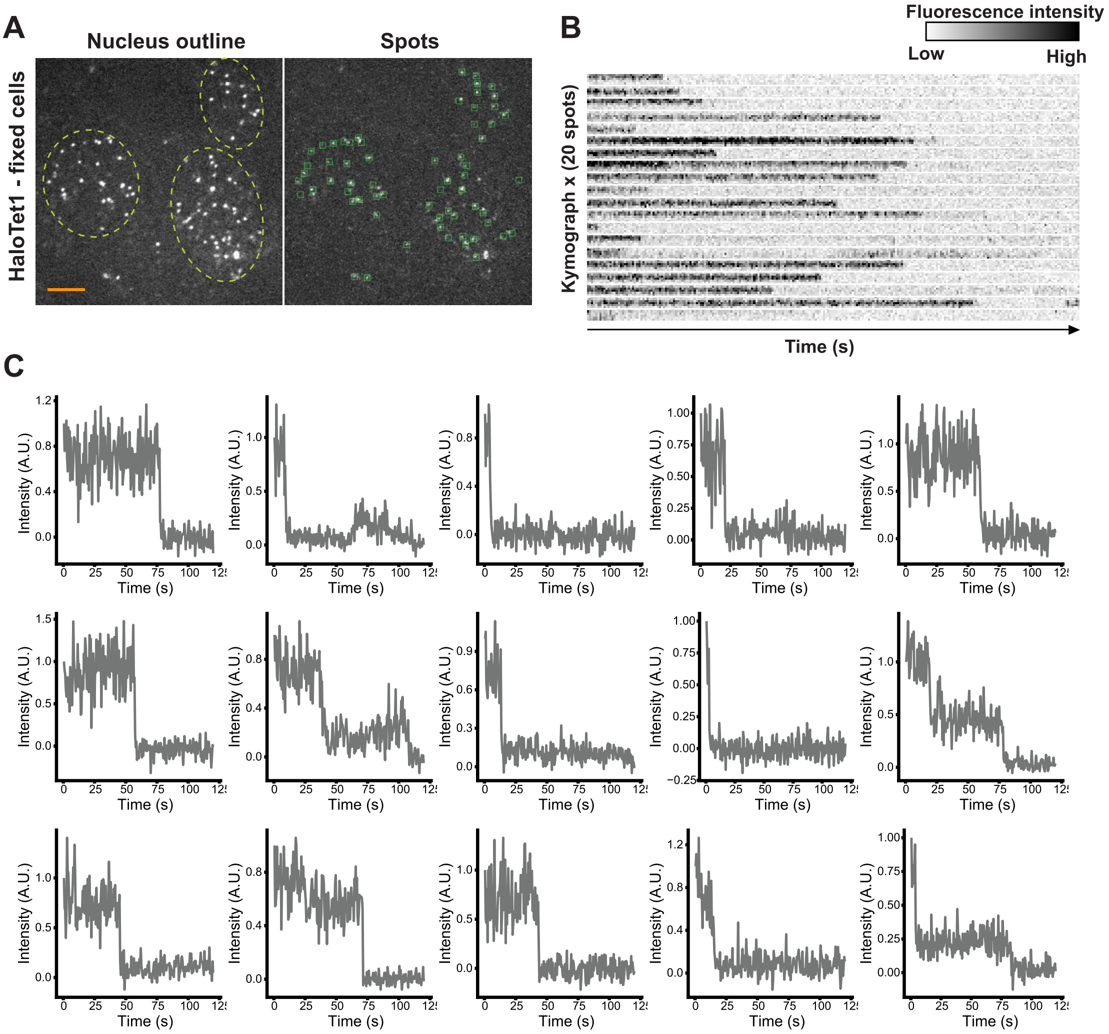


Figure 3.9 Photobleach step counting to reveal single-molecule nature of fluorescent signal. A, example image of fixed cell expressing HaloTet1, sparsely labeled with HaloJF549, nucleus outline in yellow, spots outlined in green. B, examples kymographs (mean intensity of spots outlined in green in A) throughout a time-lapse acquisition. C, example fluorescence intensity profiles over time of several spots outlined in A, showing step-like drops in intensity. Scale bar = 5 μm

Since photobleaching will artificially lead to shorter trajectories and thus decrease these residence time measurements, we performed SMT on cells expressing Halo H2B. Since H2B is expected to have a residence time on the scale of hours, the duration of SMT trajectories will be mostly limited by photobleaching, tracking errors and chromatin movement. Thus, we can use H2B trajectories for photobleach correction. By plotting all trajectories as a cumulative density function (CDF), one can calculate the survival probability (1-CDF) of trajectories as a function of time, and normalize Tet1 and Tet2 survival probability distributions to that of H2B.

We first performed this measurement on a spinning disk microscope. By fitting the survival probability to a two-component exponential, we can estimate transient and stable residence times for Tet1 and Tet2 (FIGURE 3.10 C). We calculated that the mean residence time for stably bound Tet1 is ~26 seconds. In contrast, the mean residence time for stably bound Tet2 was ~15 seconds. We note that the tail of the survival probability distribution extends into the minutes timescale, in keeping with a few rare cases of both Tet1 and Tet2 binding for extended periods of time on the minutes timescale. These data are consistent with our FRAP data, where global but incomplete fluorescence recovery is observed after one minute. Thus, Tet1 is likely bind chromatin loci for longer periods of time than Tet2.

It is important to note that even with rigorous photobleach correction, residence time measurements by SMT are still limited by tracking errors due to cell movement, drift, fluorophore blinking, etc. Thus, these data should be interpreted as representative of the lower boundary of the distribution of residence times (Hansen et al, 2017). Taken together, these residence time measurements show that Tet1 and Tet2 both show a broad and continuous distribution of residence times at chromatin loci on the seconds timescale, and that Tet1 tends to bind chromatin for longer periods of time than Tet2.

3.1.2 Raster Image Correlation Spectroscopy reveals Tet protein diffusive dynamics

The following RICS results were performed with Philipp Messer, along with the technical assistance of Nader Danaf in the group of Prof. Don Lamb (LMU Physical Chemistry).

Our FRAP and slow SMT revealed that Tet1 to binds for longer periods of time than Tet2. However, these methods do not yield information of the diffusive properties of proteins. To dissect the the diffusive properties of Tet proteins and their contribution to global protein mobility, we first used Raster Image Correlation Spectroscopy. RICS is a powerful method which uses the spatio-temporal information of fluorescence fluctuations of an image acquired on a laser scanning confocal microscope to determine the diffusive properties of proteins within living cells.

Since the pixels of an image are acquired sequentially, one line at a time, raster scanned images contain information on two different timescales: for neighbouring pixels on the same line, the time interval is ~10 µs, whereas for neighbouring pixels on two different lines, the time interval in on the order of ~ 5 ms.

Using a pixel size of 40 nm, particles that are immobile on the timescale of 4-5 lines (~20 ms) will appear as diffraction limited objects, since they remain in the excitation volume from one line to the next. However, for particles that diffuse quickly, while they may be visible on one line, they will have diffused away milliseconds later, by the time the next line is acquired. This results in images which have a striped appearance, where the persistence of the signal is different between the x-axis and the y-axis.

To get a global view of this persistence on the x- and y- axes, one can calculate the spatial autocorrelation function (ACF), the shape of which gives an indication of how similar is the image compared to the same image shifted by (i,j) pixels in x and y. For immobile particles, the shape of the ACF corresponds to the size of the excitatory volume, which is essentially the PSF of the excitation laser. For diffusing particles, however, the shape of the ACF is thinner in the y-axis, since it is less likely that the signal of a particle persists on the following lines. With a large enough sample size, and given the known timing of the acquisition between pixels and between lines, the space between each acquired pixel, and the size of the excitation volume, one can fit a diffusive model to the ACF, from which global spatio-temporal parameters can be derived such as the concentration of the protein of interest, its bound and diffusive fractions, and the diffusion coefficient of its diffusive fraction. These parameters are derived in unbiased way, in that individual data points and not selected or thresholded.

In order to gain information on the diffusive populations of Tet1 and Tet2, we performed RICS on cells expressing fluorescently labeled Tet1 and Tet2. Correlation spectroscopy measurements are exquisitely sensitive to variations/fluctuations within the sample as well as fluctuations intrinsic to the acquisition equipment. To control for this, we performed RICS on two cell lines expressing fluorescently labeled Tet proteins: mNeonGreen-Tet1 / HaloTet2, and HaloTet1 / mNeonGreen-Tet2. First, this biological replicate ensured that any information derived wasn't exclusively from a single clone. Second, the swapped fluorophores allow to rule out discrepancies based on the labeling method, since mNeonGreen is a constitutively fluorescent protein, whereas the Halo requires an ectopic enzymatic labeling step. In parallel, the swapped fluorophores also rule out discrepancies based on the excitation lasers and detectors of their respective measurement channels.

After acquiring confocal images (FIGURE 3.11 A), we calculated the ACF, which already visually suggested different mobility for Tet1 and Tet2. The ACF of Tet1 was generally broader along the y-axis compared to the ACF of Tet2, which suggests that the overall mobility of Tet1 is less than that of Tet2 (FIGURE 3.11 B, top row). We next fit a diffusive model to this ACF (FIGURE 3.11 B bottom row), in order to extract quantitative information on the biophysical behaviour of Tet1 and Tet2.

The quantitative information derived from our RICS measurements revealed some similarities between Tet1 and Tet2. Notably, the global diffusion coefficients

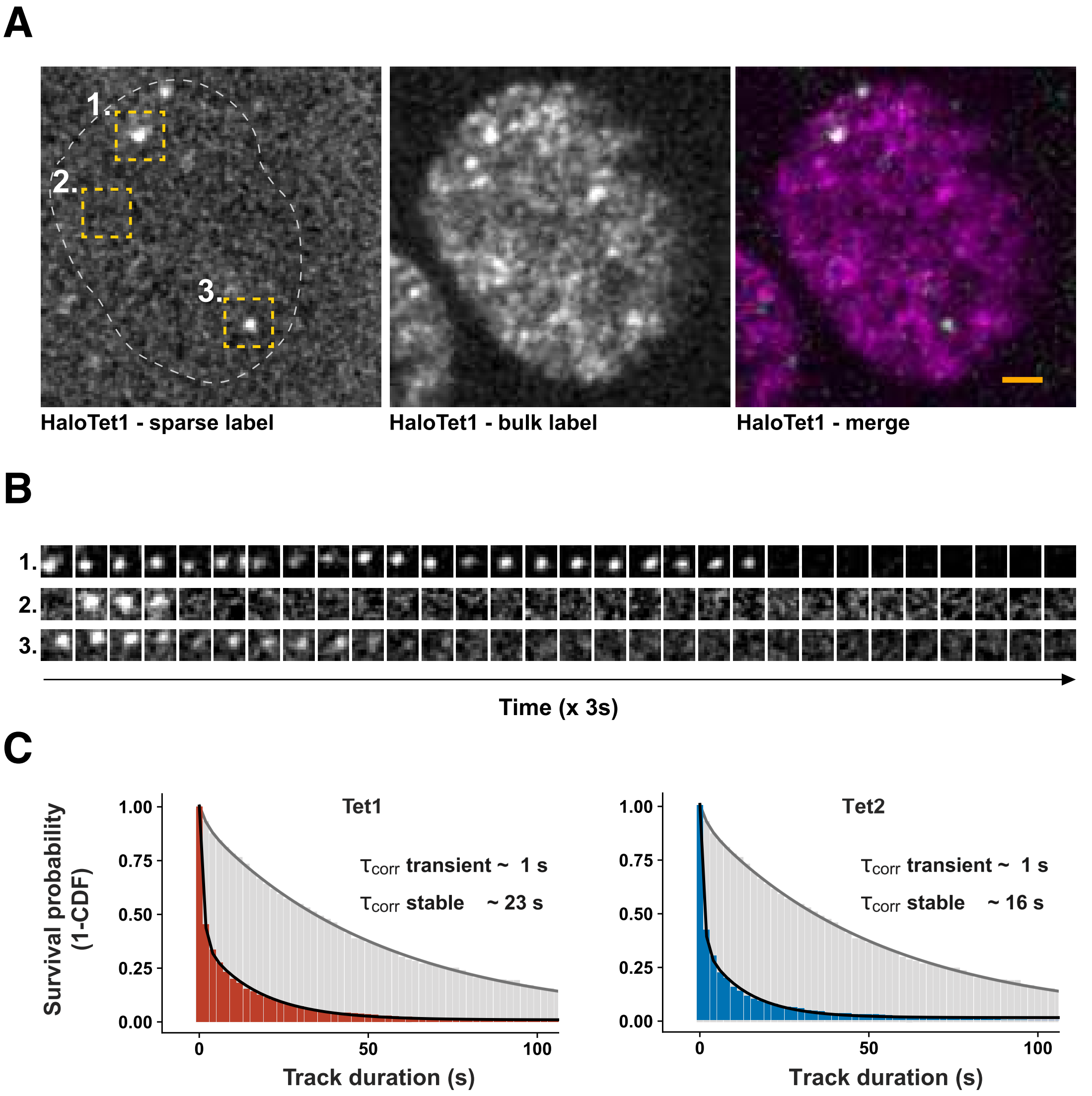


Figure 3.10 Residence time measurements using single-molecule tracking at slow acquisition rates. A, example image of a cell expressing HaloTet1 sparsely labeled with JF549 (green) and bulk labeled with JF646 (magenta) with three regions of interest outlined in white. B, time series images of regions outlined in white in A. C, survival probability plots of Tet1 (red) and Tet2 (blue), calculated from the cumulative density functions of track durations. Cumulative distribution functions were fit to a two-component exponential model, in order to determine average binding times. Grey bars show survival probability plot of HaloH2B, and its slow fit was used to correct for photobleaching and tracking errors. Scale bar = 2 μm

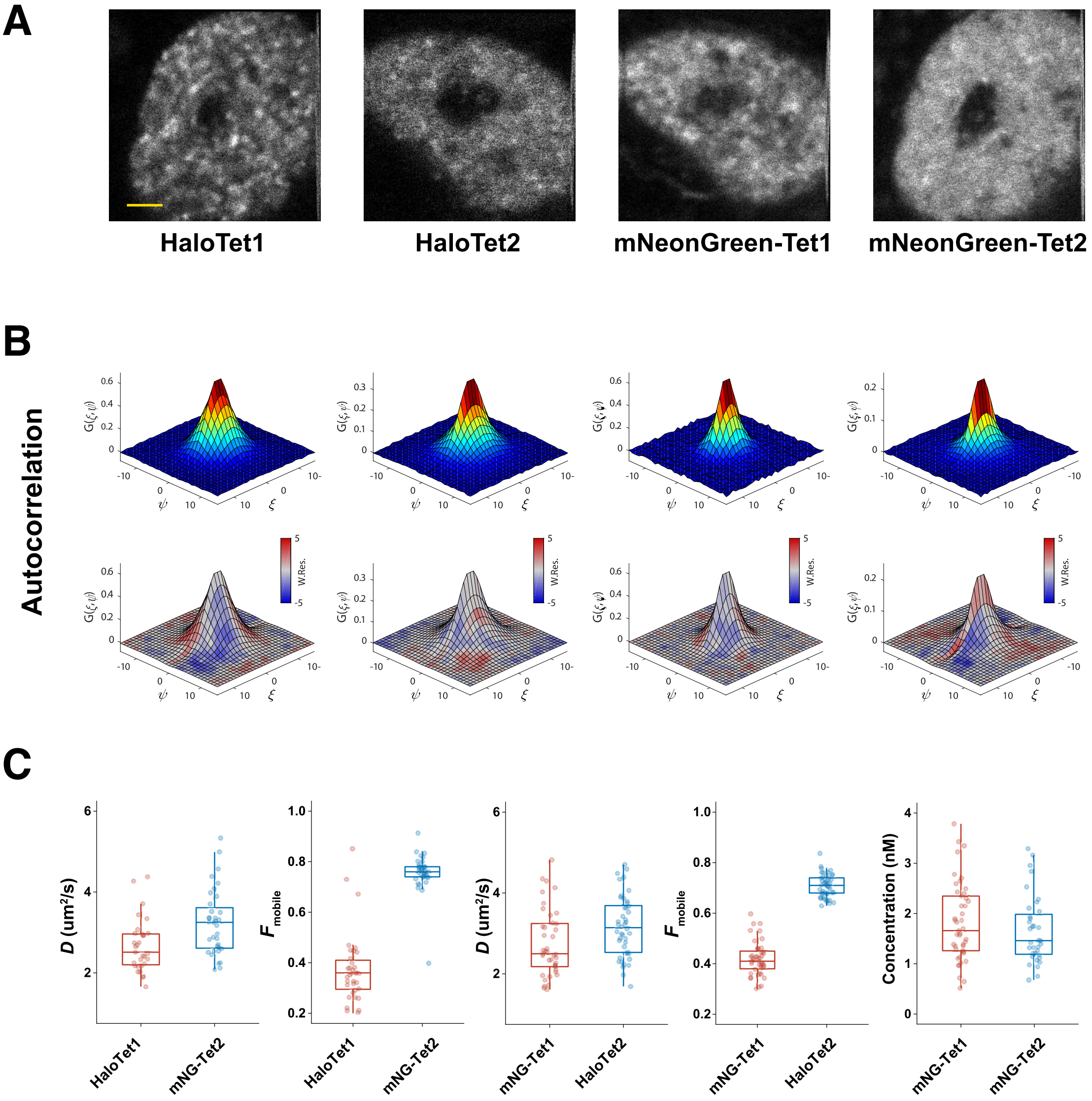


Figure 3.11 Raster Image Correlation Spectroscopy measurements of cells expressing HaloTet1 and mNeonGreen-Tet2, or mNeonGreen-Tet1 and HaloTet2. A, example sum projection images of HaloTet1, mNeonGreen-Tet2, mNeonGreen-Tet1, HaloTet2. B, autocorrelation function of fluorescence fluctuations averaged at different spatial intervals (top), and two-component diffusive model fit (bottom), with weighted residuals coloured in blue and red. C, Diffusion coefficients, mobile fractions, and concentrations of Tet1 and Tet2 derived from model fit. Scale bar = 2 μm

measured for Tet1 and Tet2 were in a similar range for both cell lines tested, (FIGURE 3.11 C and E; mean ± standard deviation: 3 ± 1.4 $\mu m^2/s$ for Tet1 and 3.2 ± 0.9 for Tet2 labeled with Halo-JF646; 2.9 ± 1.2 $\mu m^2/s$ for Tet1 and 3.2 ± 0.8 $\mu m^2/s$ for Tet2 fused to mNeonGreen). In both cell lines, interestingly, our measurements indicated that Tet2 had a slightly higher diffusion coefficient than Tet1. These data are in line with the measured diffusion rates of other epigenetic modifiers and transcription factors, which are generally found to be in the range of 1-5 $\mu m^2/s$.

The difference in diffusion coefficient was unlikely to fully explain the difference in the shape of the ACF. For example, by comparing ACFs of Tet1 and Tet2 individually selected for having similar diffusion coefficients, a thinner y-profile of Tet2 is still visible. This indicated that Tet1 and Tet2 may have a greater difference in their bound populations. Indeed, our RICS measurements revealed a dramatic difference in the bound fractions of Tet1 and Tet2. Tet1 had a much larger bound fraction compared to Tet2 (FIGURE 3.11 D and F; 0.63 ± 0.14 for Tet1 and 0.29 ± 0.05 for Tet2 labeled with Halo-JF646; 0.58 ± 0.07 and 0.24 ± 0.08 for Tet2, labeled with mNeonGreen). These data establish a model where Tet1 is much more likely to be in a bound state than Tet2, over a timescale of ~25 ms, which is the time required to acquire 5 lines on the confocal scan.

Finally, we calculated the concentration of Tet proteins from each cell line. mNeonGreen-Tet1 and mNeonGreen-Tet2 had similar concentrations in the nucleus (FIGURE 3.11 G; 3.6 ± 1.5 nM for Tet1 and 3.4 ± 1.4 nM for Tet2). We also compared HaloTet1 with HaloTet2, which revealed that both proteins were expressed at similar levels relative to one another, but at lower levels compared to mNeonGreen-labeled Tet proteins (data not shown). We reasoned that this was due to incomplete labeling with the HaloTag, since the ligand stock used to label the cells may contain unlabeled ligand or ligand covalently bound to already bleached fluorophores.

Taken together, these data show that the physical behaviours of Tet1 and Tet2 are different in live cells. Tet1 and Tet2 have diffusion coefficients around ~3 $\mu m^2/s$, similar concentrations in the range of 3 nM, but largely different bound fractions where the Tet1 is likely bound over the course of ~25 ms, whereas the majority of Tet2 is diffusive in that time window.

3.1.3 Fast Single-molecule tracking of Tet proteins

Following our RICS results, we sought to examine these binding and diffusive behaviours at the single-molecule level in living cells. To this end, we used a sparse labeling scheme as described earlier, in combination with HiLO imaging achieved using a TIRF illuminator, in order to acquire time-lapse image series at time intervals relevant to the diffusive behaviours of nuclear proteins.

We thus used this system to measure the dynamics of Tet1 and Tet2 in living cells. Cells expressing HaloTet1 and HaloTet2 were sparsely labeled with Halo-JF549, and bulk labeled with Halo-JF646. In parallel, we used H2B as a control for a stable binding. We then acquired time-lapse images at 30 frames per second with continuous illumination, in order to observe individually moving proteins.

We then applied the same tracking algorithm as before, but with a larger radius between successive frames (~1 µm), in order to track diffusive molecules. We determined that a maximum step size of 1 µm was sufficient to capture molecules diffusing at ~3 $\mu m^2/s$. Resulting trajectories can be directly observed, overlaid on an image, and the spatial and temporal coordinates of the steps can be used for further analysis. A zoomed-in view of several chosen trajectories are shown in FIGURE 3.12.

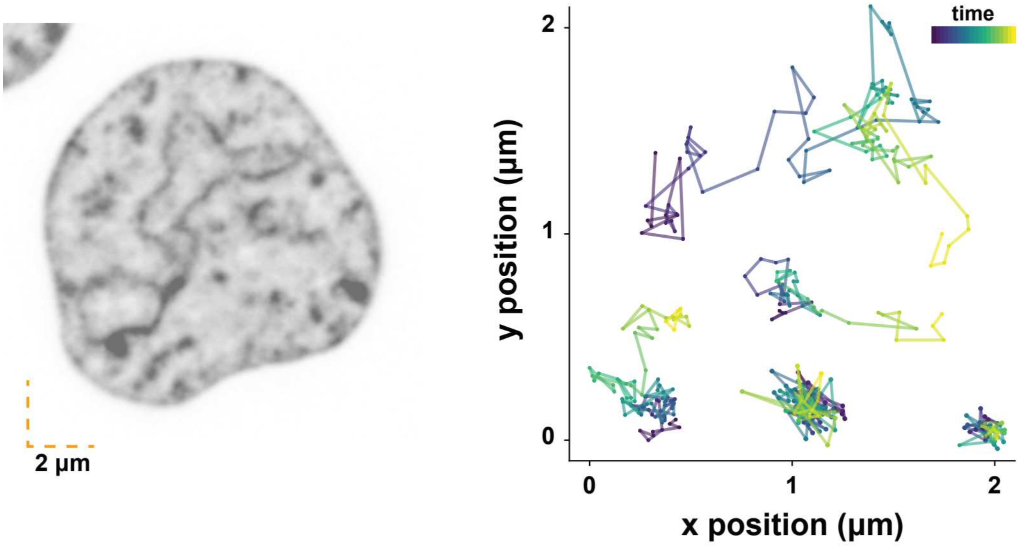


Figure 3.12 Example image of cell nucleus (labeled with Halo-H2B) shown as a reference for size, and example single-molecule trajectories acquired using fast acquisition rates (30 frames per second), placed in a 2 x 2 micron plane. Colour palette indicates time point within the trajectory.

We performed time-lapse imaging of sparsely labeled Tet proteins as well as H2B at fast acquisition rates (FIGURE 3.13 A). Single-molecule tracking revealed visibly different geometries of trajectories. Specifically, Tet1 tracking revealed a greater fraction of trajectories consisting of short displacements (<100 nm) and largely confined to small area. Such confined trajectories are typically interpreted as bound molecules, since the measured displacements are dominated by the localization inaccuracy of the imaging system. In contrast, Tet2 appears to show a greater proportion of trajectories corresponding to diffusive molecules, characterized by displacements larger than the localization accuracy of the system (~100 nm - 1 µm). H2B, used as a control for stably bound proteins, appeared to mostly contain trajectories of short displacements. Example trajectories from HaloH2B, HaloTet1 and HaloTet2 are shown in FIGURE 3.13 B.

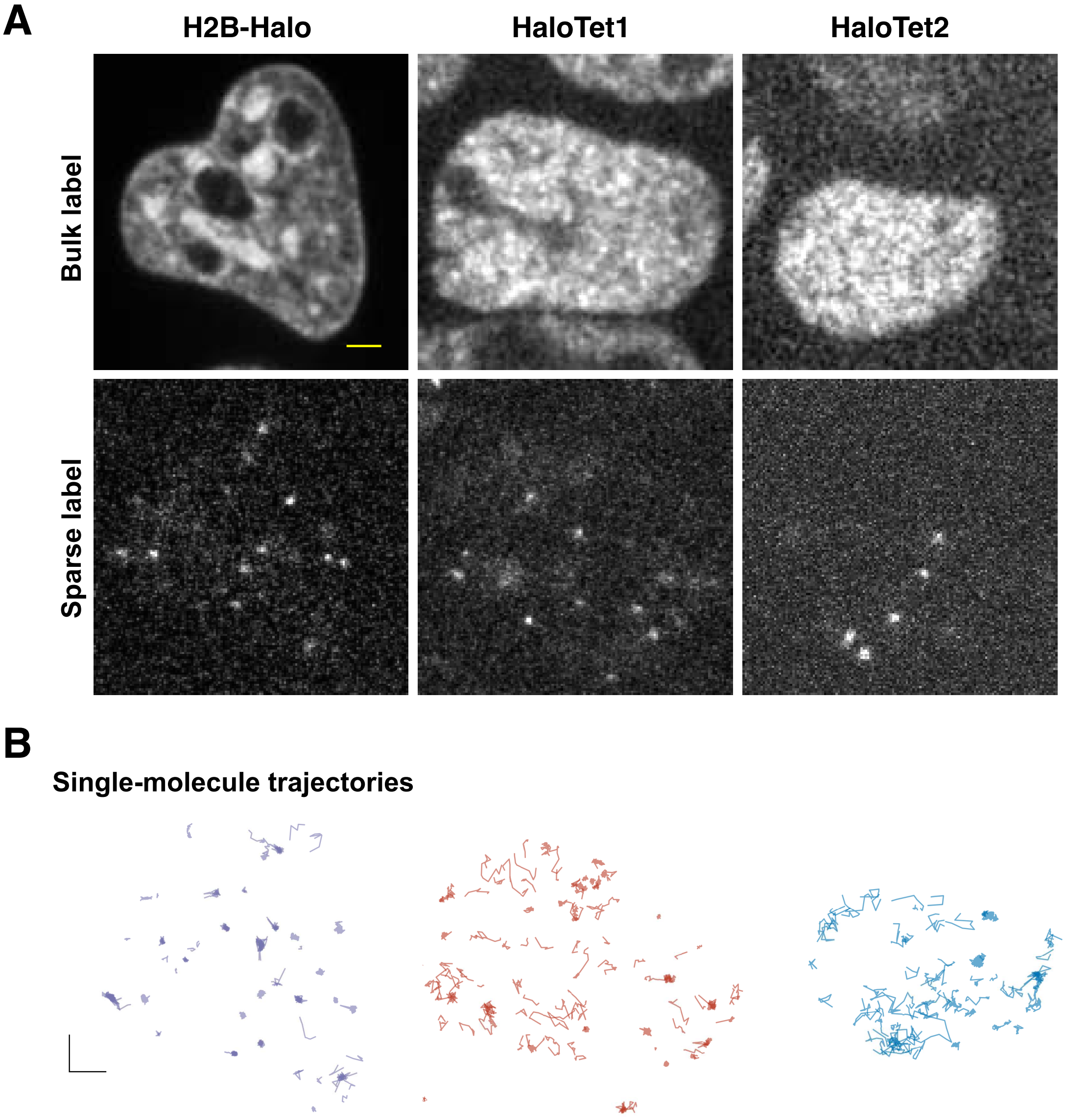


Figure 3.13 A, example images of cells expressing cells expressing H2B-Halo, HaloTet1 and HaloTet2, bulk labeled with HaloJF646 (top panels) and sparsely labeled with HaloJF549 (bottom panels). B, single-molecule trajectories measured from cells shown in A, shown as per their xy positions in the cell. Scale bars = 1 µm

To globally compare Tet1 and Tet2 trajectories, we then performed an unbiased quantitative assessment of trajectories without classification or filtering for "bound" or "diffusive" trajectories. We calculated the mean square displacement (MSD) at successive time intervals, which is typically informative about confinement in diffusive trajectories (FIGURE 3.14 A). Here, we see that the MSD of Tet2 increases more sharply with increasing time intervals than that of Tet1, and that the MSD of Tet1 increases more sharply than H2B. In this type of analysis, the degree of confinement globally corresponds to a degree of binding, since a bound molecule would lead to a small MSD at long time intervals, as seen for H2B. Taken together, a global analysis of our SMT data suggests that, consistent with our RICS data, Tet1 is more likely to be bound to chromatin loci than Tet2.

We sought to extract more information regarding the distributions of displacements, and fit them to a kinetic model. In order to quantify the binding and diffusive behaviours Tet1 and Tet2 from single-molecule trajectories, we used the recently developed SpotON software package (Hansen et al. 2017) to fit a kinetic model. By compiling the distributions of all displacements at several time intervals, data sets of single-molecule trajectories can be fit to a two- or three- state kinetic model. One can then derive the fraction of bound molecules (Fbound) as well as the diffusion coefficient of the freely diffusing fraction of molecules (Dfree) from this model. We found that a three-state model best fit the data acquired at 30 frames per second, consisting of a bound state, a slow state, and a fast state. From this model, we found again that a much larger fraction of Tet1 is bound to chromatin compared to Tet2 (FIGURE 3.14 B; 0.29 ± 0.004 vs 0.1 ± 0.002), and thus a much lower freely diffusing fraction (0.37 ± 0.008 vs 0.71 ± 0.008). Finally, the freely diffusing fraction of Tet1 has a slightly lower Diffusion coefficient than Tet2 (FIGURE 3.14 C; 1.47± 0.05 $\mu m^2/s$ for Tet1; 1.95 ± 0.04 $\mu m^2/s$ for Tet2). Distributions of displacements at different time intervals used to fit Fbound and Dfree are given in FIGURE 3.14 D.

While the absolute values obtained from these data are different than those obtained from the RICS measurements, they converged well in terms of the relative differences in mobility. Both SMT and RICS show that Tet2 has a slightly higher diffusion coefficient than Tet1. The SMT measurements revealed slower diffusion rates than RIC. This is possibly due to the lower signal to noise for quickly diffusing molecules, which affects SMT to a greater extent than RICS. Furthermore, since RICS operates at a faster rate than SMT experiments, transient binding events would more likely to affect the fit diffusion coefficient of SMT measurements. In terms of chromatin binding, both RICS and SMT show that Tet1 has a greater bound fraction than Tet2. In this case, the lower overall fraction is largely due to the three-state fit. The slowly diffusing populations of Tet1 and Tet2 could be a mix of slowly moving molecules, confinement, or transient binding. Altogether, our FRAP, RICS, and SMT results demonstrate that as fast timescales, Tet1 is mostly bound to chromatin loci, whereas Tet2 is mostly diffusive, and at longer timescales, Tet1 binds for a longer period of time compared to Tet2. Tet1 and Tet2 have similar diffusion coefficients, and Tet1 is likely to have a higher on-rate that Tet2.

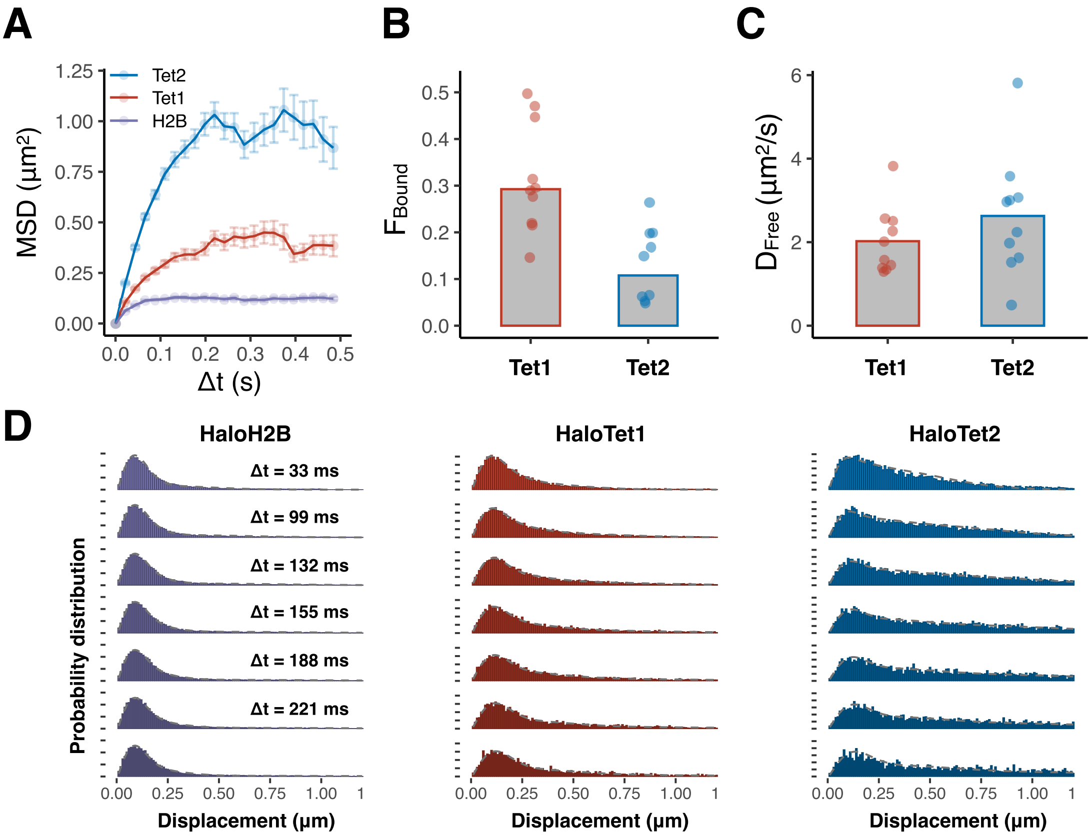


Figure 3.14 Quantification of single-molecule trajectories of HaloH2B (purple), HaloTet1 (red) and HaloTet2 (blue) from time-lapse images acquired at 30 frames per second. A, mean squared displacement calculated at sequential time intervals. B, Fraction of bound molecules and C, average diffusion coefficient of diffusing molecules, derived from a 3-state kinetic model fit using SpotON software. D, distribution of step sizes at different time intervals used to fit kinetic model.

3.1.5 Tet1 and Tet2 dynamics are methylation-independent

3D-acMFM imaging was performed at the Janelia Research Campus along with Christopher Mulholland and Miha Modic, along with the assistance of the Advanced Imaging Center team (Jesse Aaron, Satia Khuon, Ina Pavlova).

Having established a coarse of model of Tet protein mobility in the nucleus, we then asked what are the molecular determinants of Tet protein chromatin association. In addition to having many interaction partners and genomic binding sites, Tet proteins are primarily described in association with changes in methylation levels, with regards to their catalytic domain. To test the influence of DNA methylation of Tet protein localization and dynamics, we inserted the coding sequence of the HaloTag at the Tet1 or Tet2 genomic locus in DNMT triple knockout (DTKO) cells, using the same Cas9-mediated genome engineering as outlined in section 2.1.1. These cells have no DNA methylation, and are thus devoid of TET's enzymatic substrate.

HaloTet1 and HaloTet2 in DTKO cells showed similar relative global subnuclear distribution as in wt cells. Both Tet proteins localized mostly in low-compaction regions, with Tet1 having a more spotty appearance, and Tet2 showing a more diffuse appearance (FIGURE 3.15 A). We quantified this by measuring the coefficient of variation, and which showed that indeed Tet1 had a more clustered and less uniform distribution than Tet2 in the absence of methylation (FIGURE 3.15 B). Both Tet proteins in DNMT-TKO cells also showed mitotic localization as seen wt cells: Tet1 was enriched on mitotic chromosomes, whereas Tet2 was mostly depleted (FIGURE 3.15 A).

We performed FRAP on DMNT-TKO cells expressing HaloTet1 or HaloTet2. As in the wild-type scenario, we saw a similar marked difference in kinetics between Tet1 and Tet2 in DNMT-TKOs (FIGURE 3.15 C), reminiscent of Tet1 and Tet2 FRAP traces in wild type. However we did not observe a substantial difference in recovery of Tet1 between DNMT-TKO and WT cell types (FIGURE 3.15 D). These results suggested that DNA methylation does not influence global chromatin binding of either Tet1 or Tet2.

We then sought to observe the behaviour of Tet proteins in live WT and DNMT-TKO cells at the single-molecule level in 3D. To this end, we used a custom-built aberration-corrected 3D Multifocus Microscope (3D-MFM), at the Advanced Imaging Facility of the Janelia Research Campus. We employed a sparse labeling scheme as described earlier. We incubated cells expressing HaloTet1 or HaloTet2 with 1 μM Halo-JF549 ligand, in order to only label a fraction of Tet molecules and imaged them at 30 frames per second on the 3D-MFM. On this system, samples are epi-illuminated, and the collected emitted light from different focal planes are focused on different areas of the camera, using a multifocus grating and a prism (Abrahamson et al, 2013). Thus in a single image acquired on a camera, nine focal planes are simultaneously imaged (FIGURE 3.16). With deconvolution and reconstruction, 3D stacks of 9 focal planes can be generated from a single image exposure (FIGURE 3.16). By running the camera at its fastest acquisition rate (in this case, 30 frames per second), and at high laser intensities (~ 50 mW in infinity space behind the back aperture of the objective), we could then observe the mobility of individual proteins in the nuclei of living cells in 3D (FIGURE 3.17).

We first verified the behaviour of Tet1 and Tet2 in wild-type ESCs on the 3D-MFM system. To this end, we labeled cells expressing HaloTet1 or HaloTet2 with Halo-JF549, and performed time-lapse imaging on the 3D-MFM as fast acquisition rates. We reconstructed and deconvolved 3D stacks, and performed single-molecule tracking in 3D. Examples of trajectories from naive ESCs in interphase expressing HaloTet1 or HaloTet2 are shown in FIGURE 3.18 A, respectively. To get an overall view of the dynamics of Tet1 and Tet2, we plotted the distribution of step sizes of sequential spots in a trajectories detected (FIGURE 3.18 B). As expected based on our measurements on different systems, Tet1 had a larger "left-shifted" population of steps, consistent with a more consistently bound molecule, whereas Tet2 had a wider distribution of step sizes, consistent with a protein that has a larger diffuse fraction.
We then calculated the mean squared displacement, and again found that Tet2 was more likely to travel a larger distance than Tet1 at consecutive time intervals, consistent with a larger bound fraction for Tet1, and a larger diffuse fraction for Tet2. Diffusive molecules are more difficult to track compared to stably bound molecules, since their movement leads to motion blur, and frame-to-frame tracking errors. Thus, as a control, we performed 3D single molecule tracking on cells in mitosis, where a greater fraction of molecules is diffusing throughout the volume of the cell, especially in the case of Tet2. Trajectories of Tet molecules in mitotic cells visibly cover longer distances than in interphase cells (FIGURE 3.19 A), which is reflected in the larger distribution of step sizes, owing to reduced overall binding (FIGURE 3.19 B). MSD measurements show that Tet proteins in mitotic cells tend to cover a larger distance at sequential time intervals (FIGURE 3.19 C), compared to Tet proteins in the nucleus of cells in interphase (FIGURE 3.18 C). Thus, 3D SMT from images acquired 3D-MFM show clear differences in the binding and diffusive behaviours of Tet1 and Tet2, as well as a greater diffusive fraction in mitotic cells.

We then looked at cells expressing HaloTet1 or HaloTet2 in wild-type cells, DTKOs, as well as cells where Tet1 or Tet2 was rendered catalytically inactive by genetic mutation. We induced the transition from a naive pluripotent state to a primed epiblast-like state, since this transition is accompanied with a increase in methylation, and tested whether the presence or absence of methylation in highly methylated genome affected the binding and diffusive behaviour of Tet1 and Tet2.

To verify that the kinetics of Tet proteins were also largely unchanged in DTKOs, we measured the binding and diffusive behaviour of Tet1 and Tet2 in cells devoid of DNA methylation. We applied the same sparse-labeling, imaging, as described above, but performed 2D single-molecule tracking of 3D image stacks to look at these kinetics. Both Tet1 and Tet2 had a left-shifted distribution of step sizes compared to naive cells, suggesting a higher fraction of bound molecules (FIGURE 3.20). Tet2 appeared again to have a greater fraction of larger steps than Tet1, consistent with previous observations. However, we did not observe any substantial difference in step size distribution of Tet1 between wild type, DTKO, and Tet1 catalytically inactive cell types. We also observed a consistent distribution of step size for Tet2 in all conditions. Thus, abrogating methylation, or rendering Tet1 or Tet2 catalytically inactive did not substantially affect the single-molecule trajectories of Tet1 or Tet2, suggesting that changes in chromatin organization or gene expression, such at those which occur at the epiblast state are more influential than DNA methylation on Tet protein binding and diffusive behaviours.

Taken together, these data suggest that the binding and diffusive properties of Tet proteins are likely intrinsically encoded within the proteins themselves, and are influenced by larger structural changes in chromatin than a simple increase in DNA methylation.

3.2 The N-terminal domain of Tet1 drives its localization and binding

3.2.1 Localization and dynamics of Tet1 deletion mutants

MIN-tag lines were generated by Paul Stolz, Burak Ozan, and imaging was carried out with the help of Katharina Brandstetter.

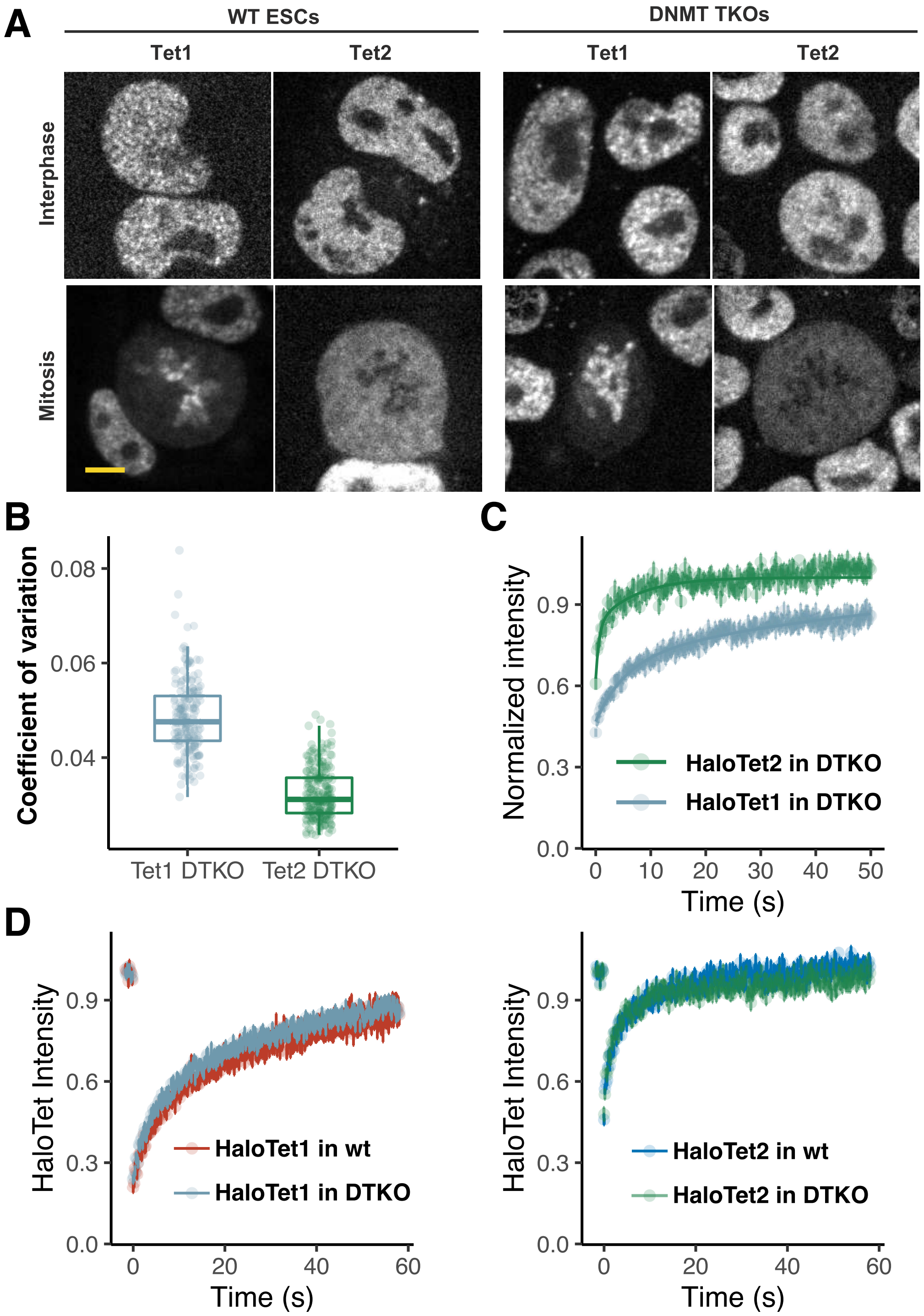


Figure 3.15 HaloTet1 and Tet2 in DNMT triple knock-out cells. A, example images of HaloTet1 and HaloTet2 expressed in wild-type J1 embryonic stem cells (left) and DTKO cells (right), in interphase (top) and during mitosis (bottom). B, coefficient of variation of HaloTet1 (grey) and HaloTet2 (green) in DTKO cells. C, FRAP of HaloTet1 and HaloTet2 in DTKO cells. D, FRAP of HaloTet1 in DTKO cells (grey) and in wild-type J1 cells (red). Scale bar = 5 µm.

Since we didn't observe any dramatic changes in the dynamics of Tet1 and Tet2 in cells lacking DNA methylation, we hypothesized that the binding and diffusive properties of Tet proteins were largely encoded within the protein itself. In this regard, the N-terminal domain of Tet1 is thought to tightly regulate the localization and binding of Tet1 (Zhang et al, 2016), restricting its catalytic activity to euchromatin areas, though a direct observation of this was still lacking.

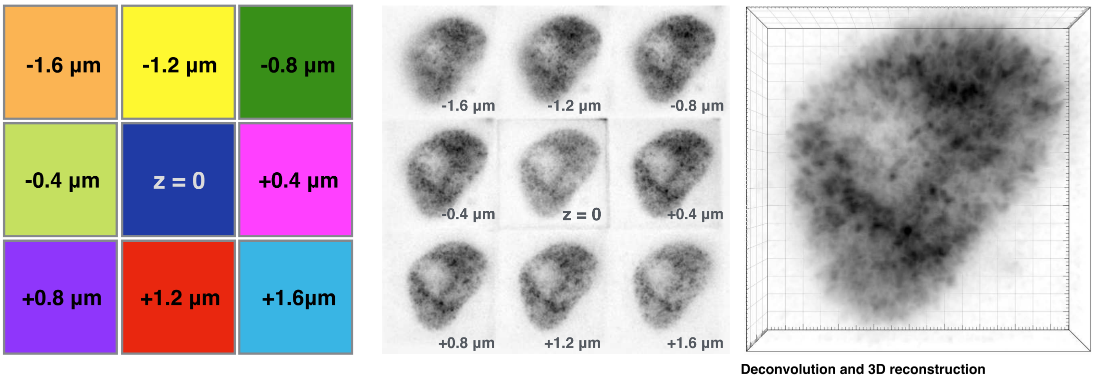
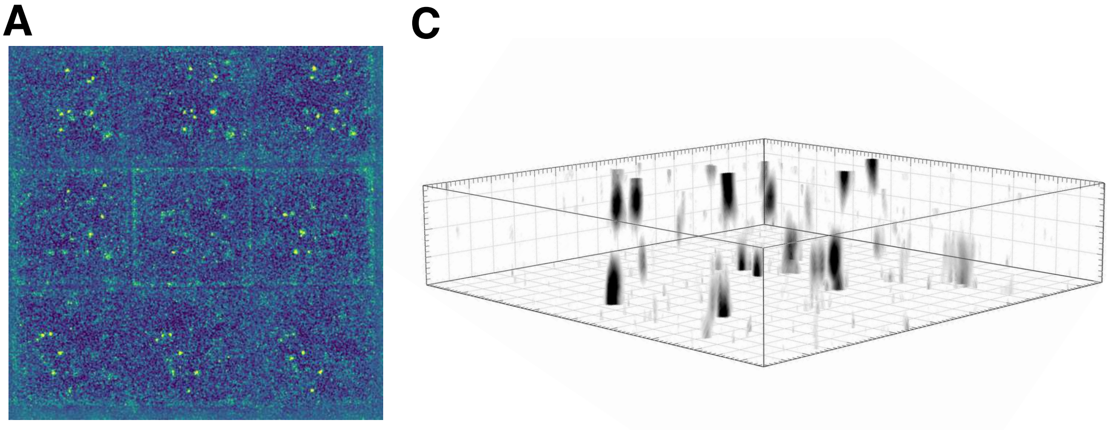


Figure 3.16 Acquisition scheme on 3D aberration-corrected multifocus microscope. Left, Light from different focal planes is focused on different parts of the camera. Middle, example single snapshot image from multifocus microscope of a cell expressing HaloTet1. Right deconvolution and 3D reconstruction from simultaneously acquired focal planes, shown in 3D perspective view rendered in Imaris. Each field of view has a width of 20.4 µm.

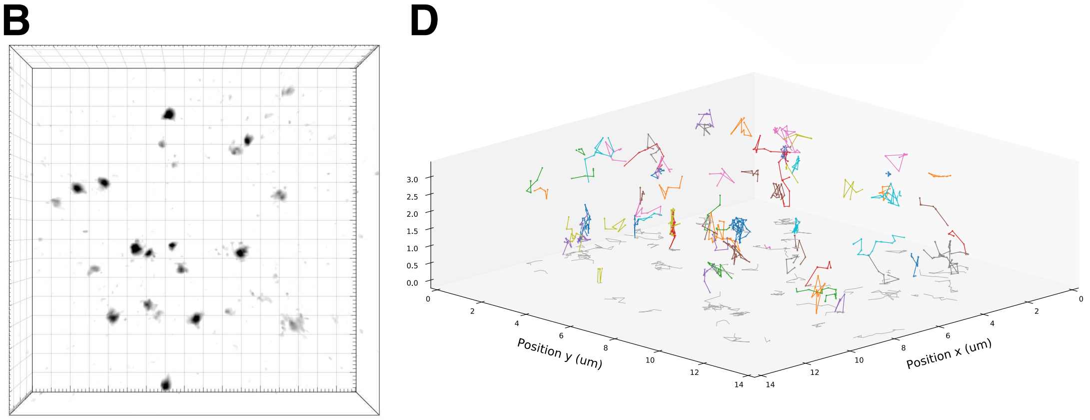
We next looked to quickly screen for changes in localization and dynamics in embryonic stem cells expressing different deletion mutants of Tet1. To do this, we turned to the MIN-tag system (Mulholland et al 2015), and used a cell line harbouring a MIN-tag integration site at the endogenous Tet1 locus. In this integration site, we then inserted cDNA fusion constructs consisting of the coding sequence of mNeonGreen in frame with the coding sequence of Tet1 or a deletion mutant of Tet1, followed by a stop codon and poly-A signal. This system allowed to rapidly express different Tet mutants in parallel at low levels in embryonic stem cells, in order to rapidly screen for global changes in their localization and mobility.

Consistent with our observations of endogenously labeled Tet1, a construct of mNeonGreen fused to the full length cDNA of Tet1 localized in a spotty pattern, mostly in low DAPI density areas, largely excluded from nucleoli and dense heterochromatin areas. These data suggested that the cDNA construct of Tet1 was behaving similarly to the endogenously labeled Tet1 protein described used in previous sections.



HaloTet1

3D perspective



Deconvolution and 3D reconstruction

3D single-molecule trajectories

Figure 3.17 A. Left, example image acquired in a single snapshot of a cell expressing HaloTet2 after continuous acquisition for 20 seconds, leading to spatially separate single molecules. B, deconvolution and 3D reconstruction from simultaneously acquired focal planes, shown in 3D perspective view rendered in Imaris. C, End-on view of 3D reconstruction of images shown in A. D, Visualization of 3D trajectories of single-molecules imaged on the 3D acMFM. Each field of view has a width of 20.4 μm, and the volume reconstruction has dimensions of 20.4 x 20.4 x 3.6 μm.

It had been suggested that the N-terminal domain was critical for genomic localization of Tet1, given the euchromatin enriched binding profiles found by ChIP-Seq studies, in comparison to the heterochromatin enriched localization of the Tet1 catalytic domain (Zhang et al, 2017). In order to visualize differences in localization and dynamics in living ESCs, we inserted cDNA fusion constructs of mNeonGreen with the N-terminal domain of Tet1 (amino acids 1-1363) or the catalytic domain of Tet1 (amino acids 1364 - 2008). Interestingly, the N-terminal domain of Tet1 had a localization pattern visually similar to the full length Tet1 construct, with a spotty appearance scattered throughout mostly euchromatin, or low DNA compaction areas. Strikingly, the catalytic domain of Tet1 displayed a remarkably different localization. Tet1 CD appeared as a diffuse continuous signal throughout the volume

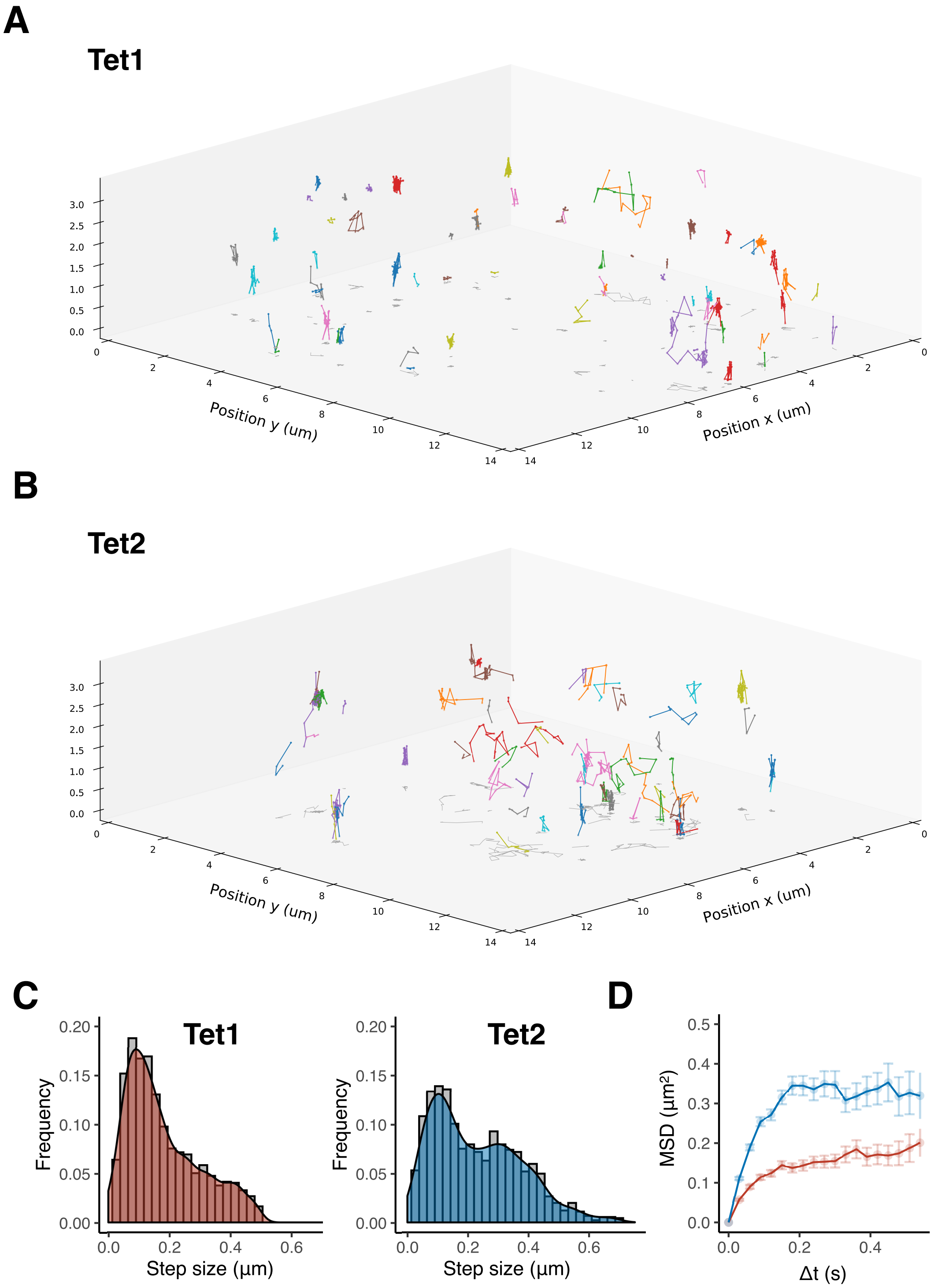


Figure 3.18 A, Example 3D trajectories of HaloTet1 in interphase cells, acquired of the 3D acMFM. B, example trajectories of HaloTet2 in interphase cells. C, distribution of sizes of single sequential steps from Tet1 (left) and Tet2 (right) trajectories. D, mean squared displacement analysis of Tet1 (red) and Tet2 (blue), from tracks acquired and analysed in 3D.

of the nucleus and nucleolus, and strikingly accumulated at heterochromatin-rich

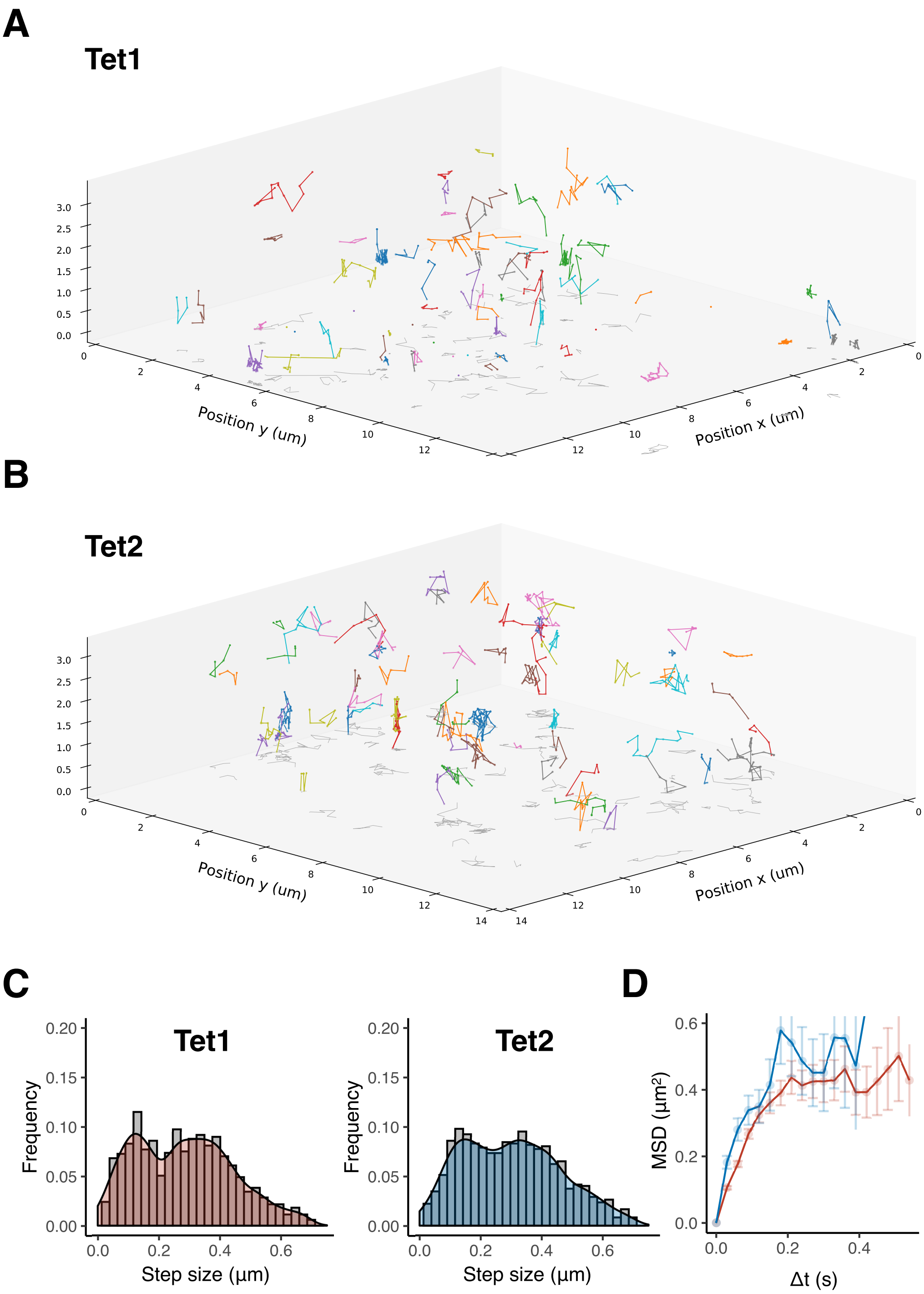


Figure 3.19 A, Example 3D trajectories of HaloTet1 in mitotic cells, acquired of the 3D acMFM. B, example trajectories of HaloTet2 in mitotic cells. C, distribution of sizes of single sequential steps from Tet1 (left) and Tet2 (right) trajectories. D, mean squared displacement analysis of Tet1 (red) and Tet2 (blue), from tracks acquired and analysed in 3D.

chromocenters. These data show for the first time that the N-terminal domain of Tet1 constrains the localization of Tet1 to mostly euchromatin loci, and hinders its entrance into both the nucleolus and heterochromatin areas.

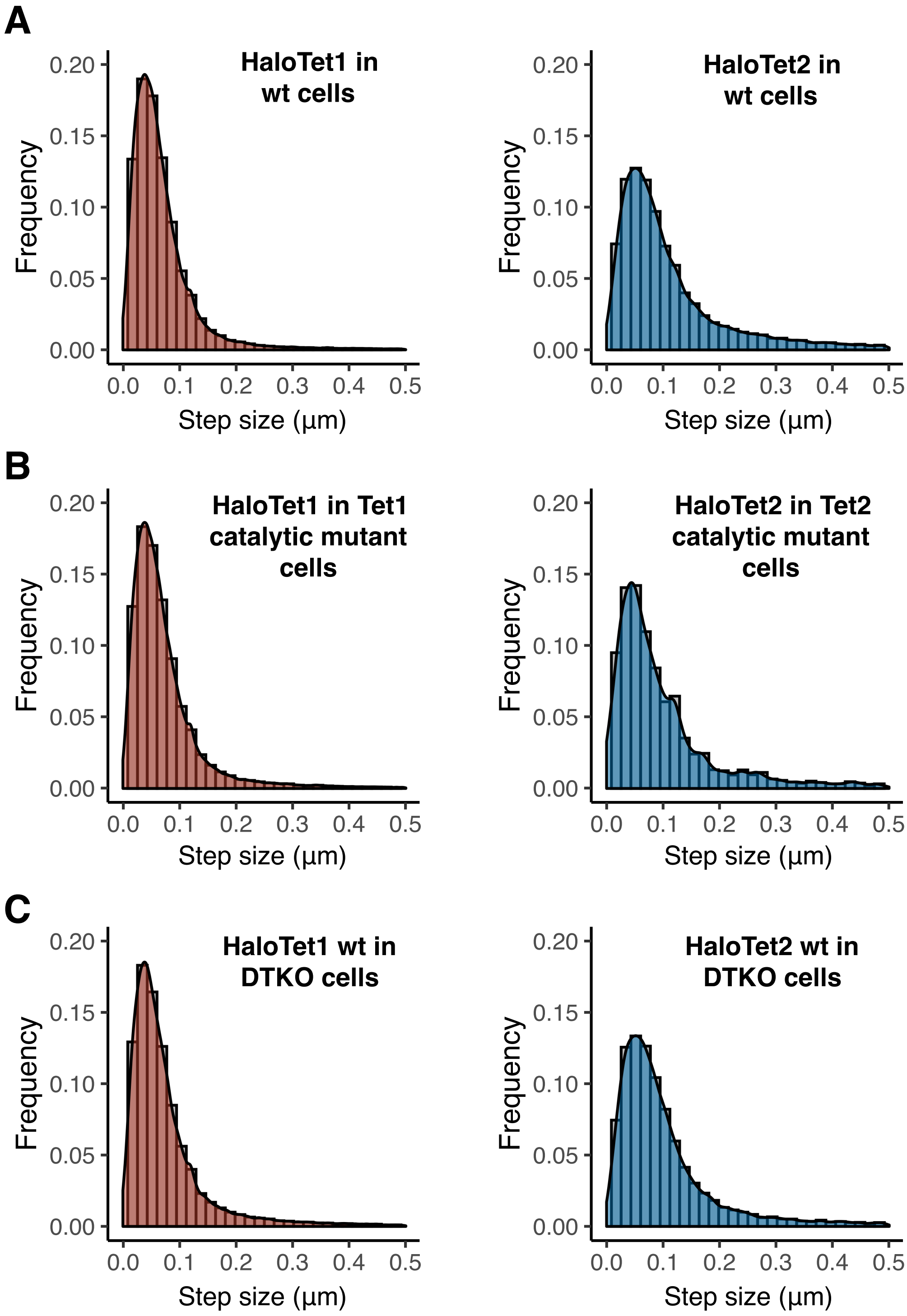


Figure 3.20 A, Distribution of step sizes of HaloTet1 (left) and HaloTet2 in epiblast-like cells, in wt J1 ESCs, B, in cells where Tet1 or Tet2 was rendered catalytically inactive, or C in DTKO cells.

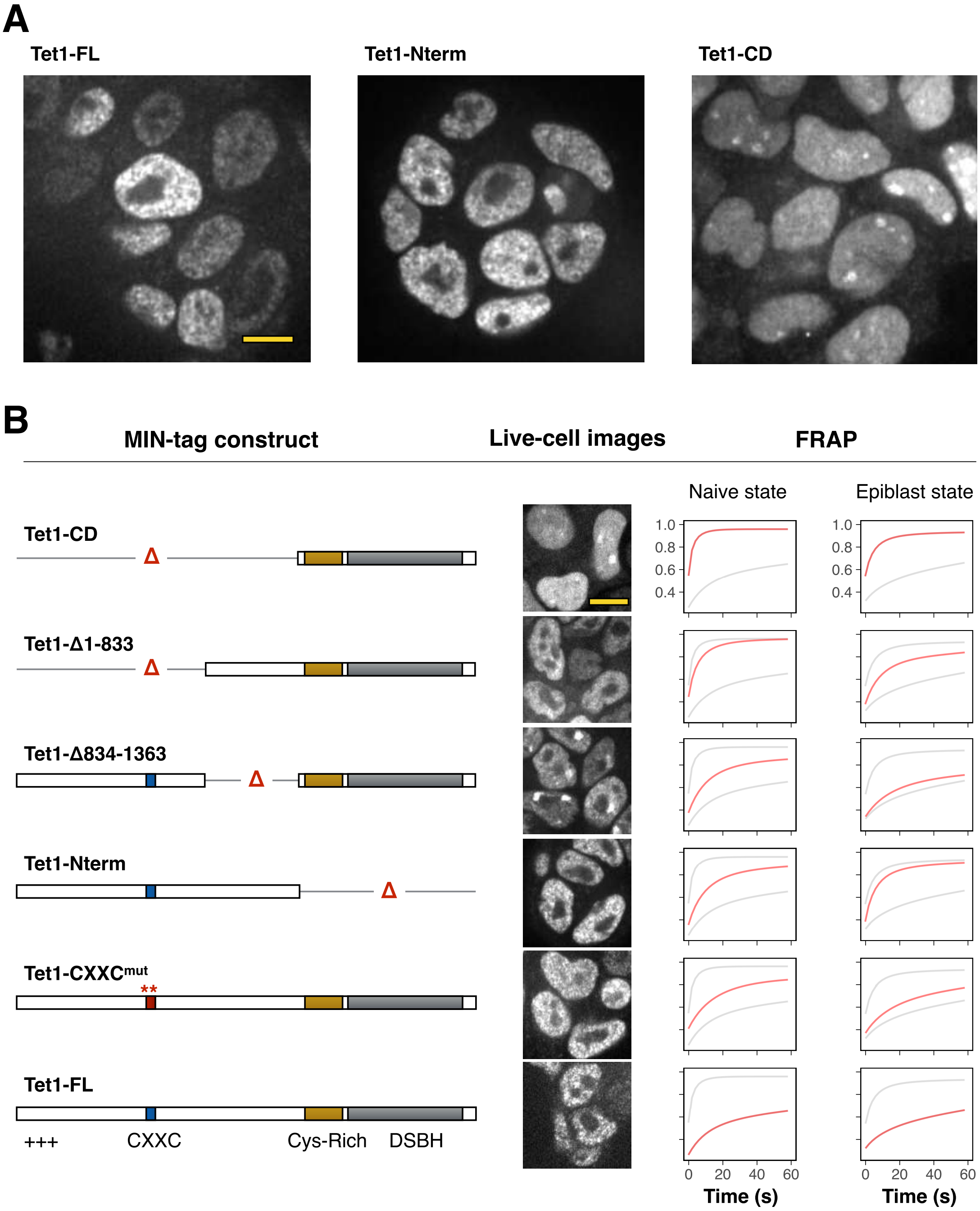


Figure 3.21 Overview of Tet1 constructs generated with the MIN-tag strategy (left), example images of cells expressing mNeonGreen fusion constructs, and FRAP recovery fit (red) in naive and epiblast-like cells, with Tet1CD and Tet1-FL in grey as reference curves. Scale bars = 10 µm.

We then examined other deletion mutants of Tet1, by cloning fragments, which were available in the plasmid database of the Leonhardt group, into MIN-tag compatible vectors. A construct removing the first 833 amino acids (thus removing the CXXC domain), Tet1 Δ1-833, showed a diffuse distribution compared to the full length protein, yet was excluded from nucleoli and heterochromatin. Curiously, a construct in which an internal fragment was removed, Tet1 Δ834-1363, localized both to euchromatin and heterochromatin, while being excluded from nucleoli. Finally, a construct with a mutated CXXC domain had an overall appearance similar to full length Tet1 (Figure 3.21).

These qualitative observations, namely the spotty appearance, nucleolus exclusion and heterochromatin exclusion typical of the full length protein hinted at differences in binding and mobility of these deletion mutants. To measure this, we performed Fluorescence Recovery After Photobleaching (FRAP) on cells expressing these Tet1 constructs, each fused to mNeonGreen. As shown in Figure 3.21 and in greater detail in Figure 3.22, in naive pluripotent stem cells, consistent with their contrasting localization, the catalytic domain of Tet1 showed the fastest recovery. Then, Tet1 Δ1-833 showed the second fastest recovery curve. Then, Tet1 Δ834-1363, the Tet1 N-terminal domain, and Tet1 CXXCmut showed similar recovery kinetics. Finally the full length Tet1 protein showed the slowest recovery kinetics.

These differences in localization and mobility between the N-terminal and catalytic domains directly demonstrate that the N-terminal domain plays a role in enriching Tet1 in low DNA compaction regions of the nuclear volume, and driving chromatin association in this volume of the nucleus (the N-terminal domain modulates the 3D volume in which Tet1 can bind and diffuse). One possible model for this would involve chromatin association primarily via the N-terminal domain, followed by a stabilization of this association via the catalytic domain.

We then pushed the cells to transition from a naive to a primed epiblast state, and performed FRAP on these constructs in this different pluripotency state (Figure 3.21

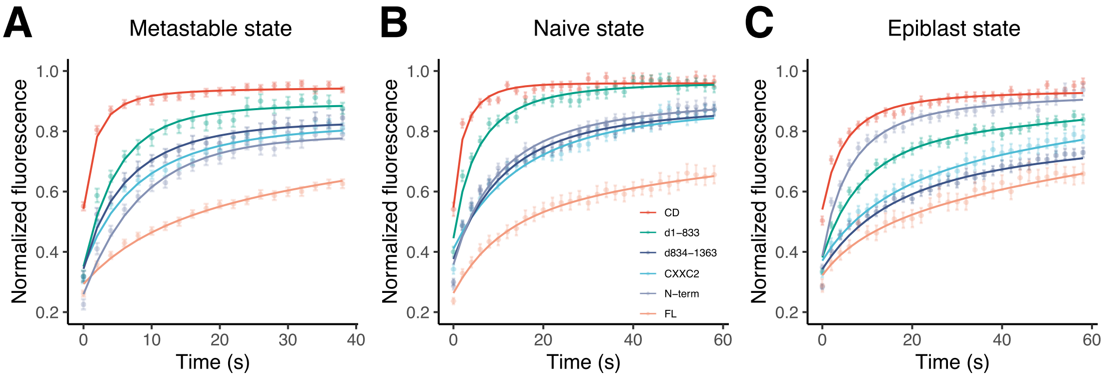


Figure 3.22 FRAP recovery traces of Tet1 mNeonGreen constructs expressed via the MIN-tag system in J1 ESCs maintained in different media conditions. Right panel: metastable cells maintained in serum 2i-LIF; middle panel: naive N2B27 2i LIF media; left panel: cells after 48 hours in differentiation media. Data points are the averaged double-normalized recovery, and lines represent single-component exponential fits. Error bars indicate standard error of the mean. For each construct at each condition, n > 25 cells.

and Figure 3.22). At the epiblast state, the recovery trace of full length Tet1 remained mostly unchanged. However, we observed a lower recovery rate for all deletion mutants which still harboured the catalytic domain,: Tet1CD, Tet1 Δ1-833, Tet1 Δ834-1363, Tet1 CXXCmut. In contrast, the mobility of the N-terminal domain increased following the transition to the epiblast state. The N-terminal domain is the only construct we expressed which does not harbour the catalytic domain of Tet1. Thus, several domains may contribute unevenly to chromatin binding, in a manner that depends on the pluripotency and genomic state of the cells.

These results suggested that Tet1 may have multiple chromatin-association domains, including some yet unknown sites. On the one hand, it is known that the catalytic domain can bind DNA, since a binding event occurs during the catalytic reaction, stabilized by a cysteine-rich domain (Hu et al 2013), and the deletion of the catalytic domain lead to increased mobility. The CXXC appeared to contribute to Tet1 chromatin association, as indicated by the faster recovery of the CXXC-mutated construct. The deletion of an internal, Tet1 Δ834-1363, lead to a strange hybrid localization, where Tet1 was found in a spotty appearance both at heterochromatin loci and euchromatin loci, thus how that domain contributes enriches euchromatin localization is yet unknown. Finally, the construct where the first 833 amino acids were deleted showed the second highest mobility, faster than simply mutating the CXXC domain. This hinted at another chromatin- or DNA-binding site near the N-terminus of Tet1.

This idea was further explored by generating cells expressing smaller deletion constructs. It was shown that a deletion of the first 384 amino acids recapitulated the mobility of the deletion of the first 833 amino acids (Katharina Brandstetter, Mastersarbeit 2017).

Thus, different domains of Tet1 contributed differently to chromatin (or other protein) binding. During naive-to-epiblast differentiation, changes in kinetics are unequal between different mutants, suggesting a developmental change in targeting modalities between these two states.

3.2.2 Deletion of a short positively-charged domain of Tet1

Cell line generation and imaging was carried out with the help of Evi Ntouliou.

Our data provides direct evidence for an uncoupling of enzymatic activity with localization and dynamics, which is consistent with a hypothesis that the N-terminal domain of Tet1 drives the localization of the protein. What's more, FRAP data on Tet1

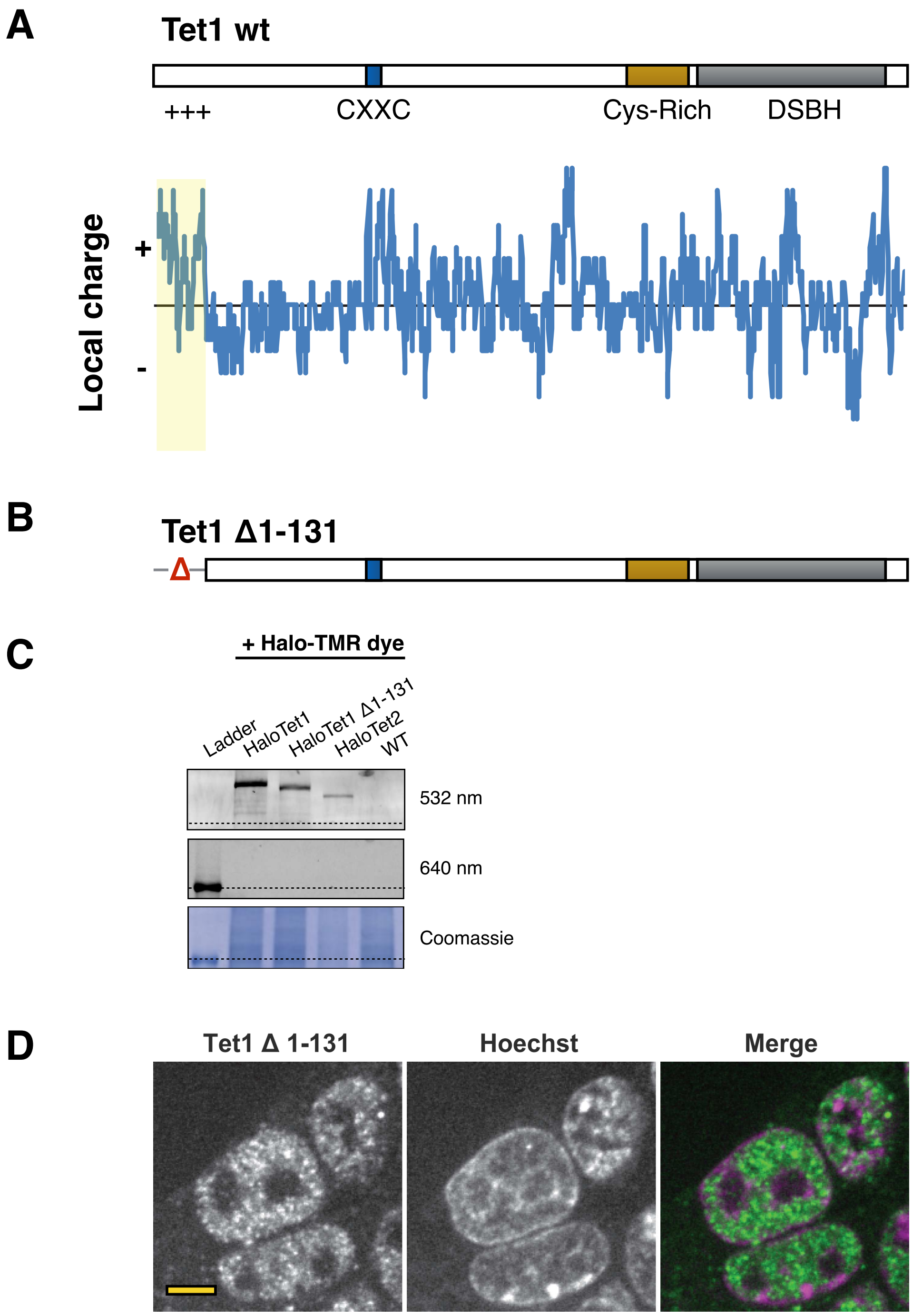


Figure 3.23 A, schematic overview of known domains in Tet1 (top), and average charge of local amino acids (bottom). B, schematic of Tet1 Δ1-131. C, in-gel fluorescence assay on lysates from Halo-TMR labeled HaloTet1, HaloTet1 Δ1-131, HaloTet2 and wt cells. Dotted line indicates 170 kDa band. D, example images of cells in which Cas9 was used to insert Halo or SNAP tags at the start codon of Tet1 and remove the genomic sequence encoding the first 131 amino acids of Tet1. Scale bar = 5 μm.

deletion constructs hinted at the existence of a domain strongly contributing to DNA binding, in addition to the CXXC domain. We thus examined the N-terminal domain of Tet1 for chromatin-binding domains. In addition to a DNA-binding CXXC zinc-finger domain, a conserved positively-charged amino acid sequence was recently identified, and bulk assays suggested this domain might enhance non-specific binding of Tet1 to chromatin (FIGURE 3.23 A).

To test this in living ESCs, we generated a cell line wherein the endogenous coding sequence of the first 131 amino acids of Tet1 was deleted (FIGURE 3.23 B). To this end, we used the same gRNA as used for cell line generation in section 2.1.1, but with a donor template containing the Halo or SNAP tag coding sequence in frame with the coding sequence of Tet1 starting at amino acid 132. Recombination with this donor template would lead to the deletion of the coding sequence of the first 131 amino acids. In parallel, we generated wild-type HaloTet1 and HaloTet2 cell lines in the same background, to control for passage number and clonal variation. We verified that Tet1, Tet1 Δ1-131 and Tet2 were labeled with the HaloTag by labeling cells with Halo-TMR, and running their lysate on a poly-acrylamide gel, which yielded fluorescent bands at their expected sizes ~250 kDa (FIGURE 3.23 C).

3.2.3 The 1-131 domain of Tet1 enhances its on-rate

We then performed live-cell imaging to look at the spatial distribution of Tet1 Δ1-131 in ESCs. While Tet1 Δ1-131 cells appeared visually similar to Tet1 wt, forming visible clusters in low DNA density areas, the coefficient of variation of Tet1 Δ1-131 was significantly lower than that of Tet1, though higher than that of Tet2 (FIGURE 3.24). This suggests that Tet1 Δ1-131 may have a larger freely diffusing fraction than Tet1, while maintaining a larger bound fraction than Tet2.

We looked at Tet1 Δ1-131 localization during mitosis to look at its behaviour at longer timescales. Strikingly, Tet1 Δ1-131 appeared depleted from mitotic chromosomes compared to Tet1 (FIGURE 3.25 A). Quantitative analysis of mitotic chromosome enrichment showed that Tet1 Δ1-131 is less enriched than Tet1, and more enriched than Tet2. This indicates that Tet1 Δ1-131 has a lower on-rate but potentially similar dwell-time than wt Tet1. We then verified whether the positively charged fragment was sufficient to drive mitotic chromosome binding on its own. We overexpressed this short fragment (Tet1 1-131) and we saw that it was clearly bound to mitotic chromosomes (FIGURE 3.25 B).

In order to look at the global mobility of Tet1 Δ1-131 we then performed FRAP on cells expressing Halo-labeled Tet1, Tet1 Δ1-131, and Tet2 (FIGURE 3.26 A). We found a higher recovery for Tet1 Δ1-131 compared to Tet1 both at the initial timepoint and after one minute of recovery. The recovery of Tet1 Δ1-131 was slower than that of Tet2, indicating that chromatin binding was only partially impaired in this deletion mutant. These data suggest that Tet1 Δ1-131 has a larger diffuse fraction than Tet1, although it is unclear whether this would be due to a lower on-rate, or due to a shorter residence time.

To investigate the residence time of Tet1 Δ1-131, we performed single-molecule tracking with low laser intensities, and long exposure time, as performed in section 2.1.3, using track duration as a readout of residence time. We calculated the survival probability, and these measurements revealed that the distribution of residence times for Tet1 Δ1-131 are similar to that of Tet1, indicating that if Tet1 Δ1-131 binds to chromatin, it is likely to bind for durations long as Tet1 (FIGURE 3.26 B). Taken together, these data suggest that Tet1 Δ1-131 can stably bind chromatin loci for tens of seconds. However, the loss of this positively-charged domain appears to decrease the frequency of binding events, thus the on-rate of the protein.

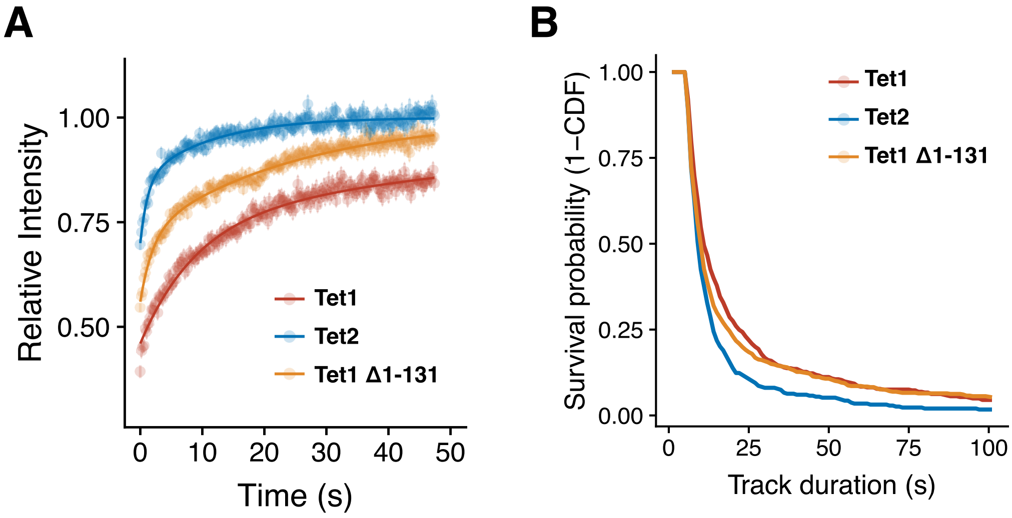


Figure 3.26 A, Normalized average fluorescence intensity traces of FRAP experiment carried out on all three cell lines, n > 60 cells per condition. B, survival probability plot of HaloTet1, HaloTet1 Δ1-131, and HaloTet2, as measured by single-molecule tracking with slow acquisition rates.

To investigate the diffusive properties of Tet1 Δ1-131 as faster timescales, we performed single-molecule tracking with shorter exposure times. We performed sparse labeling with JF646, and bulk labeling with JF549. Here, we used three different exposure times, to make sure large steps and subtle kinetics weren't missed, and that data isn't obscured at certain time intervals. We performed single-molecule tracking on Tet1, Tet2, and Tet1 Δ1-131, with exposure times of 8, 18 and 30 ms, leading to time intervals between images of 12.5, 22.5 and 33 ms, respectively.

To make sure images acquired with 8 ms exposure times could capture fast dynamics, and that diffusive molecules were not moving too quickly to be detected, we acquired images from cells expressing HaloTet1, HaloTet2, and HaloTet1 Δ1-131 undergoing mitosis. We restricted the area of tracking analysis to the cytosol, ignoring areas corresponding to condensed chromosomes, thereby enriching for molecules freely diffusing in the cytosol. We found that we could detect visibly diffusive molecular trajectories at this acquisition rate in mitotic cells (FIGURE 3.27 A). We then measured the mean squared displacement, which revealed steadily increase squared displacements at sequential time intervals for all three Tet species measured (FIGURE 3.27 B). We then used SpotON analysis software to fit a kinetic model to these data, which revealed very low bound fractions near 5% for HaloTet1, HaloTet2, and HaloTet1 Δ1-131 (FIGURE 3.27 C), as well as diffusion rates of the diffusive fraction in the range of 2-3 μm^2/s (FIGURE 3.27 D). This model was fit using the distributions of displacements at different time intervals, shown in Figure 3.27 E.

Having shown that our imaging system is capable of detecting rapidly diffusing particles, we then measured the binding and diffusive properties of HaloTet1, HaloTet2 and HaloTet1 Δ1-131 in the nucleus of cells in interphase. Example trajectories measured from interphase cells are shown in Figure 3.28 A, generally

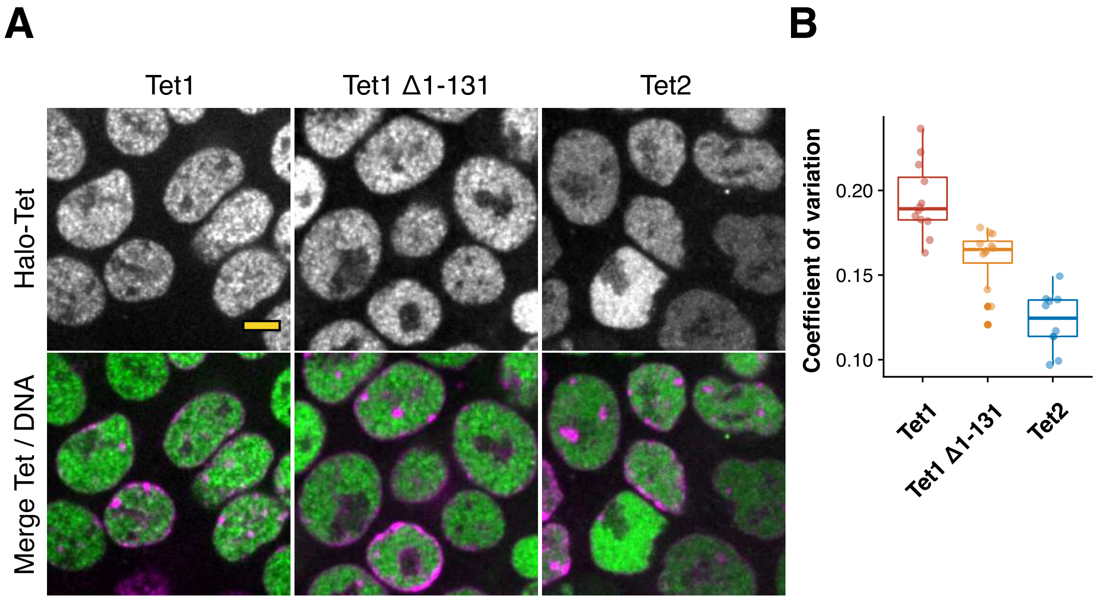


Figure 3.24 A, examples of live ESCs expressing HaloTet1, HaloTet1 Δ1-131, or HaloTet2, generated from the same culture of wild-type ESCs. B, coefficient of variation of each cell type. Scale bar = 5 μm.

reveal a smaller fraction of diffusive trajectories compared to mitotic cells (shown in FIGURE 3.27 A). Mean squared displacement analysis revealed a more constrained diffusion compared to mitotic cells (FIGURE 3.28 B). Furthermore, as shown in other mean squared displacement analyses in this work, Tet2 showed the largest displacement at sequential intervals compared to Tet1, and Tet1 Δ1-131 showed

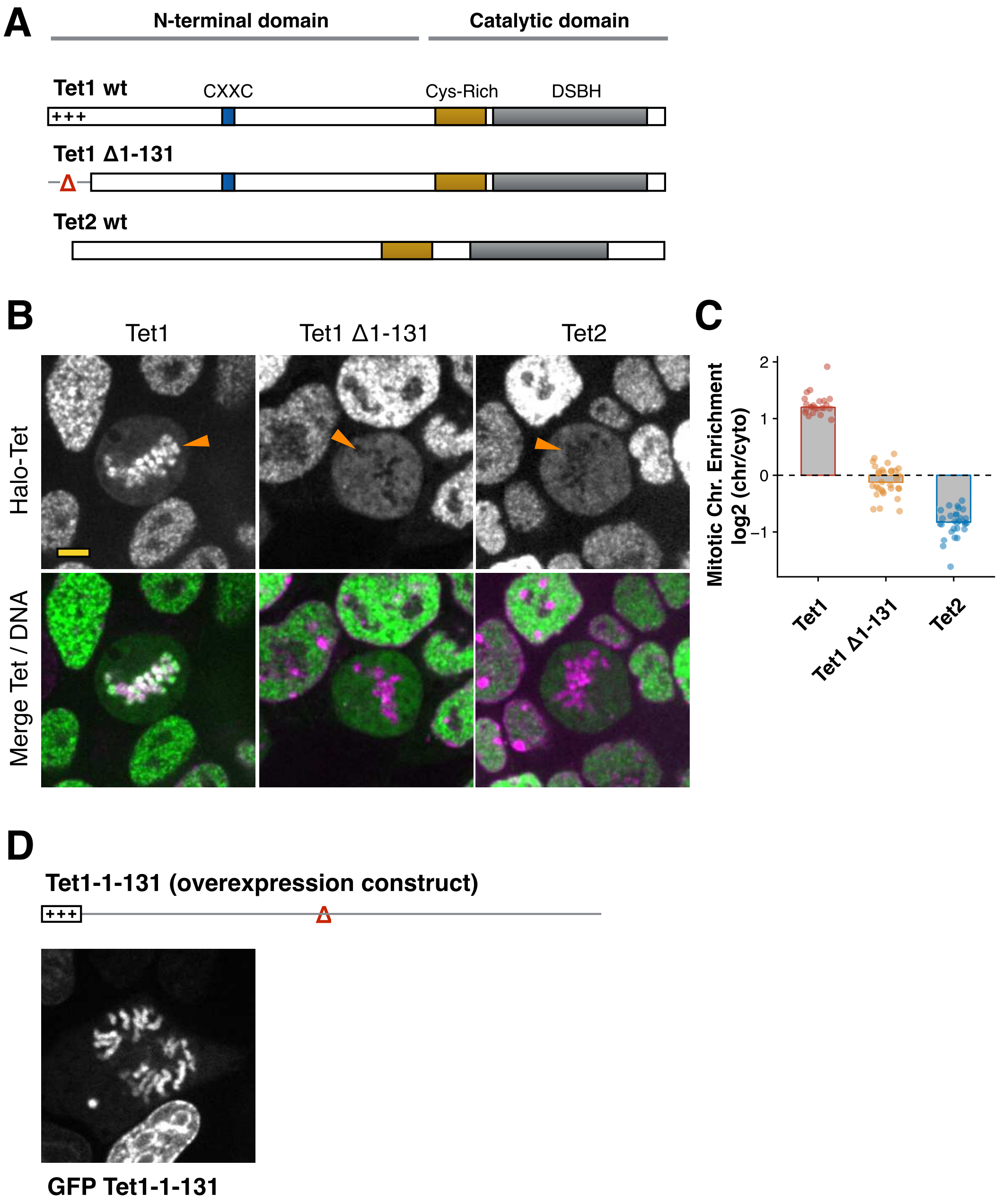


Figure 3.25 A, Protein schematics of Tet1, Tet1 Δ1-131, Tet2. B, examples of live ESCs in mitosis expressing HaloTet1, HaloTet1 Δ1-131, or HaloTet2, arrow heads point to location of mitotic chromosomes. C, mitotic chromosome enrichment from all three cell types. D, mitotic chromosome localization of the short positively charged fragment Tet1 1-131. Scale bars = 5 μm.

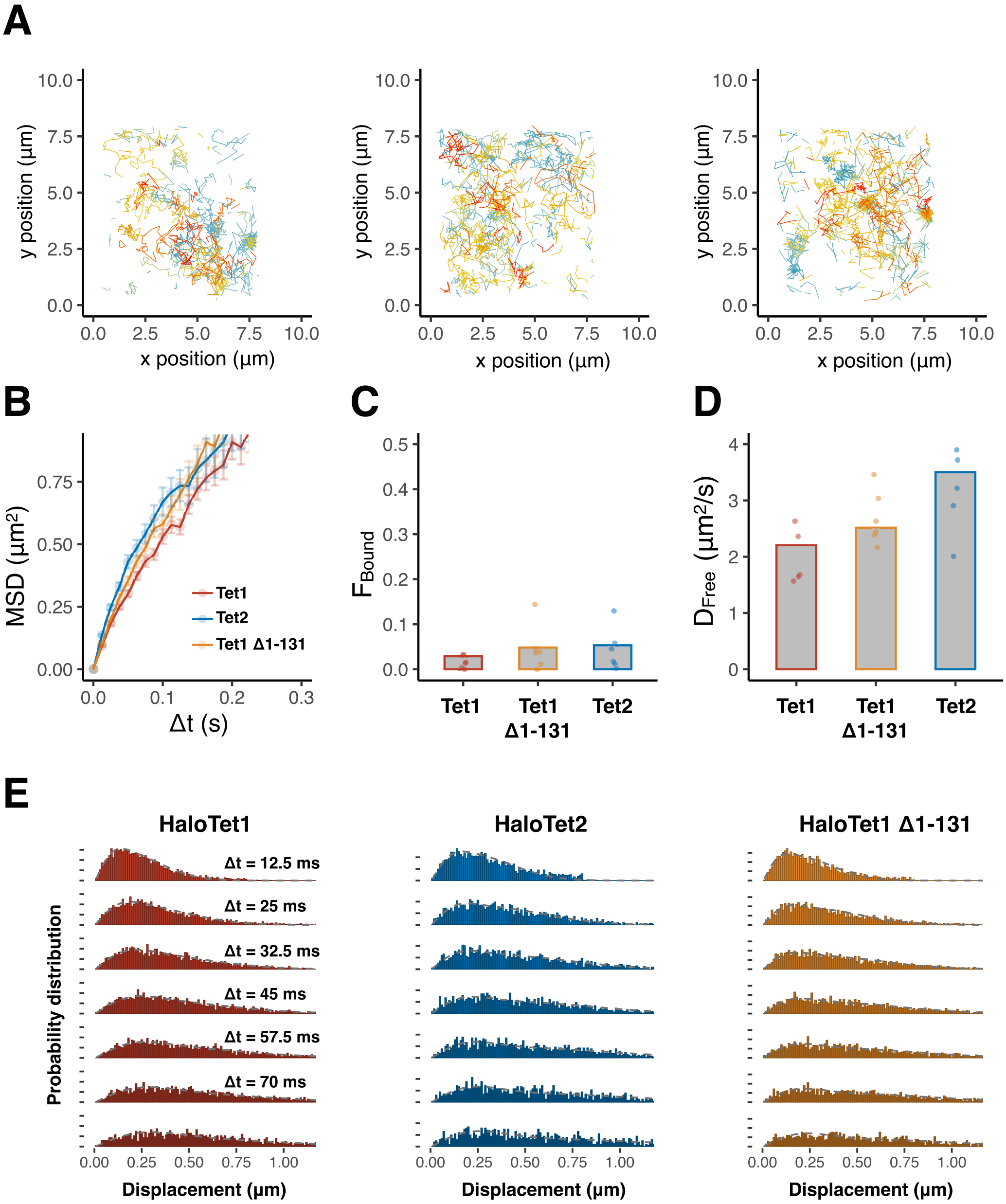


Figure 3.27 A, single-molecule trajectories of HaloTet1 (left), HaloTet1 Δ1-131 (middle), and HaloTet2 (right), in mitotic cells, extracted from time-lapse images with 8ms exposure time. B, mean squared displacement analysis of trajectories. C, bound fraction of molecules following three-state kinetic model fit in SpotON. D, diffusion rate of diffusive fraction, obtained from fitting. E, distribution of displacements at different time intervals used for kinetic model fitting.

slightly larger displacements compared to Tet1. A kinetic model calculated with SpotON revealed again a greater bound fraction for Tet1 compared to Tet2, and Tet1 Δ1-131 having a bound fraction larger than Tet2 but lesser than Tet2, consistent with

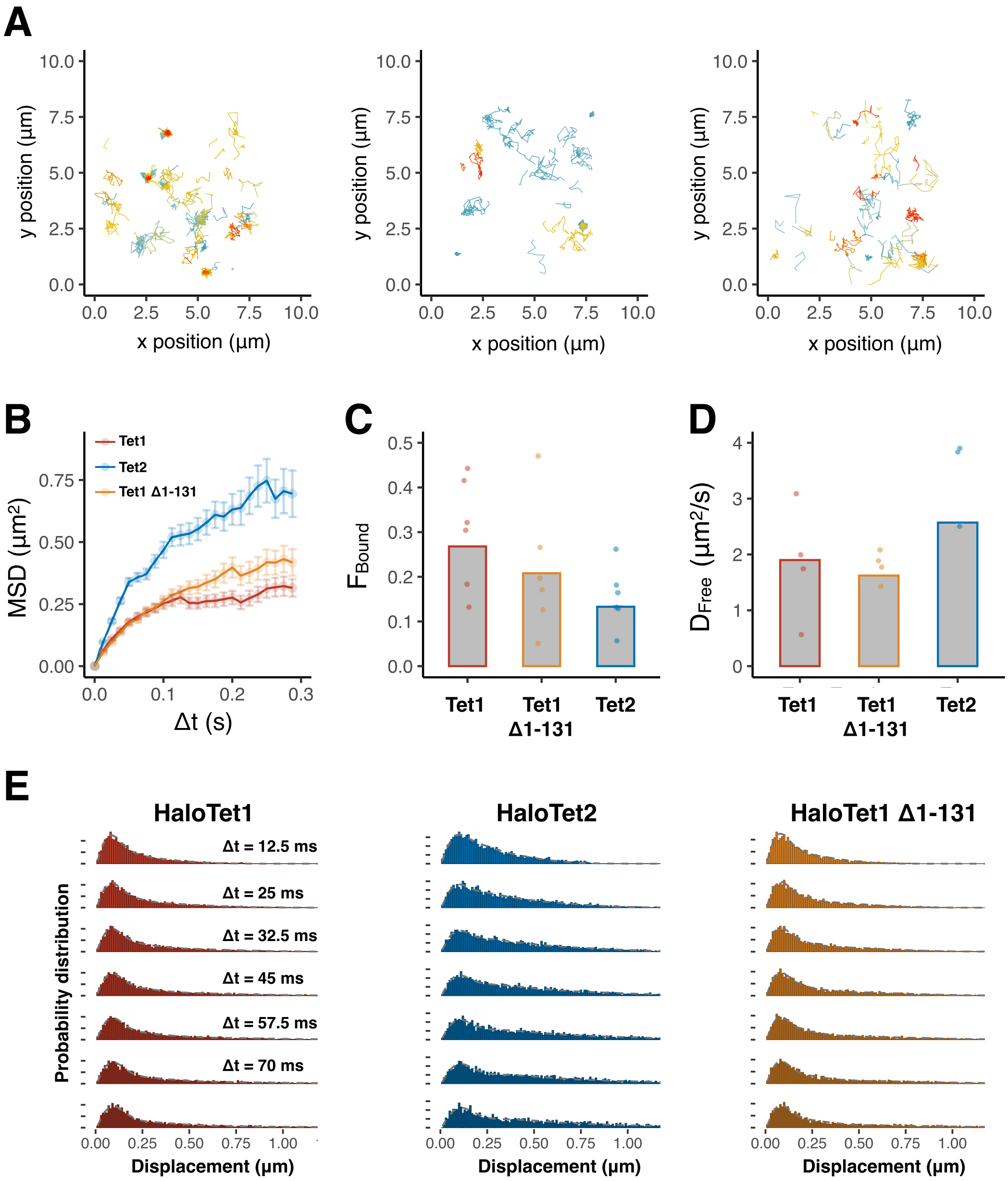


Figure 3.28 A, single-molecule trajectories of HaloTet1 (left), HaloTet1 Δ1-131 (middle), and HaloTet2 (right), in the nucleus of interphase cells, extracted from time-lapse images with 8ms exposure time. B, mean squared displacement analysis of trajectories. C, bound fraction of molecules following three-state kinetic model fit in SpotON. D, diffusion rate of diffusive fraction, obtained from fitting. E, distribution of displacements at different time intervals used for kinetic model fitting.

Tet1 Δ1-131 being more mobile than Tet1 (FIGURE 3.28 C). Diffusive rates calculated were in the range of 1.5 to 3 µm²/s (FIGURE 3.28 D). The distribution of displacements at different time intervals used to calculate the diffusive model are shown in Figure 3.28 E.

Since acquisition at 8 ms exposure times required illuminating the cells with the maximal laser intensity available on the system (~ 25 mW), we performed SMT at 18 and 30 ms exposure times using lower laser intensity (~ 12 mW). While decreasing the laser power, the increase in exposure lead to higher total illumination power per exposure, thus leading to a higher signal-to-noise and lower photobleaching rate than our 8ms exposure acquisition setup, at the cost of a lower temporal resolution. Both 18 and 30 ms revealed binding and diffusive behaviour similar to what we observed with 8 ms exposure times. Mean squared displacement analysis showed that Tet1 Δ1-131 had a larger displacement than Tet1, but less than Tet2 (FIGURE 3.29 A and D). SpotON analysis revealed that Tet1 Δ1-131 had a slightly smaller bound fraction than Tet1, but still higher than Tet2 (FIGURE 3.29 B and E), and diffusive rates between 1 and 2 $\mu m^2/s$ (FIGURE 3.29 C and F), suggesting that some rapidly diffusing particles are not detected in these acquisition setups. Displacement distributions at different time intervals are shown for 18 and 30 ms exposure time acquisitions in Figure 3.29 G and H, respectively.

Thus, at three different exposure times and laser intensity settings, these measurements consistently revealed that Tet1 Δ1-131 has a larger diffusive fraction and a smaller bound fraction than Tet1. In parallel, the diffusive fraction of Tet2 was consistently larger than both Tet1 and Tet1 Δ1-131. Taken together, given the similar residence times of Tet1 and Tet1 Δ1-131, their similar recovery half-time in FRAP, and the larger diffuse fraction of Tet1 Δ1-131 as measured by single-molecule tracking, we conclude that the positive stretch of amino acids increases the on-rate of Tet1, thereby ensuring frequent bouts of binding to chromatin loci. These data suggest that a high on-rate, capable of maintaining chromosome binding during mitosis, is largely driven by a positively charged stretch of amino acids. The deletion of this domain lead to a larger diffuse fraction and a lower on-rate of Tet1 Δ1-131 compared to wt Tet1. This deletion, however, did not lead to kinetics as diffusive as Tet2, suggesting that other mechanisms, such as the CXXC domain, likely contribute to chromatin binding of Tet1. Since the CXXC domain has been shown to target unmethylated CpG islands (Xu et al, 2018), we propose that the positively charged domain of Tet1 drives transient non-specific binding, and the CXXC domain, and likely other uncharacterized domains, drive specific binding to target promoters.

2.3 Tet protein target DPPA3 impairs maintenance methylation by interacting with UHRF1

2.3.1 DPPA3 localization and dynamics in live-cells

The genomes of mammals such as rodents and primates undergo an extensive wave of DNA demethylation during embryonic development, shortly after fertilization (Reik et al, 2001). Curiously, other animals, such as boreoutherian mammals, as well as fish and amphibians, do not undergo such waves of DNA demethylation, despite having similar conserved DNA methylation machinery as well as Tet proteins. As a result, it is unclear what drives this wave of DNA demethylation in rodents and primates, since most of the machinery known to generate and modify methylation exists in animals that do not undergo such demethylation.

The demethylating activity of TET proteins is generally seen as a result of both its catalytic activity. TET catalytic activity, as described earlier, directly modulates methylcytosine, and can lead to passive and active DNA demethylation. However, the non-catalytic activities of TET proteins are not fully understood.

To dissect catalytic and non-catalytic targets of Tet proteins in stem cells as they transition from a naive to a primed state, Christopher Mulholland generated a series of cell lines where Tet1, Tet2 or both Tet1 and Tet2 were mutated, rendering them catalytically inactive. By performing RNA-seq to survey the transcriptome of these cells, as well as reduced representation bisulfite sequencing to profile their methylome, a collection of Tet catalytic activity dependent targets were found. One of these candidates found to be regulated by Tet1 and Tet2 catalytic activity was DPPA3 (also known as Stella). DPPA3 was an interesting candidate for further investigation, since it was shown to be involved with maintenance methylation in the early embryo (Li et al, 2018), and thought to interact with UHRF1, a protein involved maintenance methylation (Funaki et al, 2014, Du et al, 2019).

To this end, further characterization of the behaviour of DPPA3 was required to assess whether and how it could contribute to modulating methylation levels in embryonic stem cells. To gain insight into the cellular localization of DPPA3, the HaloTag was inserted via CRISPr/Cas9 at the endogenous DPPA3 locus in embryonic stem cells. Surprisingly, live-cell imaging revealed a largely cytosolic localization (FIGURE 3.30), which was unexpected for a protein thought of as a transcription factor (Bortvin et al, 2004) which modulates DNA methylation.

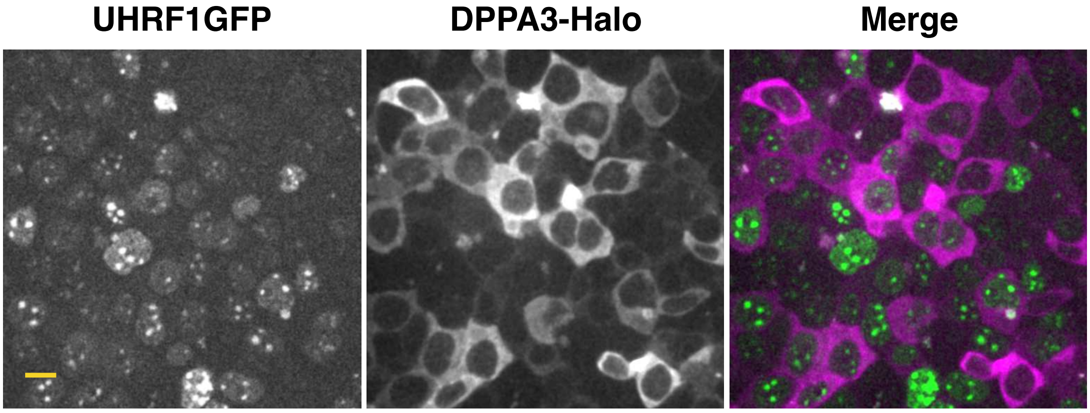


Figure 3.30 Example images of embryonic stem cells expressing UHRF1-GFP and DPPA3-Halo. Left panel shows UHRF1-GFP, middle panel shows DPPA3-Halo, and right panel shows merged image. Scale bar = 10 µm

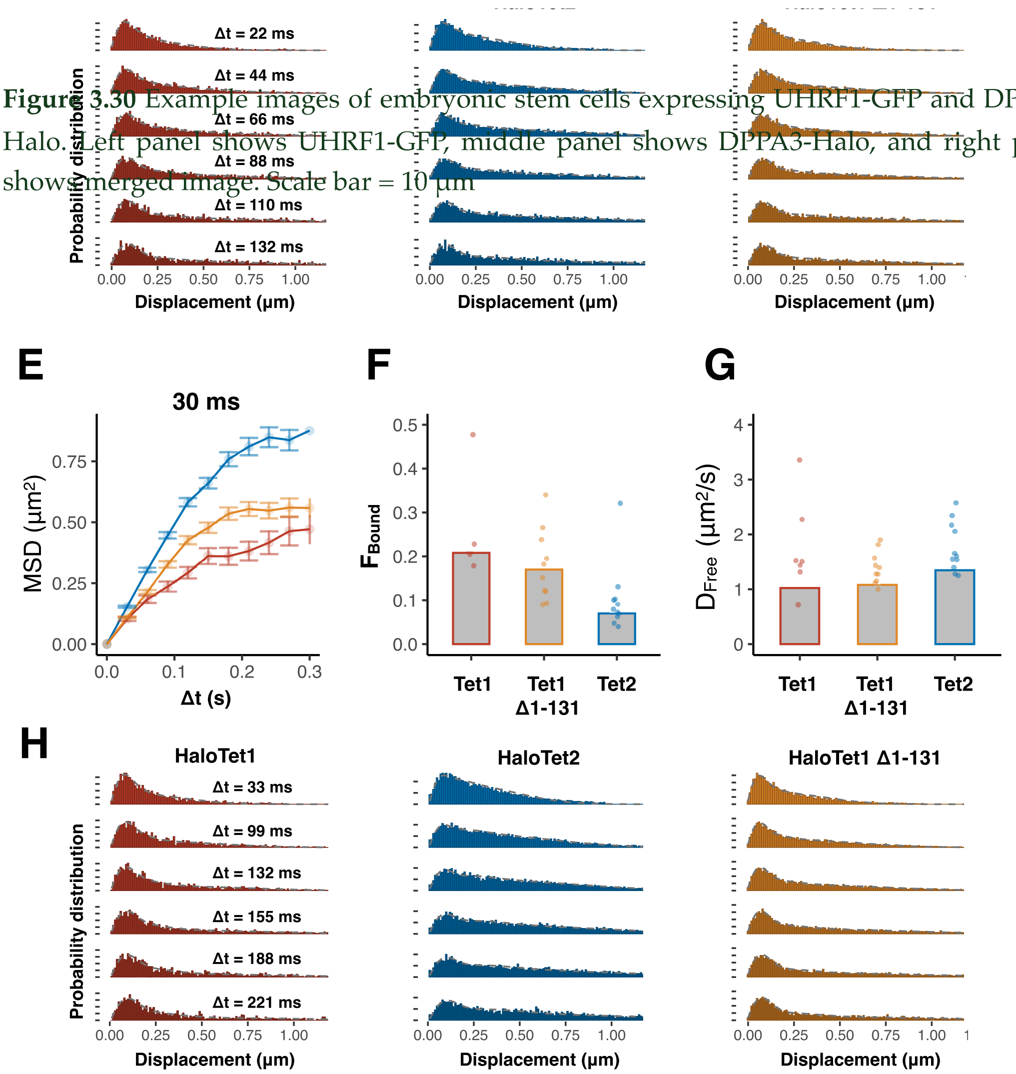


Figure 3.29 A and D mean squared displacement analysis of trajectories of HaloTet1, HaloTet1 Δ1-131, and HaloTet2, in the nucleus of interphase cells, extracted from time-lapse images with 18 ms and 30 ms exposure time. B and E, bound fraction of molecules following three-state kinetic model fit in SpotON. C and F, diffusion rate of diffusive fraction, obtained from fitting. G and H, distribution of displacements at different time intervals used for kinetic model fitting.

2.3.2 UHRF1 localization disrupted by expression of DPPA3

At first glance, this predominant cytosolic localization was at odds with the notion of DPPA3 interacting with UHRF1, a DNA methylation factor usually found in the nucleus. We then sought to characterize how this interaction occurred in ESCs. Cell lines were generated in which UHRF1 was labeled with GFP, endogenous DPPA3 was knocked-out, and fluorescently-labeled DPPA3 was then re-expressed under a doxicycline-inducible promoter. In cells where DPPA3 was knocked-out, UHRF1 was largely found in the nucleus, with a clustered appearance at heterochromatin foci (FIGURE 3.31 A). However, upon doxycycline induction and expression of DPPA3, the distribution and localization of UHRF1 changed dramatically. Qualitatively, after induction of DPPA3, UHRF1 distribution in the nucleus appeared more diffuse, which lead to accumulation in the cytosol.

As a control, we performed the same experiment in UHRF1-GFP / DPPA3 KO cells in which we expressed a mutant of DPPA3 where in vitro interaction with UHRF1 is deficient (DPPA3-KRR). In these cells, we measured the coefficient of variation of UHRF1-GFP (FIGURE 3.31 B). We found that UHRF1-GFP had the most diffuse distribution in cells expressing DPPA3-wt, and the most clustered distribution in cells deficient for DPPA3. Cells expressing DPPA3-KRR had a slightly lower coefficient of variation than DPPA3-KO cells, suggesting some residual interaction between this mutant and UHRF1. Taken together, these data hinted at a strong interaction between DPPA3 and UHRF1. Despite their contrasting localization, we hypothesized that a small fraction of DPPA3 enters the nucleus, interacts with DPPA3, and leads to its export into the cytosol.

We performed FRAP on these induced cells, and found that induction of wt DPPA3, but not the mutant, lead to UHRF1 being more mobile within the nucleus, suggesting that DPPA3's interaction with UHRF1 inhibited UHRF1 binding to chromatin (FIGURE 3.31 C).

2.3.3 DPPA3 forms a diffusive complex with UHRF1 impairing its chromatin binding

RICS results presented in this section were performed with Christopher Mulholland and Ivo Gluck, with technical assistance from Nader Danaf, and analysis advice from Philipp Messer in Prof Don Lamb's group.

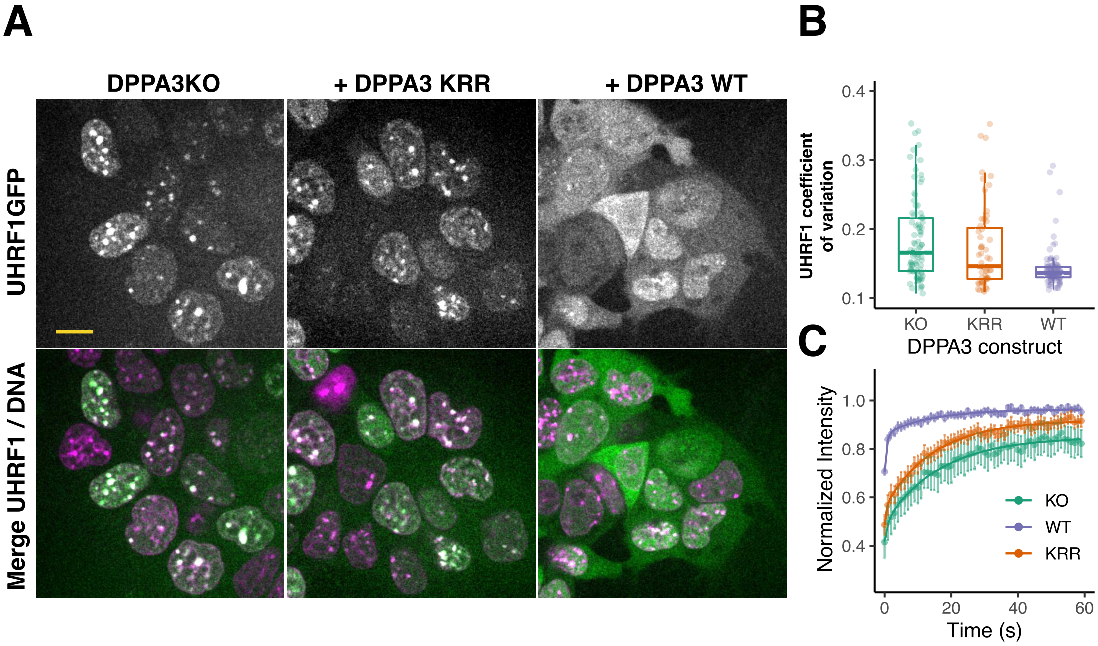


Figure 3.31 A, Example images of embryonic stem cells expressing UHRF1-GFP, in which endogenous DPPA3 is knocked out (left), and in which constructs of DPPA3-KRR (middle) or wild-type DPPA3 (right) are over-expressed under the control of a doxycycline promoter. B, coefficient of variation of UHRF1-GFP measured in DPPA3-KO, DPPA3-KO with re-expression of DPPA3-KRR, and DPPA3-KO with re-expression of wild-type DPPA3. C, FRAP intensity profiles of UHRF1 in these three genetic backgrounds. Scale bar = 10 μm

In vitro biochemical methods had suggested that DPPA3 and UHRF1 interact with high affinity. However, it was unclear if this interaction occurred in live-cells, given the different proteomic environment of the cell compared to a sample of recombinant proteins, and since given the contrasting localization of DPPA3 and UHRF1 in live cells. To test this interaction in living cells using a quantitative approach, we turned to RICS and cross-correlation RICS, using Pulsed Interleaved Excitation (PIE-RICS), in cells expressing both fluorescently labeled DPPA3 and UHRF1.

We first established the global mobility and cross-correlation properties of freely diffusing eGFP and freely diffusing mScarlet, as visible in the thin autocorrelation functions (FIGURE 3.32 A). As expected, both proteins showed a highly diffusive behaviour, with diffusion rates nearing 40 $\mu m^2/s$ and almost total mobile fraction (FIGURE 3.32 B). Finally, we measured the cross-correlation between these two protein, which indicates the likelihood of finding these two proteins diffusing together as a complex (FIGURE 3.32 C). We found virtually no cross-correlation, as expected for freely diffusing independent proteins.

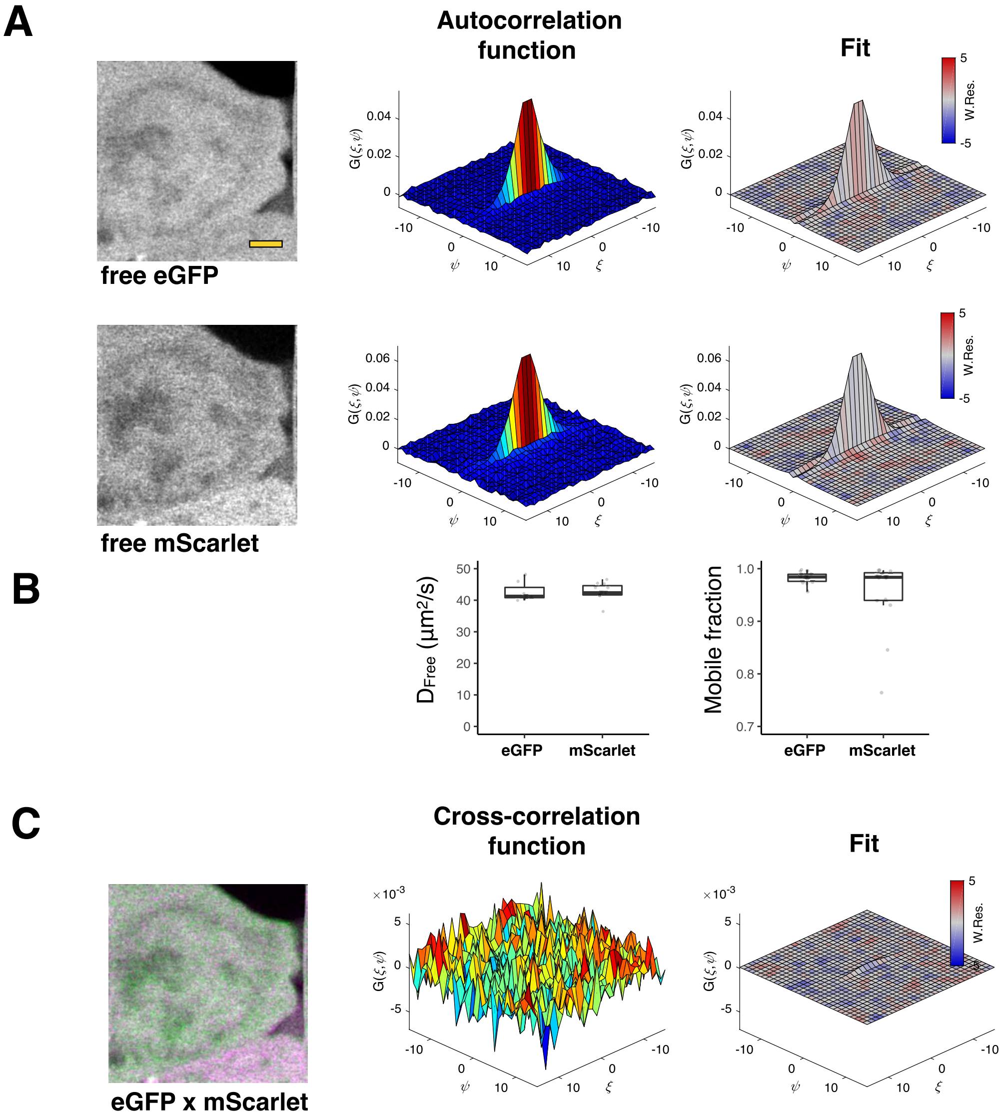


Figure 3.32 Raster Image Correlation Spectroscopy of embryonic stem cells expressing free eGFP as well as free mScarlet, as a control for quickly diffusing independent proteins. A, example images (left), autocorrelation functions (middle) and two-component fits (right) for eGFP (top) and mScarlet (bottom) expressed simultaneously in ESCs. B, diffusion rates and mobile fractions derived from fits. C, merged image of eGFP and mScarlet (left) cross-correlation function of eGFP and mScarlet (middle), and model fit (right). Scale bar = 2 µm

We then measured the binding and diffusive properties of DPPA3 and UHRF1 in live ESCs by performing RICS on cells expressing UHRF1-GFP and inducibly expressed DPPA3-mScarlet, and limiting our analysis to signal in the nucleus. For DPPA3, we looked at the wild-type protein in cells expressing UHRF1-eGFP, the wild-type protein in UHRF1 knock out cells, and a mutant of DPPA3 in the cells expressing UHRF1-eGFP. As controls, we also measured cells expressing a tandem eGFP-mScarlet construct, as well as cells where UHRF1 was knocked out and expressing wild-type DPPA3-mScarlet and freely diffusing eGFP. We calculated the autocorrelation functions of these fluorescently-labeled proteins, and found that UHRF1-GFP had a “thinner” profile in the presence of wild-type DPPA3 compared to DPPA3-KRR, suggesting that UHRF1 is more mobile in the presence of wild-type DPPA3 (FIGURE 3.33 A). The tandem eGFP-mScarlet construct and freely diffusing eGFP showed the thinnest ACF profiles, as expected given their highly mobile nature. Wild-type and mutant DPPA3 showed thin profiles, suggesting high mobility (FIGURE 3.33 B).

We then fit a two-component diffusive model, consisting of a bound fraction and a diffusive fraction. This analysis revealed a model where, following induction of DPPA3-KRR, the mutant lacking UHRF1 interaction in vitro, UHRF1-GFP had mobile fraction of 32.4 ± 10%, suggesting that over the timecourse of acquisition, the majority of UHRF1 is bound to chromatin in the presence of DPPA3-KRR (FIGURE 3.33 C). However, following the induction of wild-type DPPA3, we measured a substantial increase in the mobile fraction of UHRF1-eGFP, nearly doubling to 60.6 ± 13.7%. Freely diffusing fluorescent proteins (eGFP, and tandem eGFP-mScarlet), were almost totally mobile. We measured a diffusion rate of 7.2 ± 3.9 $\mu m^2/s$ for UHRF1 in the presence of wild type DPPA3 (FIGURE 3.33 D). The diffusion rate of UHRF1 in the presence of DPPA3-KRR is difficult to accurately measure, given the large aggregates of UHRF1 at heterochromatin, which dominate the autocorrelation function.

In these same cells, we analysed the mobility of mScarlet DPPA3 variants, as well as freely diffusing mScarlet. This analysis revealed a model in which the majority of wild-type DPPA3-mScarlet (88.4 ± 5.2%) was freely diffusing in the nucleus (FIGURE 3.33 E). We measured a diffusion rate of 7.2 ± 1.9 $\mu m^2/s$ over the timescale of the measurements (FIGURE 3.33 F). The DPPA3-KRR mutant behaved similarly as its wild-type counterpart, with a mostly mobile fraction diffusing at a rate of 8 ± 2.4 $\mu m^2/s$. Then, we measured the diffusive properties of DPPA3 in UHRF1-KO cells. Surprisingly, we found that DPPA3 in the absence of UHRF1 has similar diffusive properties in cells expressing UHRF1, with a diffusion rate of 8.3 ± 2.3 $\mu m^2/s$. Finally, we measured DPPA3-KRR, a mutant of DPPA3 which no longer binds UHRF1 in vitro. Here again, the diffusion coefficient and mobile fraction of DPPA3-KRR was similar to that of DPPA3-WT. These data, derived from living cells at the single-cell level, validated globally weak genome binding profiles derived from large ensemble-averaged methods such as ChIP-Seq. Together, these data show that DPPA3 does not strongly bind chromatin in ESCs, and likely does not compete with UHRF1 for chromatin binding.

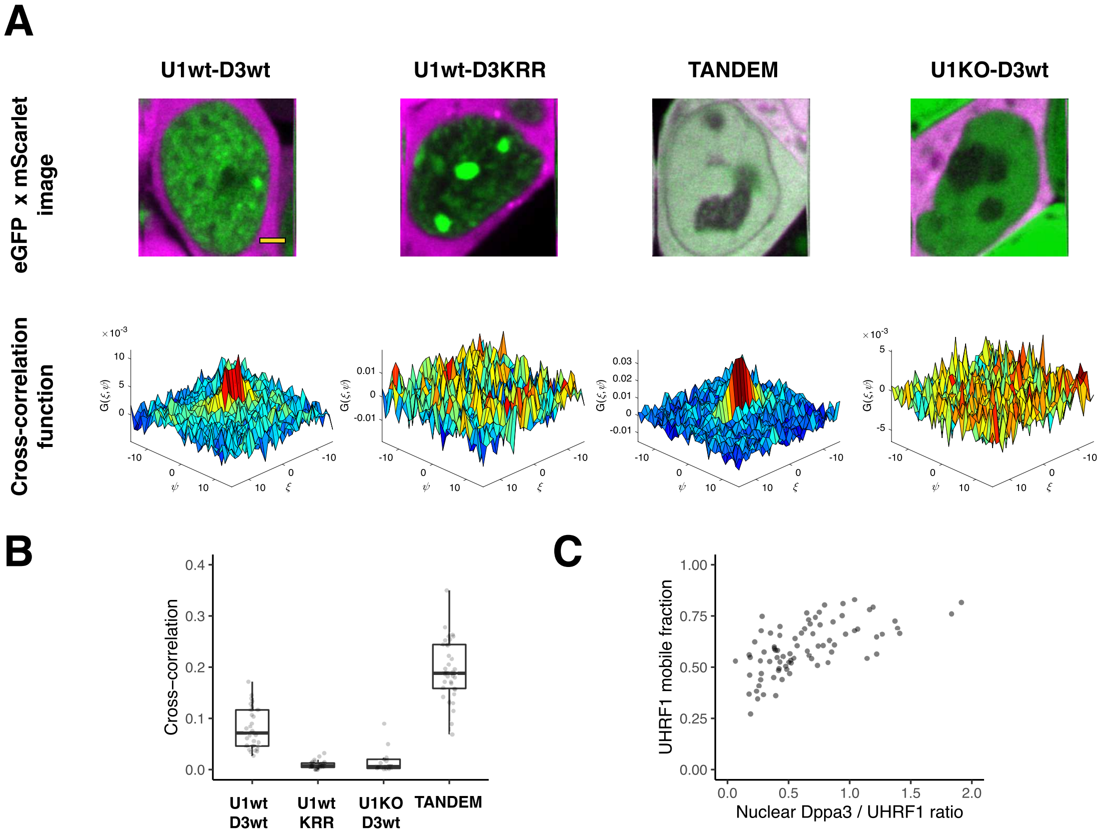


Figure 3.34 Cross-correlation measurements of cells expressing wt UHRF1-GFP and wt DPPA3-mScarlet, wt UHRF1-GFP and DPPA3-KRR-mScarlet, a tandem eGFP-mScarlet fusion construct, or UHRF1 KO cells expressing wt DPPA3-mScarlet. A, example merged images (top), and cross-correlation functions (bottom) of eGFP and mScarlet labeled molecules in each cell type. B, normalized amplitude of cross-correlation of eGFP and mScarlet species in cell type. C, Scatter plot showing the mobile fraction of UHRF1, (derived from two-component diffusive model fitting of the autocorrelation function shown in Figure 3.33 A), over the ratio of the fluorescence intensities of DPPA3 and UHRF1. Each data point corresponds to one cell. Scale bar = 2 μm

Since data was acquired quasi-simultaneously on the PIE-RICS system, we then analysed these datasets for cross-correlation between the mScarlet and eGFP channels, using the cross-correlation amplitude as a readout for potential interaction (FIGURE 3.34 A). We first validated the system's capacity to robustly detect fluorescent molecules in complex in ESCs expressing a tandem construct of mScarlet fused to eGFP, or in ESCs expressing freely diffusing, non-interacting mScarlet and eGFP. As a positive control, in cells expressing a tandem construct of mScarlet fused to eGFP, we measured a strong cross-correlation amplitude between the mScarlet and eGFP channels. As a negative control, cells expressing freely diffusing non-interacting mScarlet and eGFP, showed a low cross-correlation amplitude, indistinguishable from the noise at different spatiotemporal intervals, consistent with proteins which do not robustly interact. These results set the upper and lower limits of the system to detect cross-correlation. We then measured the cross-correlation between DPPA3 and UHRF1. We measured a substantial cross-correlation amplitude between wild-type DPPA3-mScarlet and UHRF1-eGFP, corresponding primarily to the freely diffusing fraction of UHRF1, suggesting that these two proteins form a freely diffusing complex. In striking contrast, we consistently detected low cross-correlation amplitudes between DPPA3-KRR mutant and UHRF1-eGFP, similar to those of freely diffusing, non-interacting fluorescent proteins.

To make sure these cross-correlation amplitudes weren't an artifact of different expression levels, we calculated a cross-correlation amplitude normalized the the amplitudes of individual autocorrelation functions (FIGURE 3.34 B), which again revealed high cross-correlation for the tandem eGFP-mScarlet construct and for UHRF1-GFP / wild-type DPPA3-mScarlet, whereas UHRF1-GFP with DPPA3-KRR showed a low normalized cross-correlation, at similar levels of independent, non-interacting proteins.

These data demonstrated a strong interaction between DPPA3 and UHRF1. We then asked if this interaction scaled with the expression levels of DPPA3 and UHRF1, which would further suggest a direct interaction. We found that the mobile fraction of UHRF1 scaled with the ratio of intensity of DPPA3 to UHRF1, which indicates a ratiometric effect of DPPA3 on the chromatin binding of UHRF1, and hints at a potential physical interaction (FIGURE 3.34 C). All together, these data clearly demonstrate, that in living cells embryonic stem cells, DPPA3 forms a highly mobile complex with UHRF1 and affects its mobility, whereas the DPPA3-KRR mutant fails to bind UHRF1 and inhibit its chromatin binding. Thus we show for the first time a quantitative biophysical framework underlying a recently discovered pathway for DNA demethylation.

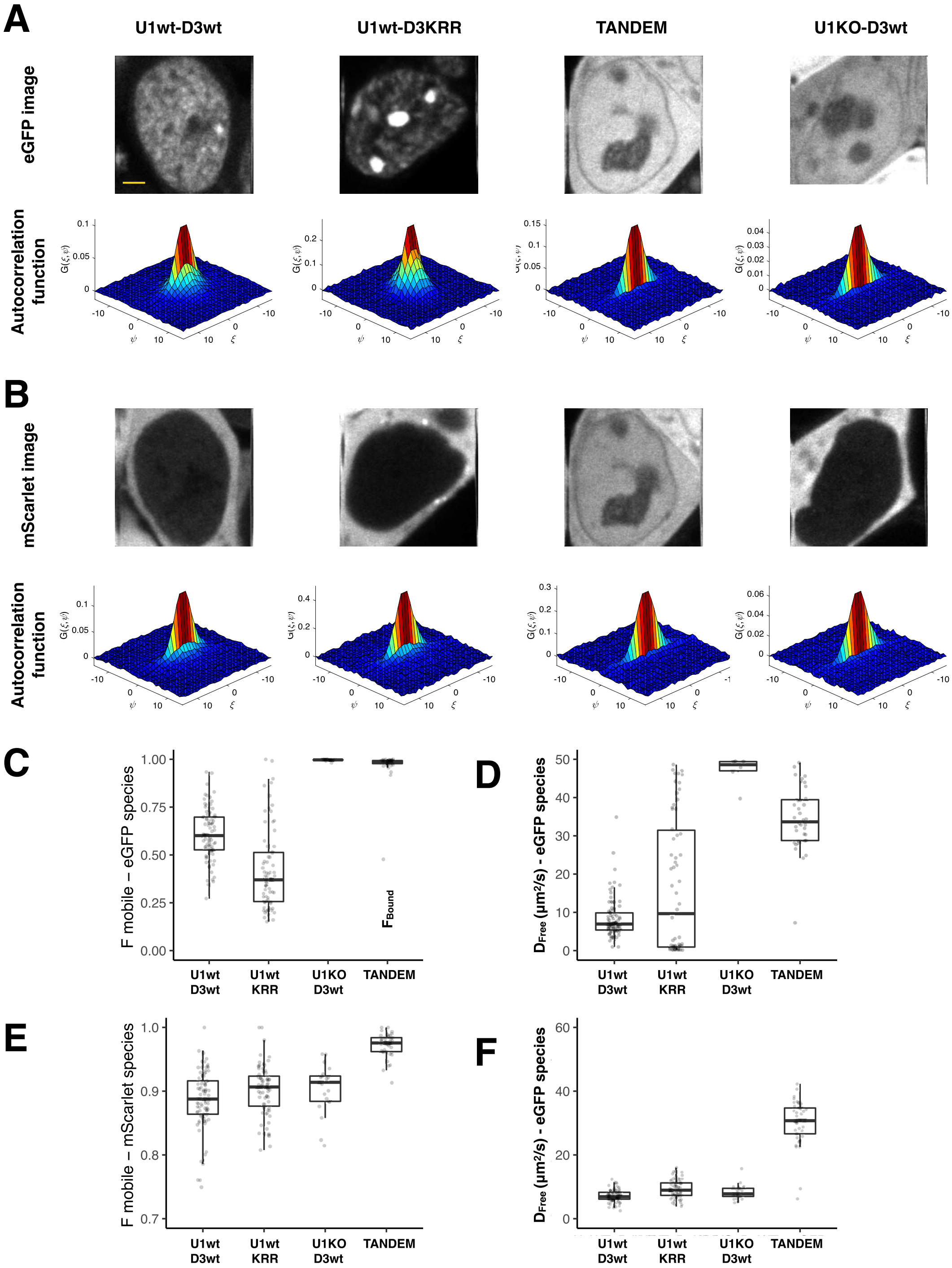


Figure 3.33 Raster Image Correlation Spectroscopy of cells expressing wt UHRF1-GFP and wt DPPA3-mScarlet, wt UHRF1-GFP and DPPA3-KRR-mScarlet, a tandem eGFP-mScarlet fusion construct, or UHRF1 KO cells expressing wt DPPA3-mScarlet. A, example images (top), and autocorrelation functions (bottom) of eGFP labeled molecules in each cell background. B Same as in A, but showing mScarlet constructs. C, and E mobile fraction of eGFP and mScarlet labeled proteins. D and F, diffusion rates of eGFP and mScarlet proteins. Data in C, D, E, and F derived from two-component diffusive model fitting. Scale bar = 2 µm

Discussion

4.1 Tet1 and Tet2 mobility in live ESCs

4.1.1 Functional cell lines to shed light on Tet protein dynamics

How Tet proteins are recruited to chromatin for epigenetic and transcriptional modulation remains poorly understood. The spatiotemporal coordination of nuclear enzymes and transcription factors, which drive changes in gene expression, ultimately shape how and when the cell progresses through development. Thus, understanding how nuclear proteins find their genomic targets is critical to better understanding epigenetic modulation and transcriptional regulation. Since Tet proteins are critical genome modifiers during development, and that their misregulation or mutation are frequently associated with Acute Myeloid Leukaemia, a further understanding of their behaviour is critical. With imaging technologies continuously evolving, observing transient interactions of chromatin-binding proteins allow to dissect the binding and diffusive properties underlying Tet protein behaviour.

Clearly, Tet proteins can bind euchromatin loci such as promoters, enhancers, distal elements, and RNA (Williams et al, 2011, Wu et al, 2011, Xiong et al 2016, Zhou et al 2019); and many chromatin associated interactors have been described, such as OGT (Vella et al 2013), Sin3A (Williams et al, 2011), HDAC1, HDAC2, Lin28 (Zeng et al, 2016), and pluripotency factors (Costa et al, 2013). In parallel, TET is frequently associated with it catalytic activity, oxdizing methylcytosine.

To our knowledge, no study has yet looked at endogenously labeled Tet1 and Tet2 in living cells. For many proteins, the pitfalls of overexpression have been discussed, but these undesired effects are often only visible with microscopy. This idea can be inferred from in vitro work on Tet1. In the vast majority of studies looking at Tet proteins in solution, only the catalytic domain is stably purified and maintained in solution, thereby missing the potential contribution of roughly two-thirds of the protein to its binding and activity. This is likely due to the difficulty of purifying such large, mostly unstructured proteins, and due to their potentially low stability in solution. Consistent with this notion, crystal structures exist of only the C-terminal catalytic domain (amino acids 1364 - 2008) and the CXXC domain (amino acids 587 - 632). ChIP experiments require reliable antibodies. While a few ChIP-seq profiles have been generated for Tet1 using antibodies against Tet1 (rather than a FLAG-tag), to date no ChIP-seq grade antibodies have not been successfully generated and used for Tet2 genome binding profiling, as discussed in Xiong et al, 2016 and Pantier et al, 2019.

Overexpression of fusion proteins could circumvent some of the issues mentioned above, for example by using a stable epitope for pull-downs, such as eGFP or the FLAG tag. However, artefacts of overexpression are to be taken seriously. Specifically, for Tet proteins, the caveats of overexpression seem particularly striking. Unpublished data from our lab showed that ectopic expression of Tet proteins using a Sleeping beauty / piggy BAC system, lead to unexpected localization phenotypes (work by Carina Trummer and Christopher Mulholland, data not shown). Transient transfection, while leading to more consistent localization patterns, suffers from low efficiency - few stem cells are transfected. Importantly, transient overexpression leads to concentrations that can be orders of magnitude higher than that of the endogenous protein, thereby making accurate quantification of diffusive and binding properties difficult.

Using Cas9-mediated genome engineering, we generated cell lines where Tet proteins were endogenously labeled with Halo- and SNAP-tags. These tags enabled the use of bright and photostable Janelia Fluor organic dyes (Grimm et al 2015), providing strong signal-to-noise, and enabling sparse labeling for single-molecule tracking. We used several microscopy systems equipped with high NA objectives and EMCCD cameras or photon-counting avalanche photodiodes, thereby allowing SMT, RICS, FRAP to directly measure the binding and diffusive properties of Tet proteins, in living cells (reviewed in Liu and Tijan 2018). The combination of genome engineering to insert the HaloTag is consistent with multiple recent studies investigating the in vivo behaviour of other nuclear proteins such as Sox2 (Chen et al 2014, Liu et al 2014), c-Myc and PTEFB (Izeddin et al 2014), telomerase and TRF2 (Schmidt et al 2016), Ezh2 and Suz12 (Youmans et al 2018), Cbx7-PRC1 (Zhen et al 2016), CTCF (Hansen et al 2017, Hansen et al 2019), RNA Polymerase II (McSwiggen et al 2019), and CRISPR-Cas9 (Knight et al, 2015).

As a first test, we used semi-quantitative analysis of live-cell images to look for differences in the spatial distribution of Tet proteins. First, the coefficient of variation, which simply measures how much the fluorescent signal intensities disperse from the mean intensity, gives a semi-quantitative readout of the spatial organization of the protein. In this regard, the “spotty” appearance of Tet1 lead to a high coefficient of variation compared to the “diffuse” appearance of Tet2. Given their similar expression levels and sub-cellular localization, this result suggested that either Tet1 was more likely bound or organized in small clusters, and that Tet2 was more likely in a diffusive state.

Next, we simply observed the behaviour of Tet proteins throughout cell division. Since nuclear envelope breakdown leads to an increase in the explorable volume of nuclear proteins, proteins that exist in a largely diffusive state will tend to be depleted from mitotic chromosomes, whereas proteins with a high on-rate will remain bound. Tet1 remained largely bound throughout the duration of mitosis, whereas Tet2 was mostly depleted, leaving a visible void of fluorescence in the volume occupied by the chromosomes.

The results from these semi-quantitative analyses were remarkably accurate predictors of the underlying physical behaviour we later observed using quantitative methods. Both coefficient of variation and mitotic chromosome binding hinted at Tet1 having a larger bound fraction than Tet2, and Tet2 having a greater diffusive fraction than Tet1. What's more, it has been shown that mitotic chromosome binding of nuclear proteins strongly correlates with their on-rate, but not their residence time (Raccaud et al, 2019). In this regard, our mitotic chromosome binding results hinted at Tet1 having a higher on-rate than Tet2, which we showed later with single-molecule tracking.

Mitotic chromosome binding evoked the notion that Tet1 may be involved in some form of mitotic bookmarking. It is indeed tempting to speculate that, along with Tet1's role in the transition of cells through pluripotency (Fidalgo et al 2016), perhaps Tet1 binding to target loci enhances a transition of cell identity from one cell division to the next.

Finally, it has been suggested that nuclear proteins bound to chromatin are protected from proteasome-mediated degradation (Coppotelli et al, 2011). While this is also a tempting hypothesis, especially given the downregulation of Tet2 during the naive to primed transition, deletion mutants of Tet1 which affect its chromatin binding do not appear to undergo substantial downregulation during the naive to primed transition.

4.1.2 Biophysical behaviour of Tet proteins in ESCs

Our live-cell imaging data using orthogonal microscopy techniques reveal a model where the binding and diffusive dynamics of Tet1 and Tet2 are starkly contrasting. While both proteins diffuse in the nucleus at rates around 2 $\mu m^2/s$, Tet1 is much more likely to be in a bound state (~60%), due to both its higher on-rate, and its longer residence time. Thus, it is frequently sampling the genome by transient binding, until it forms a stable interaction of tens-of-seconds in duration.

Tet2, on the other hand, is more likely to be in a diffusive state, binding less frequently to chromatin than Tet1, and for shorter periods of time. Thus, it diffuses through the nucleoplasm for longer periods of time before binding to a non-specific or a target to chromatin locus.

We found that Tet1 and Tet2 both diffuse in the nucleus at a rate of ~2 $\mu m^2/s$. However Tet1 was found to be more in a bound state, compared to Tet2. This preferential bound state is due to a higher on-rate and longer residence time of Tet1 compared to Tet2. Thus, Tet1 undergoes shorter bouts of diffusion between binding events, and binding events are longer in duration than Tet2, ranging from sub-second to tens-of-second. This indicates that Tet1 molecules are more likely to sample nearby chromatin sites within a local volume. In contrast, Tet2 molecules are more likely to freely diffuse for longer periods of time than Tet1, thereby sampling chromatin sites less frequently, and more sparsely distributed throughout the volume of the nucleus.

These contrasting behaviours likely influence how and where Tet protein activity takes place. The frequent binding of Tet1 is consistent with compact exploration, where a molecule samples proximal sites, which is thought to lead to higher density of local activity within a neighbouring volume, as well as clustering and bursting behaviour (Meyer et al, 2012). Tet2, on the other hand, since it is in a predominantly diffusive state, is more likely a non-compact explorer, thought to sample distant chromatin targets, since longer bouts of diffusion allow it to cover greater distances. Such behaviours have substantial implications on how molecules can exert their activity and carry information, since compact exploration would lead to local binding and activity, whereas non-compact exploration would lead to broadly distributed in the nucleus (FIGURE 4.1).

At longer timescales, these contrasting behaviours give rise to visible differences in spatial distribution. The frequent local binding behaviour of Tet1 gives rise to its spotty appearance in the nucleus, as well as mitotic chromosome binding. In contrast, the largely diffusive nature of Tet2 yields a diffuse appearance in the nucleus, and a largely depleted presence on mitotic chromosomes.

Typically, the diffusion rates of transcription factors and nuclear enzymes in the nucleus is generally on the order of 1-5 $\mu m^2/s$ (BioNumbers, see Table 1.1). Thus, the diffusion rates of Tet1 and Tet2 being roughly 2 $\mu m^2/s$ are well within the expected range. What's more, these results are derived from orthogonal methods, RICS and SMT, thereby underscoring the robustness of these findings.

RICS measurements yielded faster diffusion rates for both Tet1 and Tet2 compared to SMT, despite being performed in similar cells. We postulate that these differences are due to the nature of the experiments. RICS measurements integrate information at faster timescales than SMT: pixels are scanned every 10 us, and lines every 5 ms, as is based on the correlation of fluorescence signal at different time intervals. SMT on the other hand requires directly following a particle over multiple frames, acquired every 12.5, 22, or 33 ms. The difficulty in detecting and tracking quickly moving particles compared to slowly moving ones, known as the population exclusion effect (Izeddin et al, 2014), thus biases SMT results in a different way than RICS. Thus RICS better suited to capturing rapidly diffusing species compared to SMT, and thus we attribute a difference in measured diffusion rates to limitations in the system. Nonetheless, SMT, RICS, and FRAP all consistently revealed that Tet1 is less mobile, globally, than Tet2, again highlighting the robustness of these finding.

Furthermore, both RICS and SMT showed an interesting trend, where the diffusive fraction of Tet2 nearly always had a faster diffusion rate than Tet1. While Tet1 is slightly larger than Tet2, that difference in size is not large enough to explain the difference in diffusion rates, based on the Stokes-Einstein equation, in a manner derived and discussed for the diffusion rates of differently sized variants of RNA Polymerase II (Boehning et al, 2018). Briefly, the diffusion rate (D) of a molecule in a medium can be derived from the Stokes-Einstein equation:

$$D = \frac{k_B T}{6\pi\eta r}$$

where k_B is the Boltzmann constant, T is the temperature, η is the viscosity of the medium and r is the radius of the molecule. Assuming both Tet1 and Tet2 are diffusing within the same medium, then all parameters contributing to their diffusion rates are the same, with the exception of their radius. Tet1 and Tet2 do have slightly different molecular weights. However the radius of a molecule scales with the cube its molecular weight. Thus, the small (~5%) relative difference in molecular weight between Tet1 and Tet2 would lead to an even smaller difference in radius, which would not likely lead to a visible difference in diffusion rates.

One the one hand, this finding is in line with the idea that different biologically active nuclear proteins "see" a different environment in which to diffuse, and thus "experience" different viscosities, compartments and barriers, affecting its diffusion rate in ways which are not predicted by the Stokes-Einstein equation. In this regard, rapid bouts transient binding, or diffusion along surfaces of reduced dimensionality, could lead to reduced diffusion rates in SMT measurements (Woringer and Darzacq 2019).

We then found that the bound fraction of Tet1 is consistently greater than that of Tet2. RICS measurements revealed that at the ~20 ms timescale, roughly 70% of Tet1 molecules are in a bound state. In sharp contrast, only 40% of Tet2 molecules are found in a bound state, under identical measurement conditions.

Thus, in naive ESCs, methylation homeostasis and hmC levels are maintained by two contrasting biophysical behaviours. First, these contrasting behaviours are conceptually consistent with previous reports speculating on Tet protein behaviour. Tet2 has been suggested to be less strongly bound to chromatin compared to Tet1, due to its lack of a CXXC domain, or any other cognate DNA binding domain (Xiong et al 2016). Meanwhile, Tet1 has been suggested to a more targeted activity, binding specific promoters at different developmental stages, and capable of recruiting transcriptionally activating and repressing proteins (Wu et al, 2011).

Secondly, these distinct behaviours are consistent with a model of cooperative step-wise oxidation activity of Tet1 and Tet2, facilitated by Sall4. In this model, Tet1 would first bind enhancer regions, and oxidize mC to hmC. Tet1 and hmC at enhancers would then recruit Sall4, which would in turn recruit Tet2 for further oxidation of hmC to fC and caC.

Finally, such contrasting behaviour is consistent with the idea that these two proteins have distinct functions in the nucleus. In this regard, while Tet1 and Tet2 have similar enzymatic activity, exactly how that activity is spatiotemporally distributed in the nucleus would most probably affect their output. Thus, we propose that Tet1 is a compact explorer and Tet2 is a non-compact explorer of the nucleus.

4.1.3 Tet protein dynamics are globally methylation independent

To investigate whether methylation globally recruits Tet proteins to chromatin loci, we again performed a series of imaging experiments. First, we observed no global change in the spatial distribution of Tet1 and Tet2, in that Tet1 maintained a spotty appearance, whereas Tet2 had a more diffuse distribution. Furthermore, we observed no striking difference in their localization during mitosis, with Tet1 still binding mitotic chromosomes, and Tet2 being largely diffuse in the cytosol. These

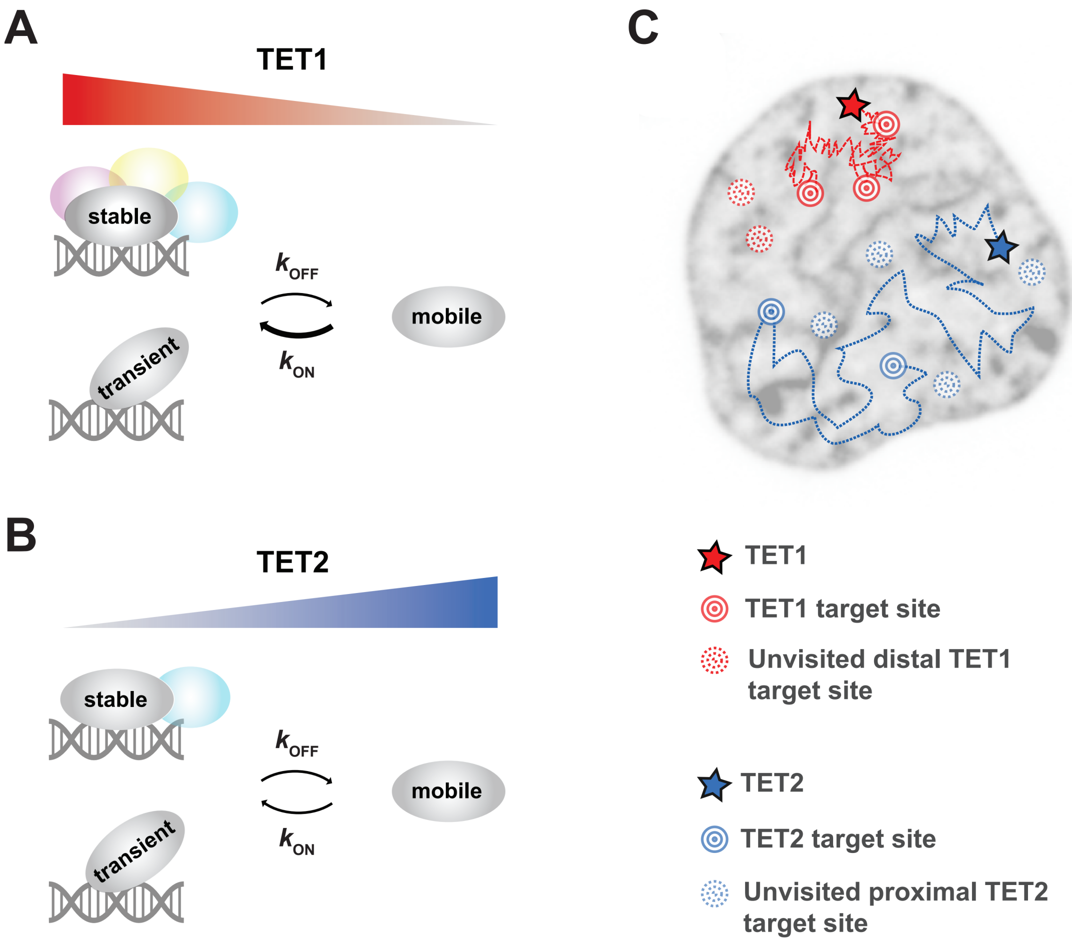


Figure 4.1 Cartoon model of Tet1 and Tet2 biophysical behaviour. A, Tet1 is more likely to be in a bound state than in a diffusive state. These binding events vary from sub-second transient events to stable tens-of-seconds binding events. Stable events most likely occur with complex formation with interactors. B, Tet2 is more likely found in a diffusive state, punctuated by transient and stable binding events, which occur less frequently than Tet1. C, as a result of these binding and diffusive behaviours, Tet1 is likely a compact explorer, more likely to bind to proximal target sites, whereas Tet2 is more likely a non-compact global explorer, diffusing through the nucleus leaving proximal target sites unvisited, where the likelihood binding one or another target is less dependent on the distance between Tet2 and these targets.

observations suggested that methylation is not a primary factor driving chromatin recruitment, nor the clustered distribution of Tet1.

Then, using SMT and FRAP, we detected no substantial difference in the binding and diffusive properties of Tet1 and Tet2. Thus, these data indicate that the catalytic activity of Tet proteins is downstream of its binding and diffusive behaviours. Specifically, this suggests that globally methylation neither affects the on-rate nor the off-rate. These data are consistent with the notion that chromatin as being a main crowding agent which influences kinetic. Thus, if transient binding occurs throughout the accessible portion of the genome, and that chromatin itself acts as a regulator of diffusive behaviours, then methylcytosine would not globally disrupt binding and diffusive dynamics, since it accounts for only a small fraction of the genome.

In parallel, these data suggest that the oxidation reaction is transient, and does not influence the off-rate of Tet protein binding. In other words, the oxidation reaction "fits" within the distribution of residence times of Tet proteins, and engaging in an oxidation reaction does not globally lengthen the residence time of a protein. Therefore, the catalytic activity of Tet proteins may be secondary to Tet protein mobility, and possibly even uncoupled.

Interestingly, other nuclear enzymes were shown to have unaffected kinetics in the absence of their substrate. Such is the case for PRC2 proteins, a histone methyltransferase complex which drives the formation of H3K27me3. It was found that acutely depleting PRC2 substrates H3K27me1, H3K27me2 and H3K37me3, using a small molecule inhibitor had only a modest effect on chromatin binding of PRC2 components Suz12 and Ezh2. (Youmans et al, 2017). These findings suggest that PRC2 components bind a variety of chromatin environments for durations which are permissive to catalytic activity.

4.1.4 The N-terminal domains controls the localization and dynamics of Tet1

By profiling the localization and mobility of a series of deletion mutants using the MIN-tag strategy, we show direct evidence the the N-terminal domain restricts the localization and binding of Tet1 to euchromatin loci. Thus, we provide the first evidence in living cells that the spatial distribution of Tet1 depends largely on the contribution of multiple peptide domains in its N-terminal region.

While it is not entirely surprising that roughly two-thirds of the protein contributes to spatial distribution, there were very few functional domains described in the N-terminal portion of Tet1 besides the CXXC domain, at the start of this project. We expressed a series of Tet1 deletion mutants fused to mNeonGreen, and we found that the catalytic domain of Tet1 is highly mobile, and that its binding patterns are heavily influenced by the N-terminal portion of Tet1.

In parallel, it was suggested that a positively-charged stretch of positively-charged amino acids at the N-terminus of Tet1 could enhance global chromatin binding (Zhang et al, 2016). Our FRAP results were consistent with the presence of a chromatin binding domain near the N-terminus of Tet, and we generated a cell line wherein the fragment encoding the first 131 amino acids was deleted.

Tet1 Δ1-131 has a nuclear distribution more similar to wild-type Tet1, with visible spots distributed throughout the nucleus. However, Tet1 Δ1-131 which showed a lower coefficient of variation of pixel intensity compared to wt Tet1, consistent with a more diffuse protein. Strikingly, deletion of this positively charged domain lead to a substantial depletion of Tet1 from mitotic chromosomes. Then, using FRAP and SMT, we showed that Tet1 Δ1-131 indeed has a greater mobility than wild-type Tet1, suggesting less global binding, and SMT suggested that this difference in binding was likely due to a lower on-rate, or a lower frequency of chromatin association for Tet1 Δ1-131. Overexpression of those 131 amino acids fused to GFP revealed clear binding to mitotic chromosomes as well as a nuclear distribution reminiscent of DNA labeling, indicative of non-specific DNA binding.

Curiously, the catalytic domain of Tet1 associates with heterochromatin, and the 1-131 domain associates unspecifically to DNA, including at heterochromatin. Thus, more work on the contributions of different domains of Tet1 would shed light on how the protein maintains euchromatin localization, despite potent heterochromatin binding domain.

Thus, we propose that Tet1 protein motion is largely driven by transient binding events in the nucleus, which allow it to slide and hop along chromatin to find target sites, and form functional interactions.

Interestingly, recent work has shown that a globally positive charge in a peptide can greatly influence its diffusion rate. It was recently shown that by adding positively charged amino acids to a peptide sequence, the diffusion rate of the protein could be decreased by an order of magnitude (Xiang et al, 2019). This was shown by single-molecule tracking using a novel illumination scheme, and performed on photoconvertible fluorescent proteins, generally seen as “biologically” inert. Adding positively charged residues, either in an undirected manner, or in the context of a nuclear localization signal, lead to decrease in the diffusion rate in the nucleus by roughly an order of magnitude. The underlying mechanisms of this is unclear, however it is thought that these stretches of positive charges increase the likelihood of transient binding interactions in the nucleus. In this regard, a canonical nuclear localization signal, also containing positively charged residues, is thought to mediate transient interactions with DNA.

Thus, we speculate that the positively charged domain at the N-terminus of Tet1 enhances Tet1 chromatin binding by enriching for transient binding events. These transient binding events could then lead to more stable binding via the contribution of other DNA-binding domains, such as the CXXC and Cysteine-rich domains, or through interactions with other DNA-binding proteins, such as Sin3A and Nanog. Furthermore, the multiple modes by which Tet1 can bind chromatin also highlight the multi-functionality of Tet1, and its potential non-catalytic functions. Tet1 harbours multiple phosphorylation, ubiquitination , and O-GlcNAcylation (Bauer et al 2015), and thus its activity and interaction potential can likely be controlled by these post-translational modifications. Thus, its multiple modes of binding could reflect Tet1's ability to both promote transcription and silencing, maintain pluripotency and contribute to transitioning.

Finally, it is tempting to speculate of the function of this domain with regards to propagating cell identity through multiple rounds of cell division. We show that Tet1 binds mitotic chromosomes in a manner that is largely dependent on the costively-charged N-terminus. Removal of this domain leads to a visible depletion of binding to mitotic chromosomes following nuclear enveloped breakdown and expression of this domain on its own leads to strong mitotic chromosome binding.

It is worth noting that this positively charged domain, along with the CXXC domain, are absent in a recently described isoform of Tet1 which is expressed in somatic cells, though at low levels. Thus, the high on-rate and mitotic chromosome binding conferred by the positively charged domain in thus lost in these cells. How this affects the activity of Tet1 remains unclear.

4.1.5 A quantitative framework in which to consider Tet protein activity

Our RICS measurements have provided a mean concentration of Tet proteins in the nuclei of embryonic stem cells. Tet1 and Tet2 each appear to be expressed in the range of 5 nM, which translates to roughly 5000 molecules of each Tet1 and Tet2 in the nuclei of ESCs. In parallel, mass spectroscopy measurements have revealed that methylcytosine and hydroxymethylcytosine compose roughly 0.6% and 0.015% of total DNA in naive embryonic stem cells and that both Tet1 and Tet2 contribute roughly equally to hmC production (Mulholland et al, 2018). Given the size of the mouse genome, this yields on the order of 10^7 instances of methylcytosine, and on the order of 10^6 instances of hydroxymethylcytosine. Thus, an approximate total of

10 000 molecules of Tet protein, bind and diffuse throughout the genome, and convert 10^6 instances of mC to hmC throughout each cell cycle.

These numbers thereby allow an approximation of the reaction rate of Tet protein catalytic activity. In parallel, cell cycle duration is roughly 10 hours in ESCs. Given these numbers, the slowest possible *in vivo* reaction rate for Tet catalytic activity is around 5 minutes per hmC. However, in this upper boundary context, each Tet protein would have to be actively engaged in productive oxidation at all times throughout the cell cycle.

Thus the reaction rate must be more rapid than this, since both Tet1 and Tet2 appear to bind on the seconds to tens-of-seconds timescale, interspersed with bouts of 3D diffusion. While it is assumed that free diffusion would not be permissive to the oxidation reaction, transient binding could be sufficient to allow the completion of the oxidation reaction.

Genome-wide binding studies revealed important insights with regards to reaction rates of epigenetic enzymes. For Tet1, ChIP-Seq studies have shown that distal genomic regions enriched for hmC do not correlate with Tet1 binding regions (Zhang et al, 2016). A similar case was observed for Suz12, a PRC2 component which drives the formation of H3K27me3. There also, it was observed that sparse H3K27me3 can it was found in regions which do not correlate with Suz12 binding regions (Hojfeldt et al 2018). What's more, it was suggested that the catalytic activity of PRC2 could even be carried by its diffusive fraction (Youmans et al, 2018). In these cases, the enzyme is not found where the product is found, indicating that reaction can occur on timescales which are too transient for detection via ChIP-Seq, which typically requires cross-linking and stringent washes.

Further clues about the in vivo reaction rate of Tet catalytic activity can be derived from hmC and Tet levels in Epiblast-like cells. In Epiblast-like cells, both mC and hmC levels are higher, however total Tet protein levels are much lower, since Tet1 is mildly downregulated, and Tet2 level reach nearly background levels. Thus, assuming a doubling of hmC and a halving of Tet molecules, the oxidation reaction must take place within ~1 minute at most, assuming again that each protein is actively engaged in cytosine oxidation at all times. Again, since molecules are seen in a diffusive state in the nucleus, this reaction rate should be seen as an upper limit to the actual reaction.

Clearly, many open questions regarding the *in vivo* reaction rates remain unanswered. For example, what are the rate-limiting steps to the oxidation reaction? One could speculate that substrate availability is rate limiting, given the changes in levels during the naive ESC to primed epiblast transition. In parallel, perhaps the availability of co-factors involved in the oxidation reaction, such as iron and alpha-ketoglutarate, could be rate-limiting to the oxidation reaction. However, a notion of substrate accessibility comes to mind, given the fractal organization of the nucleus (Bancaud et al, 2009), and since Tet proteins are enriched in euchromatin and largely depleted in methylcytosine-rich heterochromatin. Furthermore, Tet proteins appear to globally bind chromatin loci independently of global methylation levels, as shown in section 2.1.5. Thus, there appears to be an intricate intertwining of substrate availability and how it is spatially organized within the volume accessible by Tet proteins.

4.2 DPPA3 interacts with UHRF1 in living cells and impairs its recruitment to chromatin loci

Mammalian genomes undergo a wave of DNA demethylation shortly after fertilization, however the mechanism which drive this wave was unclear. The role of this wave of demethylation is unclear. On the one hand, such demethylation could expedite the establishment of the epigenetic marks required for transitioning through pluripotency and differentiation. On the other hand, since a conserved role for DNA methylation is to suppress the expression of transposable elements and endogenous retroviral elements, a wave of demethylation leaves the cell vulnerable to over expression of transposable elements, which can lead to genomic instability (reviewed by Klein and O'Neill, 2018). How such a balance was selected for and ultimately integrated into mammalian developmental processes remains unknown.

DNA methylation machinery is largely conserved throughout eukaryotes, though it has been lost in several organisms, such as *C. elegans*, *D. Melanogaster*, and *A. thaliana* (Greenberg and Bourc'his, 2019). Together, these hint at a recently evolved pathway to drive wave of DNA demethylation.

We showed using live-cell microscopy that in naive pluripotent stem cells DPPA3 expression leads to a decrease in chromatin association of UHRF1, and its export outside the nucleus. Our RICS data showed that DPPA3 forms a complex with UHRF1, which directly impairs its capacity to bind chromatin. UHRF1 binding to chromatin is not impaired when a mutant of DPPA3 which does not bind UHRF1 *in* *vitro* is expressed. In a follow-up experiment, it was shown that over expression of DPPA3 variants which bind UHRF1 lead to decreased methylation levels, whereas over expression of DPPA3 constructs lacking UHRF1-binding do not have impaired methylation levels (work by Chris Mulholland, data not shown). As a result, we provide a mechanism by which maintenance methylation is impaired during naive pluripotency, which underlies previous reports of impaired recruitment of DNA methylation machinery during replication in naive pluripotent stem cells (von Meyenn et al, 2016).

Interestingly, DPPA3 is uniquely present in mammalian genomes, and its expression in organisms which do not harbour a DPPA3, such as medaka and Xenopus (*in vitro* extract) lead to a massive decrease in DNA methylation and, in medaka, lethality. Thus, it is tempting to speculate that DPPA3 recently evolved in mammals as a novel mechanism to drive global demethylation.

4.3 Open questions for follow-up experiments

In this work, we have established a first-order approximation of Tet protein dynamics. Follow-up experiments could involve investigating known Tet protein-protein interactions, which would shed light on on the mechanistic contribution of many Tet interactors on Tet protein binding and diffusive properties. To this end, FRAP and SMT could be performed on cells expressing HaloTet1 or HaloTet2 in different genetic backgrounds, wherein known Tet interactors are depleted, either by knock out, or by acute protein degradation mediated by the Auxin-AID or a light-activated protein degradation.

How do Tet protein functions change as cells transition through pluripotency and engage in lineage commitment? While our results has focused mostly on naive pluripotency, the spatial distribution and abundance of Tet proteins change dramatically as cells transition to the epiblast state. In addition, Tet3, which was not within the scope of this work, is expressed at very early embryonic stages and at later time points (Gu et al, 2011). Thus, the elucidating the dynamic properties of Tet proteins may help unravel how the contributions of individual Tet proteins through different developmental states. To this end, a similar labeling strategy could be used for live-cell imaging, and could be combined with interactome studies at concurrent developmental times.

In parallel, changes to nuclear architecture as cells progress through pluripotency may drive different binding and diffusive properties for chromatin-binding proteins, thus a careful characterization of chromatin motion at different developmental time points would be critical.

An exciting direction in which live-cell single-molecule studies is moving concerns elucidating the context in which single-molecule spots are detected (Woringer and Darzaqc, 2019). For example, for any given Tet1 trajectory, what is the underlying chromatin distribution, and what is the underlying movement of chromatin at those Tet1 locations. Transient protein-protein binding reactions could be monitored using FRET By using two or three detectors simultaneously, one could observe multiple fluorescently labeled proteins at the same time. In this regard, Janelia Fluor dyes are available conjugated to Halo, SNAP, and CLIP tags, thus offering several multilabeling strategies, in addition to the current palette of genetically encoded fluorescent proteins.

Finally, combining these live-cell imaging approaches with sequencing methods which reveal binding profiles, transcriptomes and chromosome organization, will likely reveal generalizable principles of how transcription factors and nuclear architecture coordinate development and differentiation.

References

Abrahamsson, S., Chen, J., Hajj, B., Stallinga, S., Katsov, A.Y., Wisniewski, J., Mizuguchi, G., Soule, P., Mueller, F., Darzacq, C.D., Darzacq, X., Wu, C., Bargmann, C.I., Agard, D.A., Dahan, M., Gustafsson, M.G.L., 2013. Fast multicolor 3D imaging using aberration-corrected multifocus microscopy. Nat Methods 10, 60–63. https://doi.org/10.1038/nmeth.2277

Amouroux, R., Nashun, B., Shirane, K., Nakagawa, S., Hill, P.W.S., D'Souza, Z., Nakayama, M., Matsuda, M., Turp, A., Ndjetehe, E., Encheva, V., Kudo, N.R., Koseki, H., Sasaki, H., Hajkova, P., 2016. De novo DNA methylation drives 5hmC accumulation in mouse zygotes. Nat Cell Biol 18, 225–233. https://doi.org/10.1038/ncb3296

Antequera F., Tamame M., Villanueva J.R., Santos T., 1984. DNA methylation in the fungi. J. Biol. Chem 259(13).

Axelrod, D., Koppel, D.E., Schlessinger, J., Elson, E., Webb, W.W., 1976. Mobility measurement by analysis of fluorescence photobleaching recovery kinetics. Biophysical Journal 16, 1055–1069. https://doi.org/10.1016/S0006-3495(76)85755-4

Ayyanathan, K., 2003. Regulated recruitment of HP1 to a euchromatic gene induces mitotically heritable, epigenetic gene silencing: a mammalian cell culture model of gene variegation. Genes & Development 17, 1855–1869. https://doi.org/10.1101/gad.1102803

Bancaud, A., Huet, S., Daigle, N., Mozziconacci, J., Beaudouin, J., Ellenberg, J., 2009. Molecular crowding affects diffusion and binding of nuclear proteins in heterochromatin and reveals the fractal organization of chromatin. EMBO J 28, 3785–3798. https://doi.org/10.1038/emboj.2009.340

Bauer, C., Göbel, K., Nagaraj, N., Colantuoni, C., Wang, M., Müller, U., Kremmer, E., Rottach, A., Leonhardt, H., 2015. Phosphorylation of TET Proteins Is Regulated via O -GlcNAcylation by the O -Linked N -Acetylglucosamine Transferase (OGT). J. Biol. Chem. 290, 4801–4812. https://doi.org/10.1074/jbc.M114.605881

Bénichou, O., Chevalier, C., Klafter, J., Meyer, B., Voituriez, R., 2010. Geometry-controlled kinetics. Nature Chem 2, 472–477. https://doi.org/10.1038/nchem.622
Berg, O.G., Winter, R.B., Von Hippel, P.H., 1981. Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. Biochemistry 20, 6929–6948. https://doi.org/10.1021/bi00527a028

Bestor, T.H., Edwards, J.R., Boulard, M., 2015. Notes on the role of dynamic DNA methylation in mammalian development. Proc Natl Acad Sci USA 112, 6796–6799. https://doi.org/10.1073/pnas.1415301111

Bestor T., Laudano A., Mattaliano R., Ingram V., 1988. Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. J Mol Biol 203(4).

Betzig, E., Patterson, G.H., Sougrat, R., Lindwasser, O.W., Olenych, S., Bonifacino, J.S., Davidson, M.W., Lippincott-Schwartz, J., Hess, H.F., 2006. Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. Science 313, 1642–1645. https://doi.org/10.1126/science.1127344

Bird, A., 2002. DNA methylation patterns and epigenetic memory. Genes & Development 16, 6–21. https://doi.org/10.1101/gad.947102

Boehning, M., Dugast-Darzacq, C., Rankovic, M., Hansen, A.S., Yu, T., Marie-Nelly, H., McSwiggen, D.T., Kokic, G., Dailey, G.M., Cramer, P., Darzacq, X., Zweckstetter, M., 2018. RNA polymerase II clustering through carboxy-terminal domain phase separation. Nat Struct Mol Biol 25, 833–840. https://doi.org/10.1038/s41594-018-0112-y

Bogdanović, O., Smits, A.H., de la Calle Mustienes, E., Tena, J.J., Ford, E., Williams, R., Senanayake, U., Schultz, M.D., Hontelez, S., van Kruijsbergen, I., Rayon, T., Gnerlich, F., Carell, T., Veenstra, G.J.C., Manzanares, M., Sauka-Spengler, T., Ecker, J.R., Vermeulen, M., Gómez-Skarmeta, J.L., Lister, R., 2016. Active DNA demethylation at enhancers during the vertebrate phylotypic period. Nat Genet 48, 417–426. https://doi.org/10.1038/ng.3522

Bortvin, A., Goodheart, M., Liao, M., Page, D.C., 2004. BMC Dev Biol 4, 2. https://doi.org/10.1186/1471-213X-4-2

Boeynaems, S., Alberti, S., Fawzi, N.L., Mittag, T., Polymenidou, M., Rousseau, F., Schymkowitz, J., Shorter, J., Wolozin, B., Van Den Bosch, L., Tompa, P., Fuxreiter, M., 2018. Protein Phase Separation: A New Phase in Cell Biology. Trends in Cell Biology 28, 420–435. https://doi.org/10.1016/j.tcb.2018.02.004

Brown, C.M., Dalal, R.B., Hebert, B., Digman, M.A., Horwitz, A.R., Gratton, E., 2008. Raster image correlation spectroscopy (RICS) for measuring fast protein dynamics and concentrations with a commercial laser scanning confocal microscope. J Microsc 229, 78–91. https://doi.org/10.1111/j.1365-2818.2007.01871.x

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W., Prasher, D., 1994. Green fluorescent protein as a marker for gene expression. Science 263, 802–805. https://doi.org/10.1126/science.8303295

Chen, H., Larson, D.R., 2016. What have single-molecule studies taught us about gene expression? Genes Dev. 30, 1796–1810. https://doi.org/10.1101/gad.281725.116

Chen, J., Zhang, Z., Li, L., Chen, B.-C., Revyakin, A., Hajj, B., Legant, W., Dahan, M., Lionnet, T., Betzig, E., Tjian, R., Liu, Z., 2014. Single-Molecule Dynamics of Enhanceosome Assembly in Embryonic Stem Cells. Cell 156, 1274–1285. https://doi.org/10.1016/j.cell.2014.01.062

Chen, T., Ueda, Y., Dodge, J.E., Wang, Z., Li, E., 2003. Establishment and Maintenance of Genomic Methylation Patterns in Mouse Embryonic Stem Cells by Dnmt3a and Dnmt3b. Molecular and Cellular Biology 23, 5594–5605. https://doi.org/10.1128/MCB.23.16.5594-5605.2003

Clark, N.M., Hinde, E., Winter, C.M., Fisher, A.P., Crosti, G., Blilou, I., Gratton, E., Benfey, P.N., Sozzani, R., 2016. Tracking transcription factor mobility and interaction in Arabidopsis roots with fluorescence correlation spectroscopy. eLife 5, e14770. https://doi.org/10.7554/eLife.14770

Cole, N.B., Smith, C.L., Sciaky, N., Terasaki, M., Edidin, M., Lippincott-Schwartz, J., 1996. Diffusional Mobility of Golgi Proteins in Membranes of Living Cells. Science 273, 797–801. https://doi.org/10.1126/science.273.5276.797

Coppotelli, G., Mughal, N., Marescotti, D., Masucci, M.G., 2011. High Avidity Binding to DNA Protects Ubiquitylated Substrates from Proteasomal Degradation. J. Biol. Chem. 286, 19565–19575. https://doi.org/10.1074/jbc.M111.224782

Costa, Y., Ding, J., Theunissen, T.W., Faiola, F., Hore, T.A., Shliaha, P.V., Fidalgo, M., Saunders, A., Lawrence, M., Dietmann, S., Das, S., Levasseur, D.N., Li, Z., Xu, M., Reik, W., Silva, J.C.R., Wang, J., 2013. NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. Nature 495, 370–374. https://doi.org/10.1038/nature11925

Darzacq, X., Shav-Tal, Y., de Turris, V., Brody, Y., Shenoy, S.M., Phair, R.D., Singer, R.H., 2007. In vivo dynamics of RNA polymerase II transcription. Nat Struct Mol Biol 14, 796–806. https://doi.org/10.1038/nsmb1280

Dawlaty, M.M., Breiling, A., Le, T., Barrasa, M.I., Raddatz, G., Gao, Q., Powell, B.E., Cheng, A.W., Faull, K.F., Lyko, F., Jaenisch, R., 2014. Loss of Tet Enzymes Compromises Proper Differentiation of Embryonic Stem Cells. Developmental Cell 29, 102–111. https://doi.org/10.1016/j.devcel.2014.03.003

Dawlaty, M.M., Breiling, A., Le, T., Raddatz, G., Barrasa, M.I., Cheng, A.W., Gao, Q., Powell, B.E., Li, Z., Xu, M., Faull, K.F., Lyko, F., Jaenisch, R., 2013. Combined Deficiency of Tet1 and Tet2 Causes Epigenetic Abnormalities but Is Compatible with Postnatal Development. Developmental Cell 24, 310–323. https://doi.org/10.1016/j.devcel.2012.12.015

Dawlaty, M.M., Ganz, K., Powell, B.E., Hu, Y.-C., Markoulaki, S., Cheng, A.W., Gao, Q., Kim, J., Choi, S.-W., Page, D.C., Jaenisch, R., 2011. Tet1 Is Dispensable for Maintaining Pluripotency and Its Loss Is Compatible with Embryonic and Postnatal Development. Cell Stem Cell 9, 166–175. https://doi.org/10.1016/j.stem.2011.07.010

Digman, M.A., Brown, C.M., Sengupta, P., Wiseman, P.W., Horwitz, A.R., Gratton, E., 2005. Measuring Fast Dynamics in Solutions and Cells with a Laser Scanning Microscope. Biophysical Journal 89, 1317–1327. https://doi.org/10.1529/biophysj.105.062836

Du, W., Dong, Q., Zhang, Z., Liu, B., Zhou, T., Xu, R., Wang, H., Zhu, B., Li, Y., 2019. Stella protein facilitates DNA demethylation by disrupting the chromatin association of the RING finger–type E3 ubiquitin ligase UHRF1. J. Biol. Chem. 294, 8907–8917. https://doi.org/10.1074/jbc.RA119.008008

Ellenberg, J., Siggia, E.D., Moreira, J.E., Smith, C.L., Presley, J.F., Worman, H.J., Lippincott-Schwartz, J., 1997. Nuclear Membrane Dynamics and Reassembly in Living Cells: Targeting of an Inner Nuclear Membrane Protein in Interphase and Mitosis. J Cell Biol 138, 1193–1206. https://doi.org/10.1083/jcb.138.6.1193

Elson, E.L., Magde, D., 1974. Fluorescence correlation spectroscopy. I. Conceptual basis and theory. Biopolymers 13, 1–27. https://doi.org/10.1002/bip.1974.360130102

Feldman, N., Gerson, A., Fang, J., Li, E., Zhang, Y., Shinkai, Y., Cedar, H., Bergman, Y., 2006. G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. Nat Cell Biol 8, 188–194. https://doi.org/10.1038/ncb1353

Feng, S., Cokus, S.J., Zhang, X., Chen, P.-Y., Bostick, M., Goll, M.G., Hetzel, J., Jain, J., Strauss, S.H., Halpern, M.E., Ukomadu, C., Sadler, K.C., Pradhan, S., Pellegrini, M., Jacobsen, S.E., 2010. Conservation and divergence of methylation patterning in plants and animals. Proceedings of the National Academy of Sciences 107, 8689–8694. https://doi.org/10.1073/pnas.1002720107

Ficz, G., Branco, M.R., Seisenberger, S., Santos, F., Krueger, F., Hore, T.A., Marques, C.J., Andrews, S., Reik, W., 2011. Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature 473, 398–402. https://doi.org/10.1038/nature10008

Fidalgo, M., Huang, X., Guallar, D., Sanchez-Priego, C., Valdes, V.J., Saunders, A., Ding, J., Wu, W.-S., Clavel, C., Wang, J., 2016. Zfp281 Coordinates Opposing Functions of Tet1 and Tet2 in Pluripotent States. Cell Stem Cell 19, 355–369. https://doi.org/10.1016/j.stem.2016.05.025

Frauer, C., Rottach, A., Meilinger, D., Bultmann, S., Fellinger, K., Hasenöder, S., Wang, M., Qin, W., Söding, J., Spada, F., Leonhardt, H., 2011. Different Binding Properties and Function of CXXC Zinc Finger Domains in Dnmt1 and Tet1. PLoS ONE 6, e16627. https://doi.org/10.1371/journal.pone.0016627

Funaki, S., Nakamura, T., Nakatani, T., Umehara, H., Nakashima, H., Nakano, T., 2014. Inhibition of maintenance DNA methylation by Stella. Biochemical and Biophysical Research Communications 453, 455–460. https://doi.org/10.1016/j.bbrc.2014.09.101

Gassler, J., Brandão, H.B., Imakaev, M., Flyamer, I.M., Ladstätter, S., Bickmore, W.A., Peters, J., Mirny, L.A., Tachibana, K., 2017. A mechanism of cohesin‐dependent loop extrusion organizes zygotic genome architecture. EMBO J 36, 3600–3618. https://doi.org/10.15252/embj.201798083

Gebhardt, J.C.M., Suter, D.M., Roy, R., Zhao, Z.W., Chapman, A.R., Basu, S., Maniatis, T., Xie, X.S., 2013. Single-molecule imaging of transcription factor binding to DNA in live mammalian cells. Nat Methods 10, 421–426. https://doi.org/10.1038/nmeth.2411

Gerland, U., Moroz, J.D., Hwa, T., 2002. Physical constraints and functional characteristics of transcription factor-DNA interaction. Proceedings of the National Academy of Sciences 99, 12015–12020. https://doi.org/10.1073/pnas.192693599

Gibcus, J.H., Samejima, K., Goloborodko, A., Samejima, I., Naumova, N., Nuebler, J., Kanemaki, M.T., Xie, L., Paulson, J.R., Earnshaw, W.C., Mirny, L.A., Dekker, J., 2018. A pathway for mitotic chromosome formation. Science 359, eaao6135. https://doi.org/10.1126/science.aao6135

Good, C.R., Madzo, J., Patel, B., Maegawa, S., Engel, N., Jelinek, J., Issa, J.-P.J., 2017. A novel isoform of TET1 that lacks a CXXC domain is overexpressed in cancer. Nucleic Acids Research 45, 8269–8281. https://doi.org/10.1093/nar/gkx435

Greenberg, M.V.C., Bourc'his, D., 2019. The diverse roles of DNA methylation in mammalian development and disease. Nat Rev Mol Cell Biol 20, 590–607. https://doi.org/10.1038/s41580-019-0159-6

Grimm, J.B., English, B.P., Chen, J., Slaughter, J.P., Zhang, Z., Revyakin, A., Patel, R., Macklin, J.J., Normanno, D., Singer, R.H., Lionnet, T., Lavis, L.D., 2015. A general method to improve fluorophores for live-cell and single-molecule microscopy. Nat Methods 12, 244–250. https://doi.org/10.1038/nmeth.3256

Grimm, J.B., English, B.P., Choi, H., Muthusamy, A.K., Mehl, B.P., Dong, P., Brown, T.A., Lippincott-Schwartz, J., Liu, Z., Lionnet, T., Lavis, L.D., 2016. Bright photoactivatable fluorophores for single-molecule imaging. Nat Methods 13, 985–988. https://doi.org/10.1038/nmeth.4034

Grünwald, D., Martin, R.M., Buschmann, V., Bazett-Jones, D.P., Leonhardt, H., Kubitscheck, U., Cardoso, M.C., 2008. Probing Intranuclear Environments at the Single-Molecule Level. Biophysical Journal 94, 2847–2858. https://doi.org/10.1529/biophysj.107.115014

Gu, T.-P., Guo, F., Yang, H., Wu, H.-P., Xu, G.-F., Liu, W., Xie, Z.-G., Shi, L., He, X., Jin, S., Iqbal, K., Shi, Y.G., Deng, Z., Szabó, P.E., Pfeifer, G.P., Li, J., Xu, G.-L., 2011. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature 477, 606–610. https://doi.org/10.1038/nature10443

Halford, S.E., 2004. How do site-specific DNA-binding proteins find their targets? Nucleic Acids Research 32, 3040–3052. https://doi.org/10.1093/nar/gkh624

Hammar, P., Leroy, P., Mahmutovic, A., Marklund, E.G., Berg, O.G., Elf, J., 2012. The lac Repressor Displays Facilitated Diffusion in Living Cells. Science 336, 1595–1598. https://doi.org/10.1126/science.1221648

Hansen, A.S., Amitai, A., Cattoglio, C., Tjian, R., Darzacq, X., 2019. Guided nuclear exploration increases CTCF target search efficiency. Nat Chem Biol. https://doi.org/10.1038/s41589-019-0422-3

Hansen, A.S., Pustova, I., Cattoglio, C., Tjian, R., Darzacq, X., 2017. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. eLife 6, e25776. https://doi.org/10.7554/eLife.25776

Hansen, A.S., Woringer, M., Grimm, J.B., Lavis, L.D., Tjian, R., Darzacq, X., 2018. Robust model-based analysis of single-particle tracking experiments with Spot-On. Elife 7. https://doi.org/10.7554/eLife.33125

He, Y.-F., Li, B.-Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., Sun, Y., Li, X., Dai, Q., Song, C.-X., Zhang, K., He, C., Xu, G.-L., 2011. Tet-Mediated Formation of 5-Carboxylcytosine and Its Excision by TDG in Mammalian DNA. Science 333, 1303–1307. https://doi.org/10.1126/science.1210944

Hendrix, J., Baumgärtel, V., Schrimpf, W., Ivanchenko, S., Digman, M.A., Gratton, E., Kräusslich, H.-G., Müller, B., Lamb, D.C., 2015. Live-cell observation of cytosolic HIV-1 assembly onset reveals RNA-interacting Gag oligomers. The Journal of Cell Biology 210, 629–646. https://doi.org/10.1083/jcb.201504006

Hendrix, J., Dekens, T., Schrimpf, W., Lamb, D.C., 2016. Arbitrary-Region Raster Image Correlation Spectroscopy. Biophysical Journal 111, 1785–1796. https://doi.org/10.1016/j.bpj.2016.09.012

Hendrix, J., Gijsbers, R., De Rijck, J., Voet, A., Hotta, J., McNeely, M., Hofkens, J., Debyser, Z., Engelborghs, Y., 2011. The transcriptional co-activator LEDGF/p75 displays a dynamic scan-and-lock mechanism for chromatin tethering. Nucleic Acids Research 39, 1310–1325. https://doi.org/10.1093/nar/gkq933

Hendrix, J., Schrimpf, W., Höller, M., Lamb, D.C., 2013. Pulsed Interleaved Excitation Fluctuation Imaging. Biophysical Journal 105, 848–861. https://doi.org/10.1016/j.bpj.2013.05.059

Højfeldt, J.W., Laugesen, A., Willumsen, B.M., Damhofer, H., Hedehus, L., Tvardovskiy, A., Mohammad, F., Jensen, O.N., Helin, K., 2018. Accurate H3K27 methylation can be established de novo by SUZ12-directed PRC2. Nat Struct Mol Biol 25, 225–232. https://doi.org/10.1038/s41594-018-0036-6

Hon, G.C., Song, C.-X., Du, T., Jin, F., Selvaraj, S., Lee, A.Y., Yen, C., Ye, Z., Mao, S.-Q., Wang, B.-A., Kuan, S., Edsall, L.E., Zhao, B.S., Xu, G.-L., He, C., Ren, B., 2014. 5mC Oxidation by Tet2 Modulates Enhancer Activity and Timing of Transcriptome Reprogramming during Differentiation. Molecular Cell 56, 286–297. https://doi.org/10.1016/j.molcel.2014.08.026

Hu, L., Li, Z., Cheng, J., Rao, Q., Gong, W., Liu, M., Shi, Y.G., Zhu, J., Wang, P., Xu, Y., 2013. Crystal Structure of TET2-DNA Complex: Insight into TET-Mediated 5mC Oxidation. Cell 155, 1545–1555. https://doi.org/10.1016/j.cell.2013.11.020

Hu, S., Wan, J., Su, Y., Song, Q., Zeng, Y., Nguyen, H.N., Shin, J., Cox, E., Rho, H.S., Woodard, C., Xia, S., Liu, S., Lyu, H., Ming, G.-L., Wade, H., Song, H., Qian, J., Zhu, H., 2013. DNA methylation presents distinct binding sites for human transcription factors. eLife 2, e00726. https://doi.org/10.7554/eLife.00726

Inoue, A., Zhang, Y., 2011. Replication-Dependent Loss of 5-Hydroxymethylcytosine in Mouse Preimplantation Embryos. Science 334, 194–194. https://doi.org/10.1126/science.1212483

Ito, S., D'Alessio, A.C., Taranova, O.V., Hong, K., Sowers, L.C., Zhang, Y., 2010. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature 466, 1129–1133. https://doi.org/10.1038/nature09303

Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., Zhang, Y., 2011. Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine. Science 333, 1300–1303. https://doi.org/10.1126/science.1210597

Iurlaro, M., Ficz, G., Oxley, D., Raiber, E.-A., Bachman, M., Booth, M.J., Andrews, S., Balasubramanian, S., Reik, W., 2013. A screen for hydroxymethylcytosine and formylcytosine binding proteins suggests functions in transcription and chromatin regulation. Genome Biol 14, R119. https://doi.org/10.1186/gb-2013-14-10-r119

Izeddin, I., Récamier, V., Bosanac, L., Cissé, I.I., Boudarene, L., Dugast-Darzacq, C., Proux, F., Bénichou, O., Voituriez, R., Bensaude, O., Dahan, M., Darzacq, X., 2014. Single-molecule tracking in live cells reveals distinct target-search strategies of transcription factors in the nucleus. Elife 3. https://doi.org/10.7554/eLife.02230

Jackson, M., Krassowska, A., Gilbert, N., Chevassut, T., Forrester, L., Ansell, J., Ramsahoye, B., 2004. Severe Global DNA Hypomethylation Blocks Differentiation and Induces Histone Hyperacetylation in Embryonic Stem Cells. Molecular and Cellular Biology 24, 8862–8871. https://doi.org/10.1128/MCB.24.20.8862-8871.2004

Jaenisch, R., Jähner, D., 1984. Methylation, expression and chromosomal position of genes in mammals. Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression 782, 1–9. https://doi.org/10.1016/0167-4781(84)90099-X

Kang, J., Lienhard, M., Pastor, W.A., Chawla, A., Novotny, M., Tsagaratou, A., Lasken, R.S., Thompson, E.C., Surani, M.A., Koralov, S.B., Kalantry, S., Chavez, L., Rao, A., 2015. Simultaneous deletion of the methylcytosine oxidases Tet1 and Tet3 increases transcriptome variability in early embryogenesis. Proc Natl Acad Sci USA 112, E4236–E4245. https://doi.org/10.1073/pnas.1510510112

Keshet, I., Lieman-Hurwitz, J., Cedar, H., 1986. DNA methylation affects the formation of active chromatin. Cell 44, 535–543. https://doi.org/10.1016/0092-8674(86)90263-1

Khoueiry, R., Sohni, A., Thienpont, B., Luo, X., Velde, J.V., Bartoccetti, M., Boeckx, B., Zwijsen, A., Rao, A., Lambrechts, D., Koh, K.P., 2017. Lineage-specific functions of TET1 in the postimplantation mouse embryo. Nat Genet 49, 1061–1072. https://doi.org/10.1038/ng.3868

Klein, S.J., O'Neill, R.J., 2018. Transposable elements: genome innovation, chromosome diversity, and centromere conflict. Chromosome Res 26, 5–23. https://doi.org/10.1007/s10577-017-9569-5

Knight, S.C., Xie, L., Deng, W., Guglielmi, B., Witkowsky, L.B., Bosanac, L., Zhang, E.T., El Beheiry, M., Masson, J.-B., Dahan, M., Liu, Z., Doudna, J.A., Tjian, R., 2015. Dynamics of CRISPR-Cas9 genome interrogation in living cells. Science 350, 823–826. https://doi.org/10.1126/science.aac6572

Koh, K.P., Yabuuchi, A., Rao, S., Huang, Y., Cunniff, K., Nardone, J., Laiho, A., Tahiliani, M., Sommer, C.A., Mostoslavsky, G., Lahesmaa, R., Orkin, S.H., Rodig, S.J., Daley, G.Q., Rao, A., 2011. Tet1 and Tet2 Regulate 5-Hydroxymethylcytosine Production and Cell Lineage Specification in Mouse Embryonic Stem Cells. Cell Stem Cell 8, 200–213. https://doi.org/10.1016/j.stem.2011.01.008

Kornberg, R.D., 1974. Chromatin Structure: A Repeating Unit of Histones and DNA. Science 184, 868–871. https://doi.org/10.1126/science.184.4139.868

Kriaucionis, S., Heintz, N., 2009. The Nuclear DNA Base 5-Hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain. Science 324, 929–930. https://doi.org/10.1126/science.1169786

Laine, R.F., Tosheva, K.L., Gustafsson, N., Gray, R.D.M., Almada, P., Albrecht, D., Risa, G.T., Hurtig, F., Lindås, A.-C., Baum, B., Mercer, J., Leterrier, C., Pereira, P.M., Culley, S., Henriques, R., 2019. NanoJ: a high-performance open-source super-resolution microscopy toolbox. J. Phys. D: Appl. Phys. 52, 163001. https://doi.org/10.1088/1361-6463/ab0261

Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.-M., Edsall, L., Antosiewicz-Bourget, J., Stewart, R., Ruotti, V., Millar, A.H., Thomson, J.A., Ren, B., Ecker, J.R., 2009. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 462, 315–322. https://doi.org/10.1038/nature08514

Lee, H.J., Hore, T.A., Reik, W., 2014. Reprogramming the Methylome: Erasing Memory and Creating Diversity. Cell Stem Cell 14, 710–719. https://doi.org/10.1016/j.stem.2014.05.008

Leonhardt, H., Page, A.W., Weier, H.-U., Bestor, T.H., 1992. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. Cell 71, 865–873. https://doi.org/10.1016/0092-8674(92)90561-P

Li, Y., Zhang, Z., Chen, J., Liu, W., Lai, W., Liu, B., Li, X., Liu, L., Xu, S., Dong, Q., Wang, M., Duan, X., Tan, J., Zheng, Y., Zhang, P., Fan, G., Wong, J., Xu, G.-L., Wang, Z., Wang, H., Gao, S., Zhu, B., 2018. Stella safeguards the oocyte methylome by preventing de novo methylation mediated by DNMT1. Nature 564, 136–140. https://doi.org/10.1038/s41586-018-0751-5

Lippincott-Schwartz, J., Snapp, E.L., Phair, R.D., 2018. The Development and Enhancement of FRAP as a Key Tool for Investigating Protein Dynamics. Biophysical Journal 115, 1146–1155. https://doi.org/10.1016/j.bpj.2018.08.007

Liu, Z., Legant, W.R., Chen, B.-C., Li, L., Grimm, J.B., Lavis, L.D., Betzig, E., Tjian, R., 2014. 3D imaging of Sox2 enhancer clusters in embryonic stem cells. eLife 3, e04236. https://doi.org/10.7554/eLife.04236

Liu, Z., Tjian, R., 2018. Visualizing transcription factor dynamics in living cells. The Journal of Cell Biology 217, 1181–1191. https://doi.org/10.1083/jcb.201710038

Ludwig, A.K., Zhang, P., Hastert, F.D., Meyer, S., Rausch, C., Herce, H.D., Müller, U., Lehmkuhl, A., Hellmann, I., Trummer, C., Storm, C., Leonhardt, H., Cardoso, M.C., 2017. Binding of MBD proteins to DNA blocks Tet1 function thereby modulating transcriptional noise. Nucleic Acids Research 45, 2438–2457. https://doi.org/10.1093/nar/gkw1197

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., Richmond, T.J., 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature 389, 251–260. https://doi.org/10.1038/38444

Luo, X., van der Veer, B.K., Sun, L., Bartoccetti, M., Boretto, M., Vankelecom, H., Khoueiry, R., Koh, K.P., 2020. Coordination of germ layer lineage choice by TET1 during primed pluripotency. Genes Dev. genesdev;gad.329474.119v1. https://doi.org/10.1101/gad.329474.119

Maddox, P.S., Portier, N., Desai, A., Oegema, K., 2006. Molecular analysis of mitotic chromosome condensation using a quantitative time-resolved fluorescence microscopy assay. Proceedings of the National Academy of Sciences 103, 15097–15102. https://doi.org/10.1073/pnas.0606993103

Magde, D., Elson, E.L., Webb, W.W., 1974. Fluorescence correlation spectroscopy. II. An experimental realization. Biopolymers 13, 29–61. https://doi.org/10.1002/bip.1974.360130103

Maiti, A., Drohat, A.C., 2011. Thymine DNA Glycosylase Can Rapidly Excise 5-Formylcytosine and 5-Carboxylcytosine: POTENTIAL IMPLICATIONS FOR ACTIVE DEMETHYLATION OF CpG SITES. J. Biol. Chem. 286, 35334–35338. https://doi.org/10.1074/jbc.C111.284620

Manley, S., Gillette, J.M., Patterson, G.H., Shroff, H., Hess, H.F., Betzig, E., Lippincott-Schwartz, J., 2008. High-density mapping of single-molecule trajectories with photoactivated localization microscopy. Nat Methods 5, 155–157. https://doi.org/10.1038/nmeth.1176

Martin, R.M., Rino, J., Carvalho, C., Kirchhausen, T., Carmo-Fonseca, M., 2013. Live-Cell Visualization of Pre-mRNA Splicing with Single-Molecule Sensitivity. Cell Reports 4, 1144–1155. https://doi.org/10.1016/j.celrep.2013.08.013

Mazza, D., Abernathy, A., Golob, N., Morisaki, T., McNally, J.G., 2012. A benchmark for chromatin binding measurements in live cells. Nucleic Acids Research 40, e119–e119. https://doi.org/10.1093/nar/gks701

McNally, J.G., 2000. The Glucocorticoid Receptor: Rapid Exchange with Regulatory Sites in Living Cells. Science 287, 1262–1265. https://doi.org/10.1126/science.287.5456.1262
McSwiggen, D.T., Hansen, A.S., Teves, S.S., Marie-Nelly, H., Hao, Y., Heckert, A.B., Umemoto, K.K., Dugast-Darzacq, C., Tjian, R., Darzacq, X., 2019. Evidence for DNA-mediated nuclear compartmentalization distinct from phase separation. eLife 8, e47098. https://doi.org/10.7554/eLife.47098

Meyer, B., Bénichou, O., Kafri, Y., Voituriez, R., 2012. Geometry-Induced Bursting Dynamics in Gene Expression. Biophysical Journal 102, 2186–2191. https://doi.org/10.1016/j.bpj.2012.03.060

Mir, M., Stadler, M.R., Ortiz, S.A., Hannon, C.E., Harrison, M.M., Darzacq, X., Eisen, M.B., 2018. Dynamic multifactor hubs interact transiently with sites of active transcription in Drosophila embryos. Elife 7. https://doi.org/10.7554/eLife.40497

Miura, K. 2005. Analysis of FRAP curves. www.embl.de/eamnet/frap/FRAP6.html

Monnier, N., Barry, Z., Park, H.Y., Su, K.-C., Katz, Z., English, B.P., Dey, A., Pan, K., Cheeseman, I.M., Singer, R.H., Bathe, M., 2015. Inferring transient particle transport dynamics in live cells. Nat Methods 12, 838–840. https://doi.org/10.1038/nmeth.3483

Morisaki, T., Müller, W.G., Golob, N., Mazza, D., McNally, J.G., 2014. Single-molecule analysis of transcription factor binding at transcription sites in live cells. Nat Commun 5, 4456. https://doi.org/10.1038/ncomms5456

Mueller, F., Stasevich, T.J., Mazza, D., McNally, J.G., 2013. Quantifying transcription factor kinetics: At work or at play? Critical Reviews in Biochemistry and Molecular Biology 48, 492–514. https://doi.org/10.3109/10409238.2013.833891

Mulholland, C.B., Ryan, J., Qin, W., Bartoschek, M.D., Traube, F.R., Parsa, E., Modic, M., Nixdorf, D., Ziegenhain, C., Carell, T., Enard, W., Bultmann, S., Leonhardt, H., 2018a. TET1 drives global DNA demethylation via DPPA3-mediated inhibition of maintenance methylation. bioRxiv 321604. https://doi.org/10.1101/321604

Mulholland, C.B., Smets, M., Schmidtmann, E., Leidescher, S., Markaki, Y., Hofweber, M., Qin, W., Manzo, M., Kremmer, E., Thanisch, K., Bauer, C., Rombaut, P., Herzog, F., Leonhardt, H., Bultmann, S., 2015. A modular open platform for systematic functional studies under physiological conditions. Nucleic Acids Res 43, e112–e112. https://doi.org/10.1093/nar/gkv550

Mulholland, C.B., Traube, F.R., Parsa, E., Eckl, E.-M., Schönung, M., Modic, M., Bartoschek, M.D., Stolz, P., Ryan, J., Carell, T., Leonhardt, H., Bultmann, S., 2018b. Distinct and stage-specific contributions of TET1 and TET2 to stepwise cytosine oxidation in the transition from naive to primed pluripotency (preprint). Cell Biology. https://doi.org/10.1101/281519

Müller, B.K., Zaychikov, E., Bräuchle, C., Lamb, D.C., 2005. Pulsed Interleaved Excitation. Biophysical Journal 89, 3508–3522. https://doi.org/10.1529/biophysj.105.064766

Müller, U., Bauer, C., Siegl, M., Rottach, A., Leonhardt, H., 2014b. TET-mediated oxidation of methylcytosine causes TDG or NEIL glycosylase dependent gene reactivation. Nucleic Acids Res 42, 8592–8604. https://doi.org/10.1093/nar/gku552

Münzel, M., Globisch, D., Brückl, T., Wagner, M., Welzmiller, V., Michalakis, S., Müller, M., Biel, M., Carell, T., 2010. Quantification of the Sixth DNA Base Hydroxymethylcytosine in the Brain. Angewandte Chemie International Edition 49, 5375–5377. https://doi.org/10.1002/anie.201002033

Nakamura, T., Arai, Y., Umehara, H., Masuhara, M., Kimura, T., Taniguchi, H., Sekimoto, T., Ikawa, M., Yoneda, Y., Okabe, M., Tanaka, S., Shiota, K., Nakano, T., 2007. PGC7/Stella protects against DNA demethylation in early embryogenesis. Nat Cell Biol 9, 64–71. https://doi.org/10.1038/ncb1519

Normanno, D., Boudarène, L., Dugast-Darzacq, C., Chen, J., Richter, C., Proux, F., Bénichou, O., Voituriez, R., Darzacq, X., Dahan, M., 2015. Probing the target search of DNA-binding proteins in mammalian cells using TetR as model searcher. Nat Commun 6, 7357. https://doi.org/10.1038/ncomms8357

Okano, M., Bell, D.W., Haber, D.A., Li, E., 1999. DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. Cell 99, 247–257. https://doi.org/10.1016/S0092-8674(00)81656-6

Ovesný, M., Křížek, P., Borkovec, J., Švindrych, Z., Hagen, G.M., 2014. ThunderSTORM: a comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging. Bioinformatics 30, 2389–2390. https://doi.org/10.1093/bioinformatics/btu202

Padmanabhan, P., Martínez-Mármol, R., Xia, D., Götz, J., Meunier, F.A., 2019. Frontotemporal dementia mutant Tau promotes aberrant Fyn nanoclustering in hippocampal dendritic spines. eLife 8, e45040. https://doi.org/10.7554/eLife.45040

Pantier, R., Tatar, T., Colby, D., Chambers, I., 2019. Endogenous epitope-tagging of Tet1 , Tet2 and Tet3 identifies TET2 as a naïve pluripotency marker. Life Sci. Alliance 2, e201900516. https://doi.org/10.26508/lsa.201900516

Pastor, W.A., Aravind, L., Rao, A., 2013. TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. Nat Rev Mol Cell Biol 14, 341–356. https://doi.org/10.1038/nrm3589

Pastor, W.A., Pape, U.J., Huang, Y., Henderson, H.R., Lister, R., Ko, M., McLoughlin, E.M., Brudno, Y., Mahapatra, S., Kapranov, P., Tahiliani, M., Daley, G.Q., Liu, X.S., Ecker, J.R., Milos, P.M., Agarwal, S., Rao, A., 2011. Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. Nature 473, 394–397. https://doi.org/10.1038/nature10102

Pavani, S.R.P., Thompson, M.A., Biteen, J.S., Lord, S.J., Liu, N., Twieg, R.J., Piestun, R., Moerner, W.E., 2009. Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. Proceedings of the National Academy of Sciences 106, 2995–2999. https://doi.org/10.1073/pnas.0900245106

Persson, F., Lindén, M., Unoson, C., Elf, J., 2013. Extracting intracellular diffusive states and transition rates from single-molecule tracking data. Nat Methods 10, 265–269. https://doi.org/10.1038/nmeth.2367

Phair, R.D., Gorski, S.A., Misteli, T., 2003. Measurement of Dynamic Protein Binding to Chromatin In Vivo, Using Photobleaching Microscopy, in: Methods in Enzymology. Elsevier, pp. 393–414. https://doi.org/10.1016/S0076-6879(03)75025-3

Phair, R.D., Misteli, T., 2000. High mobility of proteins in the mammalian cell nucleus. Nature 404, 604–609. https://doi.org/10.1038/35007077

Raccaud, M., Friman, E.T., Alber, A.B., Agarwal, H., Deluz, C., Kuhn, T., Gebhardt, J.C.M., Suter, D.M., 2019. Mitotic chromosome binding predicts transcription factor properties in interphase. Nat Commun 10, 487. https://doi.org/10.1038/s41467-019-08417-5

Rasmussen, K.D., Berest, I., Keβler, S., Nishimura, K., Simón-Carrasco, L., Vassiliou, G.S., Pedersen, M.T., Christensen, J., Zaugg, J.B., Helin, K., 2019. TET2 binding to enhancers facilitates transcription factor recruitment in hematopoietic cells. Genome Res. 29, 564–575. https://doi.org/10.1101/gr.239277.118

Reik, W., 2001. Epigenetic Reprogramming in Mammalian Development. Science 293, 1089–1093. https://doi.org/10.1126/science.1063443

Rhodes, J., Mazza, D., Nasmyth, K., Uphoff, S., 2017. Scc2/Nipbl hops between chromosomal cohesin rings after loading. eLife 6, e30000. https://doi.org/10.7554/eLife.30000

Riggs, A.D., Bourgeois, S., Cohn, M., 1970. The lac represser-operator interaction. Journal of Molecular Biology 53, 401–417. https://doi.org/10.1016/0022-2836(70)90074-4
Rust, M.J., Bates, M., Zhuang, X., 2006. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). Nat Methods 3, 793–796. https://doi.org/10.1038/nmeth929

Saitou, M., Kagiwada, S., Kurimoto, K., 2012. Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells. Development 139, 15–31. https://doi.org/10.1242/dev.050849

Saxton, M.J., 2019. Diffusion of DNA-binding species in the nucleus: A transient anomalous subdiffusion model (preprint). Biophysics. https://doi.org/10.1101/742171

Schermelleh, L., Haemmer, A., Spada, F., Rösing, N., Meilinger, D., Rothbauer, U., Cardoso, M.C., Leonhardt, H., 2007. Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation. Nucleic Acids Research 35, 4301–4312. https://doi.org/10.1093/nar/gkm432

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P., Cardona, A., 2012. Fiji: an open-source platform for biological-image analysis. Nat Methods 9, 676–682. https://doi.org/10.1038/nmeth.2019

Schmidt, J.C., Zaug, A.J., Cech, T.R., 2016. Live Cell Imaging Reveals the Dynamics of Telomerase Recruitment to Telomeres. Cell 166, 1188-1197.e9. https://doi.org/10.1016/j.cell.2016.07.033

Schneider, K., Fuchs, C., Dobay, A., Rottach, A., Qin, W., Wolf, P., Álvarez-Castro, J.M., Nalaskowski, M.M., Kremmer, E., Schmid, V., Leonhardt, H., Schermelleh, L., 2013. Dissection of cell cycle–dependent dynamics of Dnmt1 by FRAP and diffusion-coupled modeling. Nucleic Acids Research 41, 4860–4876. https://doi.org/10.1093/nar/gkt191

Schrimpf, W., Barth, A., Hendrix, J., Lamb, D.C., 2018. PAM: A Framework for Integrated Analysis of Imaging, Single-Molecule, and Ensemble Fluorescence Data. Biophysical Journal 114, 1518–1528. https://doi.org/10.1016/j.bpj.2018.02.035

Schwille, P., Haupts, U., Maiti, S., Webb, W.W., 1999. Molecular Dynamics in Living Cells Observed by Fluorescence Correlation Spectroscopy with One- and Two-Photon Excitation. Biophysical Journal 77, 2251–2265. https://doi.org/10.1016/S0006-3495(99)77065-7

Shaner, N.C., Lambert, G.G., Chammas, A., Ni, Y., Cranfill, P.J., Baird, M.A., Sell, B.R., Allen, J.R., Day, R.N., Israelsson, M., Davidson, M.W., Wang, J., 2013. A bright monomeric green fluorescent protein derived from Branchiostoma lanceolatum. Nat Methods 10, 407–409. https://doi.org/10.1038/nmeth.2413

Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T.A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K., Tajima, S., Mitsuya, K., Okano, M., Koseki, H., 2007. The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature 450, 908–912. https://doi.org/10.1038/nature06397

Shen, L., Wu, H., Diep, D., Yamaguchi, S., D'Alessio, A.C., Fung, H.-L., Zhang, K., Zhang, Y., 2013. Genome-wide Analysis Reveals TET- and TDG-Dependent 5-Methylcytosine Oxidation Dynamics. Cell 153, 692–706. https://doi.org/10.1016/j.cell.2013.04.002

Slutsky, M., Mirny, L.A., 2004. Kinetics of Protein-DNA Interaction: Facilitated Target Location in Sequence-Dependent Potential. Biophysical Journal 87, 4021–4035. https://doi.org/10.1529/biophysj.104.050765

Smith, Z.D., Meissner, A., 2013. DNA methylation: roles in mammalian development. Nat Rev Genet 14, 204–220. https://doi.org/10.1038/nrg3354

Sohni, A., Bartoccetti, M., Khoueiry, R., Spans, L., Vande Velde, J., De Troyer, L., Pulakanti, K., Claessens, F., Rao, S., Koh, K.P., 2015. Dynamic Switching of Active Promoter and Enhancer Domains Regulates Tet1 and Tet2 Expression during Cell State Transitions between Pluripotency and Differentiation. Mol. Cell. Biol. 35, 1026–1042. https://doi.org/10.1128/MCB.01172-14

Sprague, B.L., Pego, R.L., Stavreva, D.A., McNally, J.G., 2004. Analysis of Binding Reactions by Fluorescence Recovery after Photobleaching. Biophysical Journal 86, 3473–3495. https://doi.org/10.1529/biophysj.103.026765

Sproul, D., Gilbert, N., Bickmore, W.A., 2005. The role of chromatin structure in regulating the expression of clustered genes. Nat Rev Genet 6, 775–781. https://doi.org/10.1038/nrg1688

Spruijt, C.G., Gnerlich, F., Smits, A.H., Pfaffeneder, T., Jansen, P.W.T.C., Bauer, C., Münzel, M., Wagner, M., Müller, M., Khan, F., Eberl, H.C., Mensinga, A., Brinkman, A.B., Lephikov, K., Müller, U., Walter, J., Boelens, R., van Ingen, H., Leonhardt, H., Carell, T., Vermeulen, M., 2013. Dynamic Readers for 5-(Hydroxy)Methylcytosine and Its Oxidized Derivatives. Cell 152, 1146–1159. https://doi.org/10.1016/j.cell.2013.02.004

Strom, A.R., Emelyanov, A.V., Mir, M., Fyodorov, D.V., Darzacq, X., Karpen, G.H., 2017. Phase separation drives heterochromatin domain formation. Nature 547, 241–245. https://doi.org/10.1038/nature22989

Sun, W., Zang, L., Shu, Q., Li, X., 2014. From development to diseases: The role of 5hmC in brain. Genomics 104, 347–351. https://doi.org/10.1016/j.ygeno.2014.08.021

Tafvizi, A., Huang, F., Fersht, A.R., Mirny, L.A., van Oijen, A.M., 2011. A single-molecule characterization of p53 search on DNA. Proceedings of the National Academy of Sciences 108, 563–568. https://doi.org/10.1073/pnas.1016020107

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., Rao, A., 2009. Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. Science 324, 930–935. https://doi.org/10.1126/science.1170116

Tamanaha, E., Guan, S., Marks, K., Saleh, L., 2016. Distributive Processing by the Iron(II)/α-Ketoglutarate-Dependent Catalytic Domains of the TET Enzymes Is Consistent with Epigenetic Roles for Oxidized 5-Methylcytosine Bases. J. Am. Chem. Soc. 138, 9345–9348. https://doi.org/10.1021/jacs.6b03243

Teves, S.S., An, L., Hansen, A.S., Xie, L., Darzacq, X., Tjian, R., 2016. A dynamic mode of mitotic bookmarking by transcription factors. eLife 5, e22280. https://doi.org/10.7554/eLife.22280

Tinevez, J.-Y., Perry, N., Schindelin, J., Hoopes, G.M., Reynolds, G.D., Laplantine, E., Bednarek, S.Y., Shorte, S.L., Eliceiri, K.W., 2017. TrackMate: An open and extensible platform for single-particle tracking. Methods 115, 80–90. https://doi.org/10.1016/j.ymeth.2016.09.016

Tokunaga, M., Imamoto, N., Sakata-Sogawa, K., 2008. Highly inclined thin illumination enables clear single-molecule imaging in cells. Nat Methods 5, 159–161. https://doi.org/10.1038/nmeth1171

Vella, P., Scelfo, A., Jammula, S., Chiacchiera, F., Williams, K., Cuomo, A., Roberto, A., Christensen, J., Bonaldi, T., Helin, K., Pasini, D., 2013. Tet Proteins Connect the O-Linked N-acetylglucosamine Transferase Ogt to Chromatin in Embryonic Stem Cells. Molecular Cell 49, 645–656. https://doi.org/10.1016/j.molcel.2012.12.019

von Meyenn, F., Iurlaro, M., Habibi, E., Liu, N.Q., Salehzadeh-Yazdi, A., Santos, F., Petrini, E., Milagre, I., Yu, M., Xie, Z., Kroeze, L.I., Nesterova, T.B., Jansen, J.H., Xie, H., He, C., Reik, W., Stunnenberg, H.G., 2016. Impairment of DNA Methylation Maintenance Is the Main Cause of Global Demethylation in Naive Embryonic Stem Cells. Molecular Cell 62, 848–861. https://doi.org/10.1016/j.molcel.2016.04.025

Wan, J., Su, Y., Song, Q., Tung, B., Oyinlade, O., Liu, S., Ying, M., Ming, G., Song, H., Qian, J., Zhu, H., Xia, S., 2017. Methylated cis-regulatory elements mediate KLF4-dependent gene transactivation and cell migration. eLife 6, e20068. https://doi.org/10.7554/eLife.20068

Weissmann, S., Alpermann, T., Grossmann, V., Kowarsch, A., Nadarajah, N., Eder, C., Dicker, F., Fasan, A., Haferlach, C., Haferlach, T., Kern, W., Schnittger, S., Kohlmann, A., 2012. Landscape of TET2 mutations in acute myeloid leukemia. Leukemia 26, 934–942. https://doi.org/10.1038/leu.2011.326

Wen, L., Li, Xianlong, Yan, L., Tan, Y., Li, Rong, Zhao, Y., Wang, Y., Xie, J., Zhang, Y., Song, C., Yu, M., Liu, X., Zhu, P., Li, Xiaoyu, Hou, Y., Guo, H., Wu, X., He, C., Li, Ruiqiang, Tang, F., Qiao, J., 2014. Whole-genome analysis of 5-hydroxymethylcytosine and 5-methylcytosine at base resolution in the human brain. Genome Biol 15, R49. https://doi.org/10.1186/gb-2014-15-3-r49

White, M.D., Angiolini, J.F., Alvarez, Y.D., Kaur, G., Zhao, Z.W., Mocskos, E., Bruno, L., Bissiere, S., Levi, V., Plachta, N., 2016. Long-Lived Binding of Sox2 to DNA Predicts Cell Fate in the Four-Cell Mouse Embryo. Cell 165, 75–87. https://doi.org/10.1016/j.cell.2016.02.032

Williams, K., Christensen, J., Pedersen, M.T., Johansen, J.V., Cloos, P.A.C., Rappsilber, J., Helin, K., 2011. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature 473, 343–348. https://doi.org/10.1038/nature10066

Winter, R.B., Berg, O.G., Von Hippel, P.H., 1981. Diffusion-driven mechanisms of protein translocation on nucleic acids. 3. The Escherichia coli lac repressor-operator interaction: kinetic measurements and conclusions. Biochemistry 20, 6961–6977. https://doi.org/10.1021/bi00527a030

Woodcock, C.L., Ghosh, R.P., 2010. Chromatin Higher-order Structure and Dynamics. Cold Spring Harbor Perspectives in Biology 2, a000596–a000596. https://doi.org/10.1101/cshperspect.a000596

Woringer, M., Darzacq, X., 2018. Protein motion in the nucleus: from anomalous diffusion to weak interactions. Biochem. Soc. Trans. 46, 945–956. https://doi.org/10.1042/BST20170310

Wu, H., D'Alessio, A.C., Ito, S., Xia, K., Wang, Z., Cui, K., Zhao, K., Eve Sun, Y., Zhang, Y., 2011. Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. Nature 473, 389–393. https://doi.org/10.1038/nature09934

Xiang, L., Chen, K., Yan, R., Li, W., Xu, K., 2019. Super-resolution displacement mapping of unbound single molecules reveals nanoscale heterogeneities in intracellular diffusivity (preprint). Biophysics. https://doi.org/10.1101/559484

Xiong, J., Zhang, Z., Chen, J., Huang, H., Xu, Y., Ding, X., Zheng, Y., Nishinakamura, R., Xu, G.-L., Wang, H., Chen, S., Gao, S., Zhu, B., 2016. Cooperative Action between SALL4A and TET Proteins in Stepwise Oxidation of 5-Methylcytosine. Molecular Cell 64, 913–925. https://doi.org/10.1016/j.molcel.2016.10.013

Xu, C., Liu, K., Lei, M., Yang, A., Li, Y., Hughes, T.R., Min, J., 2018. DNA Sequence Recognition of Human CXXC Domains and Their Structural Determinants. Structure 26, 85-95.e3. https://doi.org/10.1016/j.str.2017.11.022

Youmans, D.T., Schmidt, J.C., Cech, T.R., 2018. Live-cell imaging reveals the dynamics of PRC2 and recruitment to chromatin by SUZ12-associated subunits. Genes Dev. 32, 794–805. https://doi.org/10.1101/gad.311936.118

Zemach, A., McDaniel, I.E., Silva, P., Zilberman, D., 2010. Genome-Wide Evolutionary Analysis of Eukaryotic DNA Methylation. Science 328, 916–919. https://doi.org/10.1126/science.1186366

Zeng, Y., Yao, B., Shin, J., Lin, L., Kim, N., Song, Q., Liu, S., Su, Y., Guo, J.U., Huang, L., Wan, J., Wu, H., Qian, J., Cheng, X., Zhu, H., Ming, G., Jin, P., Song, H., 2016. Lin28A Binds Active Promoters and Recruits Tet1 to Regulate Gene Expression. Molecular Cell 61, 153–160. https://doi.org/10.1016/j.molcel.2015.11.020

Zhang, P., Rausch, C., Hastert, F.D., Boneva, B., Filatova, A., Patil, S.J., Nuber, U.A., Gao, Y., Zhao, X., Cardoso, M.C., 2017. Methyl-CpG binding domain protein 1 regulates localization and activity of Tet1 in a CXXC3 domain-dependent manner. Nucleic Acids Research 45, 7118–7136. https://doi.org/10.1093/nar/gkx281

Zhang, W., Xia, W., Wang, Q., Towers, A.J., Chen, J., Gao, R., Zhang, Y., Yen, C.-A., Lee, A.Y., Li, Y., Zhou, C., Liu, K., Zhang, J., Gu, T.-P., Chen, X., Chang, Z., Leung, D., Gao, S., Jiang, Y.-H., Xie, W., 2016. Isoform Switch of TET1 Regulates DNA Demethylation and Mouse Development. Mol. Cell 64, 1062–1073. https://doi.org/10.1016/j.molcel.2016.10.030

Zhen, C.Y., Tatavosian, R., Huynh, T.N., Duc, H.N., Das, R., Kokotovic, M., Grimm, J.B., Lavis, L.D., Lee, J., Mejia, F.J., Li, Y., Yao, T., Ren, X., 2016. Live-cell single-molecule tracking reveals co-recognition of H3K27me3 and DNA targets polycomb Cbx7-PRC1 to chromatin. eLife 5, e17667. https://doi.org/10.7554/eLife.17667

Zhou, L., Ren, M., Zeng, T., Wang, W., Wang, X., Hu, M., Su, S., Sun, K., Wang, C., Liu, J., Yang, C., Yan, L., 2019. TET2-interacting long noncoding RNA promotes active DNA demethylation of the MMP-9 promoter in diabetic wound healing. Cell Death Dis 10, 813. https://doi.org/10.1038/s41419-019-2047-6