UNIVERSITA' VITA-SALUTE SAN RAFFAELE

CURRICULUM IN CELLULAR AND MOLECULAR PHYSIOPATHOLOGY

Role of chromatin organization and accessibility on the transcription factor search mechanism upon senescence induction.

DoS: Prof. Davide Mazza

Second Supervisor: Dr. Nacho Molina

Dr. Deborah Toiber

Tesi di DOTTORATO DI RICERCA di Tom Fillot
matr. 016070
Ciclo di dottorato XXXV
SDS BIO/11

Anno Accademico 2022/2023

Annex 2

RELEASE OF PHD THESIS

Il/la sottoscritto/a / *I, the undersigned* Tom Fillot

Matricola / *registration number* 016070

nat_ a/ *born in* Colombes

il/*on* 25/04/1994

autore della tesi di Dottorato di ricerca dal titolo / *author of the PhD Thesis titled*

Role of chromatin organization and accessibility on transcription factor search mechanism upon senescence.

☑ AUTORIZZA la Consultazione della tesi / AUTHORIZE *the public release of the thesis*

NON AUTORIZZA la Consultazione della tesi per …… mesi /DO NOT AUTHORIZE *the public release of the thesis for …… months*

a partire dalla data di conseguimento del titolo e precisamente */ from the PhD thesis date, specifically*

Dal / *from* ……/……/…… Al / *to* ……/……/……

Poiché /*because*:

☐ l'intera ricerca o parti di essa sono potenzialmente soggette a brevettabilità/ *The whole project or parts of it may be the subject of a patent application;*

☐ ci sono parti di tesi che sono già state sottoposte a un editore o sono in attesa di pubblicazione/ *Parts of the thesis have been or are being submitted to a publisher or are in the press;*

☐ la tesi è finanziata da enti esterni che vantano dei diritti su di esse e sulla loro pubblicazione/ *the thesis project is financed by external bodies that have rights over it and its publication.*

Si rende noto che parti della tesi sono indisponibili in relazione all'utilizzo di dati tutelati da segreto industriale **(da lasciare solo se applicabile)** */Please Note: some parts of the thesis are not released due to the law protecting trade secrets* ***(delete if not relevant)***

E' fatto divieto di riprodurre, in tutto o in parte, quanto in essa contenuto / *reproduction of the thesis in whole or in part is forbidden*

Data /Date 20/06/2023 Firma /Signature 

Declaration

This thesis has been

- composed by myself and has not been used in any previous application for a degree. Throughout the thesis, 'I' and 'we' are used interchageably.
- written according to the editing guidelines approved by the University.

Permission to use images and other material covered by copyright has been sought and obtained. Fig.2.2.B has been created by David Goodsell undera CC-BY license.

Some figures presented are the result of work from collaborators. This includes :

- **Generation of the BJ knock-in cell line** (Results, section 4.1) were the result of the work of Daniela Gnani from the Mazza lab in San Raffaele Scientific Institute, Milan, Italy.
- **Characterisation of p53 activity and senescence in the IMR90 ER:RAS, BJ and BJ knock-in.** (Results, section 4.1, figure 4.1, 4.2, 4.3, 4.4) were the result of the work of Daniela Gnani from the Mazza lab in San Raffaele Scientific Institute, Milan, Italy).
- **Extension of the stochastic model of volume exclusion.** The derivation of the steady-state probability and likelihood was done by my second supervisor Nacho Molina, Université de Strasbourg, France.
- **DNA FISH in senescent IMR90 ER:RAS cells** (Results, section 6.2, fig 6.7) were the result of the work of Daniela Gnani from the Mazza lab in San Raffaele Scientific Institute, Milan, Italy.
- **custom Matlab code for SIM reconstruction** (Methods, section 8.) was written by Davide Mazza and Emanuele Colombo from the Mazza lab in San Raffaele Scientific Institute, Milan, Italy.

All sources of information are acknowledged by means of reference.

Tom Fillot
Anno Accademico 2022/2023

I thank first and foremost Davide Mazza for his unwavering support throughout those three four years of PhD. His mentorship has extended beyond the academic sphere to guide me through personal challenges and I am forever grateful for it.

Daniela Gnani has been instrumental for this thesis and I thank her for her guidance, kindness and relentless efforts in trying to generate the cell-line used in this work. I hope to grow into a scientist as dedicated as she is. Further thanks are due to Matteo Mazzocca and Emanuele Colombo, for helping me with the biological and technical aspects of live-cell microscopy. Alessia Loffreda and Fulvio Bonsignore further completes a supportive team in the lab which Davide has fostered, and which has always been available to advice and improve the experiments conducted in the lab.

I would also like to acknowledge Nacho Molina for his indispensable contributions to the stochastic model of diffusion, and his kind welcome during my brief stay in his lab. The wider PEP-NET network and its members, who have been great companions through these years, and hopefully many more to come.

Abstract

The interplay between transcription factors and the chromatin landscape is a crucial element of their function, influencing gene expression and thus, cell fate. To study the underlying rules of this interplay, we use an innovative multi-modal approach, integrating super-resolution chromatin imaging with single-molecule tracking in living cells. In particular, we have focused on the behavior of the transcription factor p53, renowned for its role in tumor suppression and senescence.

Using oncogene-induced senescence (OIS) as our model system, a cellular state characterized by a massive reorganization of chromatin into senescence-associated heterochromatin foci (SAHF), we study the dynamic interplay between p53, chromatin, and SAHF. As chromatin organization has been shown to regulate gene expression, it is tempting to hypothesize a functional role for the SAHF. However, like many other features of senescence, high cell-to-cell variability makes the study of their function complicated. In the face of this heterogeneity, the inherent single-cell nature of microscopy is poised to shed light on SAHF function.

The study of p53 behavior around chromatin, and around SAHF when they are present, reveals a tug-of-war between p53's affinity for chromatin and the volume exclusion from the highly condensed regions of chromatin. Mathematical modeling of the diffusion process allows for the quantification of each and underlines the existence of an intermediate state of slow diffusion of p53 when it diffuses on the surface of SAHF. Combining immunofluorescence and smFISH in fixed samples, we show that this is correlated with an increased in the efficiency of p53 activity compared to proliferating cells, where equal levels of p53 lead to higher expression of target genes.

Thus our results suggest the SAHF do not only have a silencing role, hiding repressed genes in heterochromatin, but potentially also an enhancing one, allowing for strong expression of p53 target genes even at low p53 expression.

Table of contents

List of figures

Chapter 1

The specificity conundrum at the heart of eukaryote gene regulation.

The study of the regulation of gene expression began with the foundational description of the lactose operon by Jacob and Monod. In their farseeing review of 1961, they proposed a concept of regulator genes, coding for a factor inhibiting the production of a messenger RNA. In their bacteria, this led to the inducible expression of β-galactosidase in the presence of lactose. Jacob and Monod also anticipated that mutations in either the operator or the regulator would "alter or abolish its specific affinity for the repressor", making them an important vector for changing the cellular phenotype (Jacob and Monod, 1961). In the following decade this discovery would be refined and generalized into the model now taught to first-year biology students: RNA Polymerase II binds at the promoter, while a transcription factor binds at an enhancer and, through chromosome looping, makes contact with the polymerase which starts the elongation(Mitchell and Tjian, 1989).

Like many ideas of molecular biology developed in the mid 20th century, this would prove both broadly applicable and wholly insufficient. Entire fields of study have emerged out of the dissection of every step in transcription, with many details of the process revealing themselves to be yet another layer to control fine-tuning transcription programs and, ultimately, to determine cell fate. For RNA Pol II, a non-exhaustive list includes the recruitment to the promoter, proximal pausing, elongation rates and splicing. Chromatin itself is an actor of the process, helping or preventing the recruitment of the transcription machinery, as well as helping or preventing the contact between the enhancer and the promoter. New actors, such as non-coding RNAs, Mediator, histone readers and writers have been identified and shown to be essential to transcription. Far from being details, all of these have been shown to be relevant to both development and pathology (T. I. Lee and Young, 2013), providing new therapeutic targets for medicine. As the first step in the chain, transcription factors remain a core part of our understanding of gene regulation. However,

as research moved from prokaryote to eukaryote, more and more evidence has emerged calling into question the role of specific binding in eukaryotes.

In this first chapter, I will argue that viewing gene regulation through the lens of a specific interaction between a transcription factor and a DNA sequence is insufficient in eukaryotes. To account for the deficiencies, we will explore the role of weak multivalent interactions and compartmentalization, and how they can influence transcription factor function. In particular, we will see how they affect the target search mechanism of transcription factors, meaning how transcription factors find an "operator" among the entire DNA.

1.1 Eukaryotes regulate gene expression in a different way than prokaryote

Sometimes used as a loose term for any protein capable of altering gene expression, the more rigorous definition of a transcription factor is a protein capable of (1) binding DNA in a sequence-specific manner and (2) regulating transcription (Lambert et al., 2018). They are instrumental in coordinating gene expression, to specify cell type and create patterns during development, to the point that ectopic expression of some 'master regulators' can cause de-differentiation, or trans-differentiation (Takahashi and Yamanaka, 2016). In humans, they amount to 1,639 genes, meaning about 8% of the genome is dedicated to the orchestration of the rest of the genome (Lambert et al., 2018). This crucial importance for cell fate is underlined by the close relationship across eukaryotes between the number of cell types and the number of transcription factors in the genome (Mendoza et al., 2013).

On a very general level, transcription factors are composed of at least one DNA-binding domain and one effector domain, a combination that allows segregating the action of the effector domain at a particular loci of the DNA.

The effector domain provides a first contrast between prokaryotes and eukaryotes. In prokaryotes transcription factors that act as repressors, typically have non-functional effector domain - with the sole scope of blocking the binding of other proteins, while transcriptional activators have effector domains that directly interact with RNA polymerase, leading to its recruitment (Lambert et al., 2018). In eukaryotes, however, transcription factors act by recruiting a host of cofactors. The intermediation of cofactors allows multicellular life to achieve combinatorial gene regulation (Reiter et al., 2017), where the action of a transcription factor can change depending on cofactor availability (Ernst et al.,

2016). This is made outright necessary by the higher number of genes, and – for metazoans and embryophytes – by the complexity required by embryonic development.

DNA-binding domains are responsible for recognizing a specific sequence and binding it. Typically, a transcription factor affinity for its specific sequence is 1000-fold greater than for non-specific sequence, sometimes as high as 10^7 times higher(Spolar and Record, 1994). DNA-binding domains are extremely conserved (Weirauch et al., 2014). In fact, the entire catalog of around 100 classes of DNA-binding domain is thought to derive from a small set in the eukaryote common ancestor (Mendoza et al., 2013). Their physiological role is also often conserved, like the HOX proteins, which determine the anterior-posterior body plan in all bilateria. The sequence of DNA bound by the DNA-binding domain, commonly called its motif, is similarly well conserved: nearly all motifs present in the human genome are also recognized by fruit fly transcription factors (Nitta et al., 2015). In other word, these motifs and their transcription factors have been in a relationship for the past 700 million years, before even the Cambrian.

Upon further inspection, this relationship is less functional than one might expect.

The advent of ChIP-seq has given us the ability to inspect *in vivo* the binding profile of a transcription factor across the entire genome. It has revealed a tenuous overlap between binding sites and the presence of a motif. The majority of sites bound by a transcription factor do not contain sequences with obvious matches to motifs determined *in vitro* and, conversely, most motifs are not bound by their transcription factor(Lambert et al., 2018). Moreover, the majority of observed binding sites do not affect the expression of nearby genes (Cusanovich et al., 2014). This apparent lack of correspondence between motifs and gene regulation led to the coining, in 2004, of the futility theorem: 99.9% of all specific sequences for a transcription factor are never bound and have no functional role as operators (Wasserman and Sandelin, 2004).

There are obvious – and important – confounds which could lead to this discrepancy. As ChIP-seq uses cross-linking, it detects indirect binding, which could explain the relative lack of specific motif in some binding sites. ChIP is also very sensitive to antibody cross-reaction, and antibodies of sufficient quality do not exist for many transcription factors (Lambert et al., 2018). Besides these technical concerns, a more fundamental difference between the techniques lies in the fact that motifs are determined *in vitro*, and many properties of chromatin, like its accessibility, could be used to partially explain the difference between theoretical predictions and experimental results. However, both motifs determination and ChIP-seq have been improved upon considerably, and the futility theorem still stands.

Thus the simple picture of specific binding between a cis-regulatory element and a trans-acting regulator, reminiscent of Fisher's lock-and-key model in enzymology, breaks down in eukaryotes. It is valid in prokaryotes, where there is a near identity between a motif, a binding site and an enhancer. In eukaryotes, the specificity of a sequence is a poor predictor of its function. Indeed, in many respects, it is useful to think of the strategies for gene regulation in prokaryotes and eukaryotes as distinct.

1.2 Transcription factors and DNA, an On-and-Off relationship.

The fundamental reason underlying why the prokaryote paradigm of strong, specific binding cannot work in eukaryote comes down simply to the much higher amount of DNA. In their review, Lambert and colleagues call the futility theorem 'expected in retrospect'. Indeed, an argument from information theory shows that the short motifs recognized by transcription factors – typically below 12 base pairs – are enough to specify a single locus within the 10^6 base pairs of a prokaryotic genome, but not enough within the 10^9 base pairs of an eukaryote (Wunderlich and Mirny, 2009).

In other words, for the prokaryotic paradigm of specific binding to work in the enormous genome of an eukaryote, we would expect the motifs recognized by transcription factors to be significantly longer than in prokaryotes. Instead, the exact opposite is observed, motifs in eukaryotes typically have less information content than in prokaryotes ; they are typically shorter and degenerate – meaning that transcription factors are tolerant to mismatches compared to their cognate motifs. With such motifs, in the billions of base-pairs of an eukaryotic genome, information theory predicts thousands of false hits throughout the DNA sequence (Wunderlich and Mirny, 2009). Accordingly, virtually all transcription factors studied so far do not bind the vast majority of their motifs in the genome, meaning additional information beside sequence come into play to determine which loci are functional. An interesting comparison of transcription factors repertoire across species suggest some structural constraint is at play, as the size of DNA-binding domains is remarkably constant in the entire tree of life (Charoensawan et al., 2010). The family of C2H2 zinc-finger transcription factors works around this limitation by chaining several DNA-binding domains in sequence, leading to increased DNA footprint and binding specificity. A prominent member of this family is the architectural protein CTCF, which represents the only clear example of a DNA-binding protein which seems to follow the prokaryotic paradigm as it occupies most of its 14,000 motifs in the human genome (Fu et al., 2008).

This lack of binding specificity is hard to square-off with the evident functional specificity of transcription factors. For example, the DNA-binding domains of Hox proteins, transcription factors crucial in morphogenesis, is crucial to their function. Swapping them – even between closely related ones – leads to severe homeotic transformations (Zhao and Potter, 2002).

If a single binding domain is insufficient to warrant binding specificity in the context of the eukaryotic genome, cooperative effects, such as the formation of complexes, can help bring back some specificity. Many transcription factors bind as homo- or heterotypic dimers, which could be understood as a way to multiply the motif's length and thus to increase the information content. For example, the tumor suppressor p53, on which we will focus later, contains a tetramerisation domain, allowing it to recognize an unusually long motif of 20 base pairs (Veprintsev et al., 2006). The interferon-β enhanceosome is a large complex of 8 factors which assemble in a strict manner on a highly optimized 50 base-pair long sequence (Panne, 2008). However, this type of highly ordered binding sites are by far the exception ; eukaryotic enhancers are tolerant to changes in the orientation or spacing between motifs, going against the idea of highly specific interactions. Cooperation between transcription factor can also be mediated by the DNA. Either via vibrational modes or by altering DNA shape, the binding of one transcription factor can increase the affinity of others (Lambert et al., 2018). This kind of allosteric interaction can lead to latent specificity, where the pair of transcription factor recognizes a DNA sequence that is different from the juxtaposition of their own two motifs.

However, cooperative complexes cannot rescue the prokaryote paradigm of specific binding. This is apparent when looking at the problem from the other side, that of the enhancer. As already mentioned, the catalog of DNA-binding domain families has remained constant, at least since the evolution of metazoans. Diversification has come, instead, from the expansion of a few families. The human repertoire is dominated by two families of proteins, the C2H2 zinc-finger and the homeodomain families, together accounting for over 1000 transcription factors – out of our 1639 (Lambert et al., 2018). Paralogs recognize sequences very similar to each other ; a problem compounded by the degeneracy of eukaryotic motifs. This makes enhancers very promiscuous: if a mutation in cis-regulatory element increases the affinity for a transcription factor, it will also increase its affinity for all its family members (Kribelbauer et al., 2019). An important consequence of this promiscuity is the high-affinity binding sites are not suited for the control of paralog-specific processes, since having high-affinity for one transcription factor means the regulatory element also has high-affinity to all its paralogs. Accordingly, experiments have shown that replacing low-affinity binding sites with high-affinity ones leads to ectopic gene activation (Farley et al., 2015), which suggests that paralogs in different tissues ignore the low-affinity binding sites, but will occupy high-affinity ones. Conversely, this means that

high-affinity binding sites are not suited for tissue-specific expression, and are typically seen for cell-autonomous processes.

For the prokaryotic paradigm, the final nail in the coffin comes from direct measurement of live-cell binding kinetics of transcription factors interacting with DNA, using techniques such as single-molecule tracking or fluorescence recovery after photobleaching. Instead of long binding events, most transcription factors in eukaryotes probed by these techniques appear to rapidly move on and off DNA, with binding times typically in the order of 1s. Even when these kinetics are probed at engineered genetic loci containing repetitions of the binding site, the residence time of the transcription factor on DNA does not exceed 10s, much shorter than residence times at specific binding sites of bacterial TFs, on the timescale of tens of minutes (Lammers et al., 2020).

These observations are compatible with a functional relevance in eukaryotes of low-affinity binding site – given equal association rate, lower affinity means lower residence time. It might also explain why some sites with high occupancy measure by ChIP-seq do not appear as protected from DNAse in DNase-seq assays (Sung et al., 2014): dynamic binding and unbinding obscures the footprint expected from a stably bound transcription factor.

Thus the prokaryotic paradigm of gene regulation, with at its core a single strong relationship between one cis-regulatory element and one transcription factor, does not hold for eukaryotes, not from a theoretical perspective nor in the data. Instead, the new paradigm that emerges in eukaryotes relies weak multivalent cooperative interactions.

1.3 Rescuing functional specificity from a sea of weak interactions.

Letting go of specificity is not trivial, from the conceptual standpoint. Not only is the lock-and-key model ingrained in the mind of molecular biologists, weak – thus short-lived – interactions seem at first glance incompatible with the timescale at which the rest of transcription takes place. Once bound, the enhancer needs to contact the promoter, a process that might take minutes. One burst of transcription also lasts in the order of minutes – up to several hours (Lammers et al., 2020). A strongly bound transcription factor is, quite literally, an anchor for the initiation of transcription, which can stay in place long enough to recruit the many co-factors needed to set this whole process in motion. But these timescales are not easily understood when the transcription factor is gone within seconds.

The current proposed answer to this conundrum lies in the formation of compartments, whose purpose is to increase the local concentration of transcription factor. Instead of one long-lasting bounding of a single molecule, these compartments would insure frequent

visits by many transcription factor molecules, leading overall to a significant occupancy of the enhancer. The theme of sub-nuclear compartmentalization of eukaryotic nuclei has emerged as a major theme in the last decade, and many of the discovered structures have been proposed to help in transcription.

In line with this model, regulatory elements in eukaryotes frequently contain multiple weak binding sites close to each other, instead of their cognate motif. Figure 1.1 illustrate how such binding site can be brought in physical proximity. As stated previously, these generally do not work as enhanceosome. They are robust to changes in the exact arrangement of the binding sites, even though synthetic biology studies have shown that spacing and orientation can influence transcription rates, hinting at an underlying syntax (Ezer et al., 2014). The decoding of this syntax is made more complicated by long-distance interactions between multiple enhancers, spreading regulatory function to multiple loci. Indeed, chromatin can fold in 3D and bring several regulatory elements in close proximity, forming enhancer communities (Madsen et al., 2020). These clusters are also sometimes described as super-enhancers when they are associated with an large enrichment in Mediator1 or p300 binding (Pott and Lieb, 2015). In either case, these cluster generally span a region that is one order of magnitude greater than the enhancers as we are used to conceptualize them – from a median of 703 to 8667 bp in mouse embryonic stem cells (Hnisz et al., 2013) – a clear hurdle in our ability to decode their function from the DNA sequence. The 3D architecture of chromatin is on its own a major theme relating to our question. It is arranged in nested compartment and while the relationship between transcription and chromatin architecture suffers from a 'chicken-and-egg' situation, there is evidence of some degree of control over which distal elements are brought together or to the promoter. A variety of architectural proteins, such as the previously mentioned CTCF, cells are able to control how chromatin folds onto itself, bringing distant elements together and, perhaps more importantly, preventing some response elements and promoters from coming into contact (Chakraborty et al., 2023).

While it is possible to see chromatin as the scaffold of compartmentalization, transcription factors themselves, crucially, are able to form clusters of high local concentration, *in vitro* and *in vivo*. These so-called condensates form without the need for a lipid membrane, via weak multivalent interactions. It is in condensates that the notion of weak multivalency becomes central, as it gives them properties unlike both classical compartments and stable complexes. Composed of tens to several hundreds molecules, they form spontaneously when the concentration of factors in the nucleus crosses a given threshold, and dissipate when it goes back down, allowing for dynamical assembly and disassembly of these structures (Sabari et al., 2018). Because they remain liquid, molecules inside the condensate still diffuse. Molecules can also diffuse in and out of the phase, but at an energy cost. This

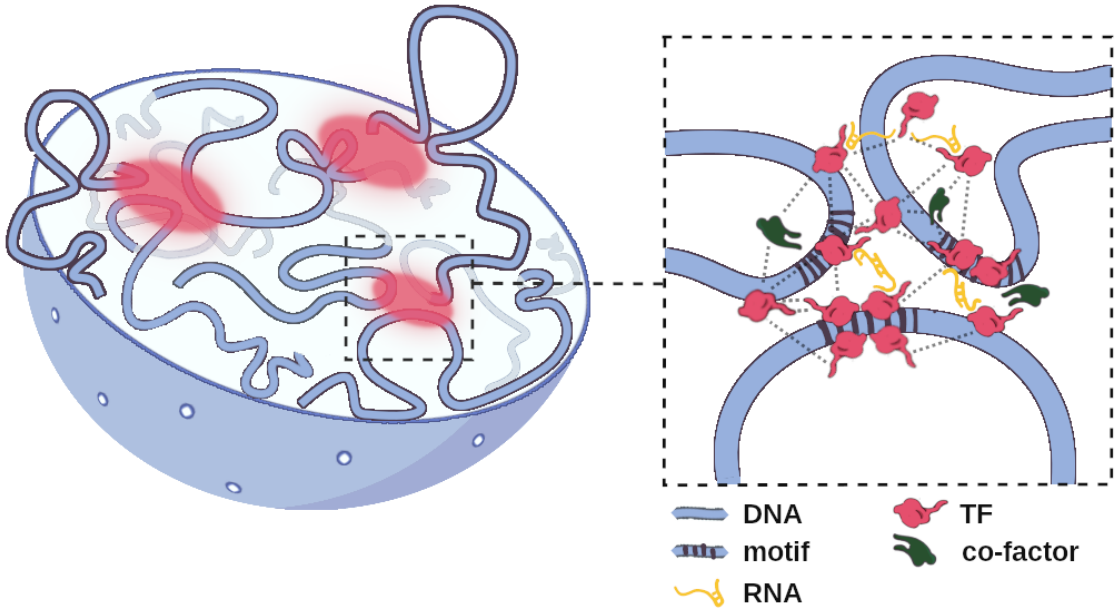

Fig. 1.1 **Weakly interacting molecules can create transient compartments crucial for gene expression. Left:** A cell is represented with its chromatin folded inside. Red circles represent compartments.. **Right:** Detail of the molecular composition of a compartment. A number of TFs interact weakly, either with co-factors such as the Mediator complex, within themselves directly or with RNA molecules. Several distant binding sites may be brought together and serve as a nucleation point for the transient compartment.

last property is crucial for solving our temporal problem in transcription, because it allows us to understand phase separated compartment as trapping zones. Transcription factors may bind transiently, but be forced to revisit multiple time the same portion of DNA, separated as they are from the rest of the nucleus (fig.1.1. As theory shows, bridging the time-scales of the different steps in transcription requires multiple rate-limiting steps, and this confinement could be one of them (Lammers et al., 2020). Another consequence of this confinement, which we will return to later, is to ensure that a binding site is effectively found by the transcription factor, by limiting the portions of DNA it needs to explore to those present in the condensate. It is important to stress that the role of condensates in transcription is heavily debated, and local accumulations of molecules have been observed without requiring phase separation(McSwiggen et al., 2019).

Thus we have replaced the stable anchor of the DNA-transcription factor bound by an ensemble of loosely tethered factors and DNA elements, which cluster together long enough for the following step of transcription to take place. But the interactions described so far are not necessarily specific. For weak binding sites, there is by definition less of a contrast between their affinity and the affinity of the surrounding non-specific chromatin. An argument can be made that the sequence specificity acts as a nucleation event, starting the cluster at least preferentially around the enhancer, but liquid-liquid phase separation

is a physical process, and a priori would not discriminate between different types of factors.

To understand how clustering brings specificity back, we need to focus further on the nature of these multivalent interactions. In particular, I will describe the two main mechanisms through which they take place, namely RNA molecules and Intrinsically Disordered Domains. I will also describe how a model of 'spacers-and-stickers' (Martin et al., 2020) provides a general understanding for how these weak multivalent interaction take place.

RNA molecules are long arrays of potentially hydrogen-bound forming nucleotides, well-fitted for weak multivalent interactions. As such, they are capable of forming condensate by themselves *in vivo*, and are found in condensates *in vitro*. For molecular biologists, they also provides a very natural mechanism for specificity, through Watson-Crick pairing. It is, however, largely a misleading image for how RNA molecules interact with each others. RNA-RNA interactions through Watson-Crick pairing tend to involve short sections of each molecules instead of one long stretch. Moreover, RNA-RNA interactions can be mediated by other mechanisms, including non-Watson-Crick base pairs, and even interactions between sugar edges (Bevilacqua et al., 2022). Another important consideration are the secondary structures which RNA can form: because these tend to be strong, they leave fewer bases available for RNA-RNA interactions (Roden and Gladfelter, 2021). All in all, these considerations mean the specific sequence of an RNA molecule incorporated into a condensate is far from explaining the molecular composition of a condensate. Accordingly, Henninger et al. (2021) provides a mechanistic function to enhancerRNA that is entirely independent of their sequence, where the presence of RNA, a negatively charged molecule, simply acts as an electrolyte which facilitates the recruitment of the positively charged DNA-interacting molecules – including transcription factors (fig 1.2). This is not to say that the identity of the RNA molecule has no impact on the properties and composition of the condensate. For example, (Langdon et al., 2018) provides evidence of how different RNA secondary structure can lead to the formation of different scaffolds, leading to the recruitment of different proteins and even providing a rationale for why all condensate don't coalesce into a single droplet. (Cerase et al., 2022) shows how reorganizing the sequence of Xist leads to a reorganization of the structure of Xist condensates around the inactive X chromosome in female. In summary, even if it doesn't rely on sequence-specific interactions between the bases of the RNA, RNA molecules can bring about specificity in clusters of molecules through multiple interactions engaging different parts of the molecule.

Intrinsically Disordered Regions are protein domains that do not fold into a stable structure. Due to the prevalence of crystallography in our early understanding of protein function, unstructured regions remained poorly understood for many years (Dyson and

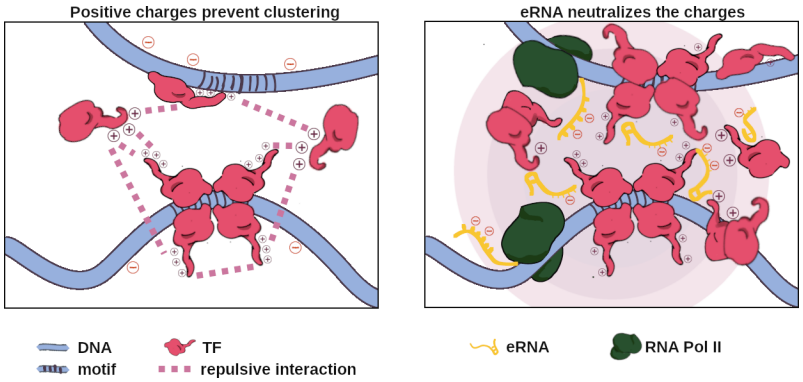

Fig. 1.2 **Enhancer RNA (eRNA) can help the formation of condensates. Left:** In the absence of RNA, the positively charged TFs cannot accumulate to a high degree due to electrostatic interactions. **Right:** The binding of TFs often leads to bi-directional transcription of the enhancer. The resulting eRNA counter-act the positive charges, and can lead to an increased local concentration of TF.

Wright, 2005). Their distinctive amino-acid composition, enriched in polar amino-acids and with aromatic residues punctuated throughout the region suggested a function however. Within the spacers-and-stickers model (Martin et al., 2020), several patterns in IDR have been identified as 'stickers' able to engage in specific interactions with other molecules. In this context, the specific sequence of the IDR doesn't matter so much as the stickers valence (i.e. how many there are on a single protein), patterning and interaction strength, all of which can drive interaction with different partners. Importantly for clustering, IDR are often necessary and sufficient to drive liquid-liquid phase separation *in vitro* (Vernon and Forman-Kay, 2019). Within IDR, Phe-Gly repeats have been identified as an important component, serving as a prominent 'sticker', because the aromatic rings can engage in cation-pi interactions (Nott et al., 2015). In this regard, (Lyons et al., 2023) is a recent study in which researchers changed the patterning of 'stickers', in the form of charged blocks within the IDR of MED1, and observed changes in which regulators of transcription partitioned inside MED1 condensates. Given this role of IDR, it is striking that transcription factors are twice as likely as the rest of the proteome to contain IDR of more than 50 amino-acids.

Importantly, these weak multivalent interactions can be heterotypic. RNA and IDR can interact with one another, as has been shown for example in the case of YTHDF1/2/3. The IDR of these three proteins have been shown to engage in a stronger manner with mRNAs which contain multiple instances of a specific nucleotide modification, N6-methlyadenosine (Ries et al., 2019), providing yet another avenue of specificity in the creation of these clusters of molecules.

As we see, both RNA and IDR's role in creating clusters of molecules can be seen through the stickers-and-spacers model, and in both case some mechanisms are identified that are able to selectively recruit some co-factors rather than others.

I will point out, however, that we started this section discussing the specificity between transcription factors and DNA-regulatory elements, but we switched to a different kind of specificity along the way. RNA and IDR can explain the specific attraction of some co-factors and proteins, which is an important step in explaining how enhancers can still be functional even with transcription factors binding for a short duration. But our more fundamental problem starts with (Wunderlich and Mirny, 2009), showing that the short and degenerate eukaryotic motifs should produce up to 10^5 spurious binding sites. The big question left unanswered is: why are some canonical motifs ignored in the first place? Additionally, some proteins, such as p53, do not appear to form microscopic clusters *in vivo*, calling into question their necessity in our developing understanding of eukaryotic transcription.

In the following section, we will explore one possible part of the answer to the specificity conundrum, which lie in the speed with which transcription factor find their targets. As we will see, ensuring some portion of DNA are visited more frequently than others is another way to increase occupancy of a site.

1.4 The Target Search Mechanism: hide and seek, at the genomic level.

The target search is the first – but often forgotten – step in how a transcription factor operates. Perhaps because the textbook pictures depict molecules on a white background, we do not often question how all the necessary actors of a biological process haors, typically have non-functional effector domain - with the sole scope of blocking theppen to be in the same place at the same time. This is reasonable in many cases; either the molecules are in abundance, or they are confined in an organelle to facilitate their interaction with other, desired molecules. It is less so, however, in the case of transcription factor searching for their binding site, if only because there are few molecules of any transcription factor compared to the amount of DNA they have to explore. As we will see in this part, the target search problem resonates strongly with fundamental aspects of our new, eukaryotic paradigm of gene regulation, dominated by weak multivalent interactions. One particularly interesting example in that regard is how the specificity conundrum I outlined so far focuses on the specific binding sites, whereas the question of how a transcription factor finds its target will bring the vast excess of unspecific DNA into the forefront. In this part I will outline why the target search is a problem from a theoretical standpoint, then present how the cell can solve this problem. We will see how some of the observed and proposed

solutions can help answer the specificity conundrum

The physical theory of how chemical reaction take place is a good starting point to understand the target search is, indeed, a problem, and quantify its importance. For diffusion-limited reactions, the general theory is the the Smoluchowski equation, which can be used to derive the association rate of the transcription factor, and the expected time τ_{search} that a single molecule of the factor needs to find its target site (McSwiggen et al., 2019).

$$\tau_{search} = \frac{V}{4\pi aD} \tag{1.1}$$

Where D is the diffusion coefficient of the transcription factor, V the volume of the nucleus, and a the diameter of the target site which, for our purposes, is a few nucleotides wide. Doing the calculation, we expect search times in the order of minutes in bacteria, and days in mammalian cells. This is, of course, incompatible with observations, in humans and bacteria alike. In bacteria, it was recognized in 1970 that the association rates of the Lac repressor to its target sequence is about 1000 times faster than the association rates of of any protein-protein association rate, and that the Lac repressor reaches its target 100 times faster than the diffusion limit. Similarly in mammals TF respond in external stimuli in minutes, while the Smoluchowski equation would predict search times in the order of days (Mazzocca, Fillot, et al., 2021) ; for example, p53 can activate transcription of its target genes in less than 30 minutes (Allen et al., 2014). This discrepancy between observed association rates or, equivalently, the time it takes for a transcription factor to find its target, and the theoretical limit is the heart of the target search problem.

An immediate answer to this discrepancy could lie in copy numbers. In a single cell, there is not a single copy of a transcription factor: proteomics studies have estimated the median mammalian transcription factor at 71,000 copies per cell (J. J. Li et al., 2014). While this number qualifies transcription factors as lowly expressed proteins (Beck et al., 2011), it seems at first glance high enough to bring the search time down significantly. Until we consider that the target site is also not alone. Not only do many binding sites need to be found and occupied, there is a gigantic excess of unspecific DNA. While transcription factors generally have 1000-fold higher affinity to specific sequences vs. unspecific ones, there is a factor 10,000 in the quantity of each. Using the ratio of specific vs. unspecific sites, we can predict that 99.9% of a transcription factors time is spent engaging in unspecific interactions with DNA ; conversely, this predicts very poor occupancy of specific binding sites.

Even if it turns out to be violated, starting from theory of diffusion-limited reaction is useful, since it tells us one or several of the assumptions in the Smoluchowski equation

do not hold in living cells. This broken assumption is generally taken to be that of free 3D diffusion, also called Brownian diffusion. Instead, anomalous diffusion is taking place. That is to say, the transcription factor isn't exploring 3D space indiscriminately, but instead can be seen as exploring a restricted space of reduced dimensionality (Woringer and Darzacq, 2018). The problem with this general, theoretical answer is that many different biological phenomenon can lead to anomalous diffusion.

In bacteria, the now accepted solution is facilitated diffusion (see fig1.4). Instead of exploring in 3D all available volume isotropically, as you would expect of particles diffusing in an ideal solution, transcription factors slide in 1D along the DNA from time to time, effectively restricting the search volume. This 1D diffusion along DNA – requiring no energy expense – has been experimentally confirmed, first *in vitro* on naked DNA(Blainey et al., 2006; Gorman et al., 2007; Granéli et al., 2006), and more recently *in vivo*(Elf et al., 2007). The theoretical side of the framework has also matured, with studies showing what speed-up is expected depending on a number of observables, such as the transcription factor's diffusion coefficient while sliding and its dissociation rates with non-specific DNA – which combine into the average length of the segment of DNA explored per sliding event – the ratio between the amount of time spent diffusing in 3D or sliding in 1D (see fig1.3). For the LacI, these values have been measured *in vivo*, with a 1D diffusion coefficient of 0.05 $\mu m.s^{-2}$, and a sliding window of 45bp, and predict with accuracy the measured speed-up in search time (Elf et al., 2007; Hammar et al., 2012). Importantly, (Elf et al., 2007) found that LacI is engaging with non-specific DNA around 90% of its time. These finding have since been expending to other bacterial proteins (Esadze and Stivers, 2018; Silverstein et al., 2014; Cravens et al., 2015).

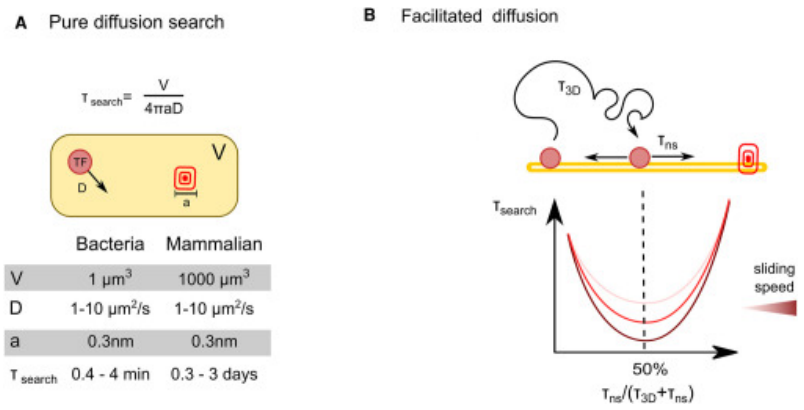

Fig. 1.3 **Facilitated diffusion help TFs reach their targets faster than the diffusion limit. A:** Expectation of search time in bacteria and mammals based on a diffusion-limited reaction. **B:** Theorety predicts that the search time is minimized for an equal time spent sliding and diffusion. From (Mazzocca, Fillot, et al., 2021).

For eukaryotes, once again, the 1000 time larger amount of DNA seemingly works against their own gene regulation. Facilitated diffusion is sensible to the amount of

unspecific DNA, and using the kinetic parameters measured in eukaryotic cells yields significantly longer search times of 500h. The presence of nucleosomes raises even more concerns, and it is unclear if transcription factors can slide between or on nucleosomes. Several human transcription factors have been shown *in vitro* to be capable of sliding on naked DNA (Tafvizi et al., 2011; Esadze and Iwahara, 2014; Zandarashvili et al., 2012), a fact that we will return in the part focusing on p53. Additionally, many transcription factors have been shown to have anomalous diffusion (Woringer and Darzacq, 2018; Pederson, 2000; Fritsch and Langowski, 2011). In particular, this is seen in mean-square displacements (MSD) curves of many mammalian proteins. Under a free-diffusion regime, it is expected that the average squared distance traveled by a randomly moving particle increases linearly with time. In many cases, some mammalian proteins – although not all – seem to be under a sub-diffusive regime ; in other words they are exploring less space than would be expected under free diffusion. For example, TEFb is shown to be sub-diffusive, but the diffusion of c-Myc seems Brownian (Izeddin et al., 2014). MSD curves, however, are ill-suited for fine analysis of the behavior of molecules in eukaryotic nuclei, owing mainly to the bigger size of eukaryotic cells. Of particular concern is the fact that the focus slice, of around 800 nm, is much narrower than the entire nuclei, leaving molecules free to diffuse out of focus. This leaves big uncertainty on the distance traveled at longer time lags (Izeddin et al., 2014). As a result of the increased experimental challenges, experimental confirmation of sliding *in vivo* is much harder to obtain.

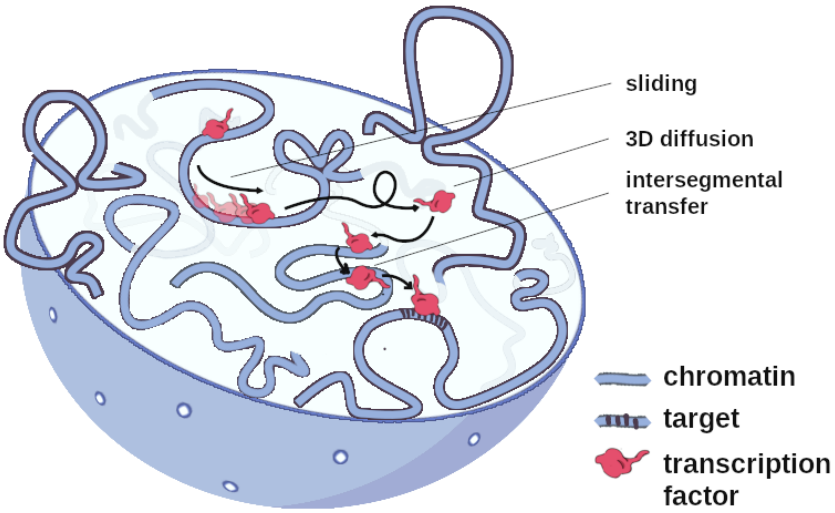

Fig. 1.4 **Facilitated diffusion helps TFs reach their targets.** To search the vast amount of DNA fast enough, TF can slide along the DNA or avoid obstacles with intersegmental transfer.

Explaining the target search in eukaryotes may thus require us to look for other mechanisms of speeding-up the search. For this, we need to dwell deeper into the reasons what makes facilitated diffusion – alternating between 1D sliding and 3D diffusion – work

on a fundamental level.

1D search is a so-called compact search: when a random-walker goes from point A to B on a 1D path, it has necessarily passed through every points in between, insuring that its target was found if it was in this segment. 3D search is a non-compact search: there's an infinity of path between A and B, and any target in space can be missed by walking around it. Thus searching in a space of reduced dimension is more likely to find its target, but its speed is highly dependent on the distance to the target: as soon as the 1D diffusion coefficient starts to be small compared to the length of DNA to explore, this becomes prohibitively slow. 3D search doesn't show this dependency on distance, but the speed of exploration comes at the cost of walking past its target without finding it. Fundamentally, facilitated diffusion works by providing a balance between those two modes of exploration, reducing the dimensionality of the search on DNA, but using 3D diffusion to skip over large linear sections of DNA, mitigating the 1D compact search dependence on distance. This general switching between compact and non-compact exploration is called guided exploration.

The alternation between compact and non-compact search is the heart of the solution to the target search problem. And broadly speaking, this amounts to periodically reducing the dimensionality of the search space. Intermittent sliding is one way to achieve this, but not the only one, and other phenomenon could cause the same behavior in eukaryotes.

The key here is, once again, to create local compartments. One of the main ways in which anomalous diffusion can manifest is the diffusion in a fractal media. The nucleus is a space packed with molecules, but an organized space nonetheless, with many sub-structures including the nucleolus and PML bodies. These many structures, in addition to chromatin organization, can affect diffusion by simple steric hindrance and create the fractal media giving rise to anomalous diffusion. The interaction of transcription factor with these compartments may not be limited to passive steric hindrance however. For example, nucleosomes can organize in clutches, and thus expose DNA on the surface of the clutch ; with diffusion on a surface being another way to reduce the dimensionality of the search. Weak multivalent interactions may also play a role here. As we have mentioned earlier, condensate have been proposed to work in this manner in a recent work (Hansen et al., 2020) ; by transiently trapping a molecule in a restricted volume, it is made to explore this volume more exhaustively than with simple free diffusion in the entire volume of the nucleus (fig1.5). Although phase condensates are emerging as a compelling mechanism to explain this backtracking behavior, 'trapping zones' need not rise to the level of a different phase (McSwiggen et al., 2019). Of course, it should be noted that these reduction in the dimensionality of the search only reduce the search time if the target is indeed contained

within the reduced search space. This is one of the reasons why phase condensate may fit well for some factors, as molecular clusters based on MED1 have been shown to form at enhancers (Boija et al., 2018; Sabari et al., 2018). Clusters, and in general multivalent weak interactions, may be a large reason for the sub-diffusive behavior observed in some proteins. And although, as we noted, some proteins do not form visible condensates – i.e. they do not diffuse into a condensate nor are kept there preferentially – interaction with the surface of said condensates may help in reducing the dimensionality of the search.

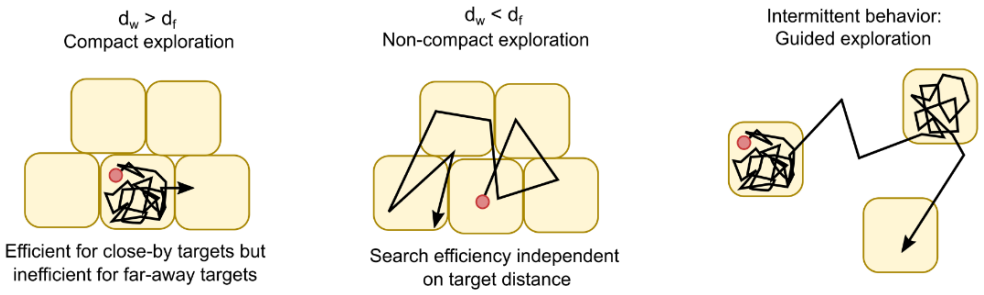

Fig. 1.5 **Compact and non-compact exploration are two strategies with different drawbacks.** Guided exploration, which comprises transiently trapping the molecules in compartments, helps by mitigating the drawbacks of both. From (Mazzocca, Fillot, et al., 2021).

As I have laid out here, there are many biological phenomenon that may give rise to anomalous diffusion and thus speed up a transcription factor's search for its target. To discriminate between them experimentally, the main experimental method is single-molecule tracking (SMT), also sometimes called single-particle tracking. This aptly named technique allows us to follow single copies of a given molecule as they diffuses through the cell, and analyze their movement over time. Various feature of the random-walk of the protein can be evidenced to discriminate between different modes of exploration and, ideally, between the different biological phenomenon that could be responsible for them. A frequent metric is the MSD curve, whose form is the defining characteristic of anomalous diffusion. The exponent of the MSD curve, generally called anomalous exponent α, directly relates to the dimension of the random-walk d_w, as $\alpha = 2/d_w$. For Brownian diffusion, $\alpha = 1$, i.e. MSD increases linearly with time. When the MSD curve increases with $\alpha < 1$, in other words when $d_w > 2$, the particle is said to be sub-diffusive. Concretely, that the dimension of the walk is above 2 means the particle goes back on its track more frequently than expected of a free diffusing particle, which increases the roughness of its path (fig.1.6). The value of d_w is crucial in determining the nature of the exploration by the transcription factor, because the search becomes compact (unlikely to miss a target, but very distance dependent), if the space available for diffusion as a dimension df is inferior to d_w. Conversely, if the dimension of the explorable space is lower than the dimension of the walk, the search is non-compact, and the particle will

oversample the local space, theoretically not leaving it until it is fully explored (Bénichou et al., 2007).

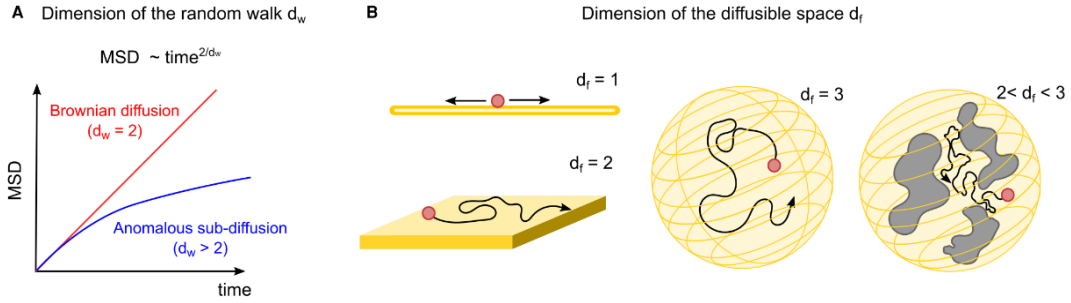

Fig. 1.6 **The characteristics of diffusion change depending on the dimensionality of the space explored compared to the dimension of the walk. A:** The dimension of the walk is determined using the mean-square displacement (MSD). **B:** Depending on the diffusible space, this translate to either compact or non-compact searches. From (Mazzocca, Fillot, et al., 2021).

The values of these relative dimension translates very directly in an MSD curve with an exponent lower than 1. However, the realities of SMT experiments in eukaryotic nuclei render the MSD unreliable, and thus other metrics tend to be more informative to provide evidence to this over-sampling of the local space. Commonly, research look at the distribution of angles between consecutive displacement: a departure from isotropic diffusion can already underlie anomalous diffusion, and the over-representation of near 180° angles is evidence of this frequent backtracking typical of compact searches (Hansen et al., 2020). Additionally, SMT allows us to quantify various kinetics parameters such as the diffusion coefficient, the fraction of molecules bound to DNA and their residence time on said DNA in any varying conditions.

This body of research has for example helped to explain how the pre-initiation complex and other general co-factors involved in starting the process of transcription assemble on their targets via confined exploration of particular sub-regions of the nucleus (V. Q. Nguyen et al., 2021), and how this confined exploration is regulated by Mediator. Another study showed that the sub-unit CBX2 of Polycomb uses condensates to attract and confine other CBX molecules (Kent et al., 2020). The kinetic parameters mentioned above also help our fundamental understanding of transcription factors, with the Elf lab recently providing evidence that the association rates are the main governing factor in sequence specificity, underlining that dissociation rates instead carry the risk of trapping the transcription factor on non-specific DNA for unsustainable amounts of time (Marklund et al., 2022). In eukaryotic environments, identifying specific endogenous binding sites by microscopy is difficult, given that the number of potential target sites is much too high to label them *in vivo*: in these settings, the fraction of immobilized molecules (i.e. the bound fraction) has be used as an indirect measure of the target search efficiency, since a higher bound fraction is related to transcription factors finding their targets easier (Y. Chen et al., 2022).

Interestingly, with regards to our previous discussion on multivalent interaction, these research have evidence a big role for the IDR of transcription factors.

In this chapter I have laid out the general concept underlying the research on target search mechanisms of transcription factors. We have seen the evidence that living cells have mechanisms to facilitate the finding of target sites, which rely on restricting the dimension of the space in which the transcription factor is allowed to diffuse. The control of this process can thus be a way to bias the search, and make it so the sites that are located inside this restricted space are found more frequently. In this context as well, we see that the amount of DNA in an eukaryotic nucleus, and the presence of more nucleosomes and obstacles, represents an additional hurdle for gene regulation.

Research on the relationship between chromatin and the target search in mammalian cells is only starting, however, and this present thesis aims at participating in this question with an innovative microscopy method.

Given that chromatin occupies up to 30% of the available volume of the nucleus (Ou et al., 2017), its role cannot be overlooked. Two roles have been proposed for chromatin: (1) chromatin could act by volume exclusion, reducing the available volume by preventing diffusion into regions of high chromatin density and (2) via unspecific, electrostatic interactions between it and transcription factor. In the next section, we will explore the various ways in which chromatin can influence the target search mechanism, both actively and passively.

1.5 Chromatin as an architect of gene expression guiding transcription factors.

Throughout the last two chapter we have highlighted the crucial role of compartmentalization in gene regulation in general, and for the target search in particular. The control of the space available for diffusion in the main deciding factor on whether or not a search is compact, non-compact, or an back and forth between the two, the so-called guided exploration. Coincidentally, the compartmentalization of chromatin has been a major area of research in the last decade, which has evidenced the controlled nature of chromatin architecture. In this part, we will take a look at those chromatin compartments starting from the top-down, and how they might affect and/or control the target search mechanism.

1.5.1 Heterochromatin & euchromatin.

The separation of the nucleus into eu- and heterochromatin is one of the earliest observation in cell biology, dating back to Flemming in the 1880s, and the idea that heterochromatin can hide genes or regulatory elements away from the transcription machinery is a long standing one as well. However, recent advances has given much more active role to chromatin than this simple one. The development of DNase-seq (A. P. Boyle et al., 2008), ATAC-seq (Buenrostro et al., 2013) and FAIRE-seq (Giresi et al., 2007) has given more quantitative data to the long standing idea that genes and regulatory elements can be hidden away in heterochromatin. They have given a powerful tool to identify elements by comparing accessibility across cell types (Thurman et al., 2012). There is a correlation between occupancy and accessibility at regulatory elements, such that regions poorly occupied by nucleosomes collect about 90% of the bound transcription factors (Thurman et al., 2012). However this doesn't explain why some perfect motifs are not regulatory elements. In addition, too many motifs remain accessible in any given cell-type to explain our target search problem (Kribelbauer et al., 2019)

Another line of argument against the simple idea that genes in heterochromatin regions are simply hidden away comes from FRAP and photo-activation experiments. FRAP experiments have for example shown that molecules of low mass, that is to say molecules that are small enough to navigate between the tightly packed heterochromatin, are effectively freely diffusing throughout the entire nucleus (Verschure et al., 2003). Above a size of 150 kDa, however, the steric hindrance of heterochromatin starts to have noticeable effects on the diffusive behavior of molecules. On top of affecting the diffusion coefficient, the fractal nature of the media in which the molecule diffuses starts to induce anomalous diffusion (Guigas et al., 2007). It is often assumed that RNA polymerase complexes are simply too large to diffuse into regions where chromatin is highly compacted, however photo-activation experiments have shown that the diffusion hindrance is moderate (Bancaud et al., 2009). In other words, it's not that proteins cannot enter per say, but that there is less diffusible space overall, given that chromatin fills up more of the space in heterochromatin. Interestingly though, molecules of similar molecular weight behave differently in this media, suggesting a phenomenon beyond simple volume exclusion. This topic will be discussed further in part 4.4, as part of the present work purports to quantifying the effect of this volume exclusion for p53.

1.5.2 Chromosome territories.

One step down from the global separation into eu- and heterochromatin are the chromosome territories. In interphase nuclei, chromosome occupy defined regions of the nucleus, generally one or a few micrometers in scale. The existence of this structured partition of

the different chromosome was revealed by the development of DNA fluorescence in situ hybridization (FISH) (Bolzer et al., 2005). The location and structure of these chromosome territory are not random, as evidence by the fact that silent heterochromatin tends to locate near the nuclear envelope (Misteli, 2020). These experiments also revealed the internal structure of these territories, as an highly branched and connected network of channels. Importantly, the existence of these territories goes hand in hand with the existence of interchromatin compartments, the channels in between two territories that are essentially devoid of DNA. These inter-chromatin channels have long been proposed as preferential channels for molecules such as RNA, or even nuclear speckles, to be transported around and out of the nucleus (Kim et al., 2019).

This structure relates to our problem in two ways, (1) the heavily branching pattern correspond to a fractal media in which to diffuse and (2) this structure places the DNA segments exposed on the surface of the channels in a preferential position to be found. In a recent paper, T. Cremer et al. (2020) go further in presenting a structural model of how chromatin architecture can influence the search mechanism of transcription factor. This model presents not a simple partition between eu- and heterochromatin, but considers the structural relationship between the inactive nuclear compartment (INC), the active nuclear compartment (ANC) and the interchromatin compartment (IC) (fig.1.7). In this framework, the INC is the highly compacted chromatin which acts through volume exclusion to hide inactive genes away. However, those domains of inactive chromatin have on their surface the active euchromatin. This view in interesting for our discussion. The interconnected channels provide pathways for the diffusion of specific molecules. More significantly, the placement of active euchromatin on the surface creates a two-dimensional search space. This contrasts with the one-dimensional search seen when molecules slide along DNA, as is typical in facilitated diffusion.

Recent 3D super-resolution and electron microscopy has shown that chromatin arranges itself in a chain of dense chromatin domains, that are separated from the surrounding, RNA filled interchromatin compartment (Miron et al., 2020).The same work correlated this different compartment with histone modification, showing that histone marks associated with transcriptionally active chromatin, and nascent RNA, seems to be located in the ANC. Importantly, a key feature necessary in this INC-ANC-IC model to be functional is the ability of cells to relocate genes to be silenced or activated inside or outside the INC respectively. There is evidence that long-term or permanently inactive genes, as well as regulatory elements, are localized inside the compacted chromatin domains representative of the INC (M. Cremer and T. Cremer, 2018). However, the orthogonal methods to study chromatin architecture have questioned the active role of chromatin architecture in gene regulation.

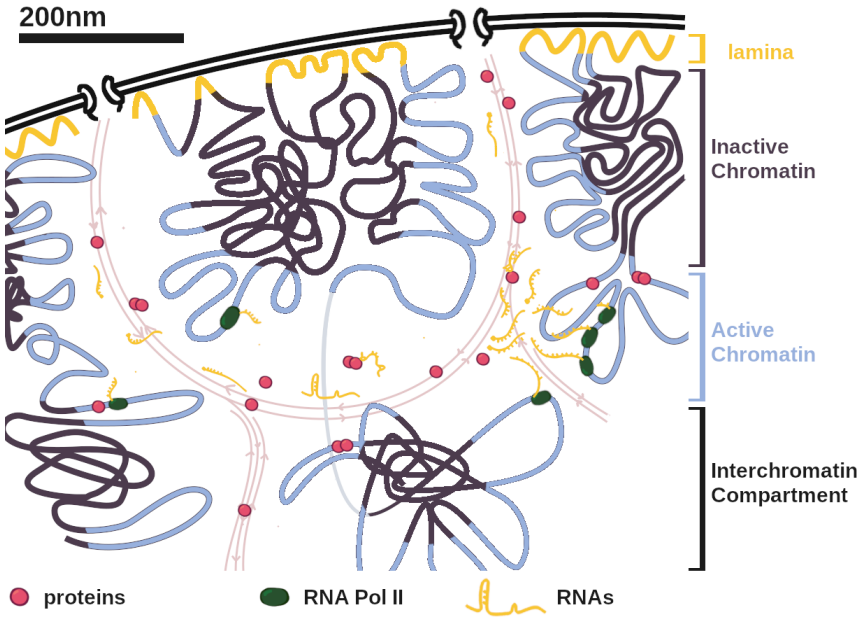

Fig. 1.7 **Global view of the nuclear lanscape separated in an inactive compartment, active compartment and interchromatin compartment.** This view refines the eu- vs. heterochromatin in two ways. It places euchromatin on the surface of dense chromatin domains ; and it highlights the role of a branching region largely devoid a DNA in which many molecules diffuse in and out of the nucleus. Inspired by (T. Cremer et al., 2020).

1.5.3 Topologically Associated Domains.

At the sub-megabase level, the structure that has attracted a lot of attention lately are the topologically associated domains (TADs) (Nora et al., 2012; Sexton et al., 2012). These structures were revealed with the advent of chromosome conformation capture (3C) and their derivatives like Hi-C or micro-C (Szabo et al., 2018). Those methods rely on cross-linking the nucleus, and using annealing followed by sequencing, measure how much different regions of chromatin are in physical contact with each others. This literature is vast, but one of the early take-away was that some regions form domains of preferential contact in a controlled and cell-specific manner. TADs are generally composed of one or several loops of chromatin held at the base by cohesin, with a fairly robust understanding of the loop extrusion process which creates those loops. Overall, TADs coincide to an extent with the mesoscale chromatin domains seen by electron scanning microscopy in (Miron et al., 2020), however a big conceptual difference between the two is that TADs are population averages. While it is possible to treat the contact maps given by Hi-C experiments as a proximity matrix and reconstruct a 3D structure from that (Lesne et al., 2014), however the resulting chromatin architecture is probably more of a theoretical construct than an actual conformation that was realized in any single cell.

Efforts have been made to link the features of TADs to gene expression, with considerable debate over which of the two precedes the other. On the one hand, deletion of a TAD boundary leads to aberrant expression (Lupiáñez et al., 2015) ; indeed the insulating effect

of TAD boundaries upon the enhancer-promoter contact, where the presence of a boundary prevents a regulatory element within one TAD to activate the expression of a gene in a neighboring TAD is probably the best understood function of TADs. On the other hand, the loss of cohesin or CTCF, which are the two factors responsible for the maintenance of TADs, only has a moderate effect on gene expression, despite the signature of TADs being clearly erased from the contact maps (Hyle et al., 2019). Looking at embryonic development in *D.melanogaster*, in which transcription only starts after 12 rounds of nuclear division, showed that while some TADs are present before transcription starts, the activation of transcription afterwards is correlated with the appearance of new TADs (Hug et al., 2017). Subsequent inhibition of transcription however doesn't erase those TADs completely, suggesting that other factors beside active transcription itself may be decisive in the formation of TADs. Overall, the degree with which transcription influences architecture and vice versa is largely unresolved.

The relationship between TADs and the target search mechanism, and more generally to the specificity conundrum, have been explored in some work, but remains under-determined. As we have mentioned, studies have shown how the clustering, in 3 dimensions, of loci containing low-affinity motifs may create hubs where occupancy is higher than expected (Kribelbauer et al., 2019). We mentioned phase condensates as a possible explanation for this detainment, however some have argued that the local chromatin architecture can contribute to this trapping by itself, because of a significant probability of reattachment – a probability compounded when several binding sites are in close proximity (Amitai, 2018). An interesting theoretical work showed that, assuming a facilitated diffusion model in which transcription factors would periodically slide in 1D, loci situated at the base contact point of chromatin loops ; on the other hand if the compaction of the TADs exceeds at certain amount, the transcription factors traffic tends instead to be restricted to the surface of the domain (Cortini and Filion, 2018). Lastly, another work showed, using polymer physics, that DNA double-strand break tend to passively diffuse at the border of chromatin domain, which they confirm experimentally in yeast (Amitai et al., 2017). Framed in the terms of the model proposed in Cremer, 2020, this is an example of a locus translocating from the INC to the ANC on its surface, although in this case it is not a regulatory element. It is noteworthy that most of these work are theoretical in nature, as simultaneous imaging of chromatin architecture and transcription factor movement is challenging.

1.5.4 Nucleosomes.

At the smallest scale of chromatin architecture is the nucleosome, with histones at their core. Nucleosomes are the first level of compaction required to organize the 2m of DNA in a coherent way, but they are most importantly a signaling hub, with many histone

readers and writers able to modify their tail to facilitate the recruitment or eviction of some other factors. The degree with which neighboring histones attract or repel each other is highly depended on electrostatic interactions. This dependence has fueled a long standing debate on whether or not nucleosomes associate in a 30nm fiber (Tremethick, 2007), with the consensus now shifting against the existence of such a structure, notably on the basis of electron microscopy of nucleus (Fussner et al., 2012). Instead, the higher-order interactions between nucleosomes seem to be more flexible structures called clutches, with some cell-type dependency (Ricci et al., 2015). While a single nucleosome spans around 160 to 240 bp – with 146 bp wrapped around the histone proper –clutches are estimated to comprise typically 4 to 8 nucleosomes, thus spanning a few kilobases (Ricci et al., 2015). Of note, the compaction status of these clutches is dependent on the modifications on their histone tails, which correlates with an increased size of domains that are epigenetically active, as opposed to repressed domains (Maeshima, Rogge, et al., 2016) (Nozaki et al., 2017).

Both single nucleosomes – and the histone therein – and clutches pose interesting problems to the target search. One I already mentioned is the simple fact that histone can act as a roadblock to sliding along the DNA, limiting the possibility of facilitated diffusion. Accordingly, in yeast, it has been shown that nucleosome removal speeds up the target search (Mehta et al., 2018). Several factors could mitigates this. For example, in or around regions of high chromatin density, the block represented by a nucleosome could be mitigated by having many other DNA strands nearby, on which the transcription factor could rapidly hop in a process called intersegmental transfer (Kamagata et al., 2020; Itoh et al., 2018). Some transcription factors may also bring along histone modifiers such as acetyltransferase, whose function is to acetylate the histone tail. The resulting effect is to loosen the wrapping of DNA around the histone, potentially increasing the accessibility of chromatin in that region (Fierz and Poirier, 2019; Michael and Thomä, 2021). It is not clear if transcription factors are able to diffuse inside nucleosome clutches. Some transcription factors, called pioneer, by definition have the ability to dislodge individual histones, however evidence suggests clusters may be solid enough to prevent diffusion, and thus access to the DNA inside the clutch, although the extent of this hiding depends on the transcription factor size (Axpe et al., 2019; Maeshima, Kaizu, et al., 2015). In the case of total exclusion, however, we come back to the possibility of the DNA exposed on the surface of the clutch, in a space of reduced dimensionality. This type of exposure may explain why some factors, such as p53, bind nucleosomal DNA with more affinity than naked DNA.

Through this exploration of the various nested structures inside the eukaryotic nucleus, we have seen the various compartment associated with chromatin at different scales. Although I was fairly descriptive, I believe a major theme emerges in the way these may

affect the target search mechanism, which is the importance of interfaces between different compartment. This appears at various scales, but is exemplified most clearly in Cremer's idea of an active chromatin compartment which is the surface of more compact, less active or downright silent chromatin domains. The fractal nature of these compartments, and the small scales of some of those, has so far limited our study of the interplay between chromatin and the target search mechanism, with a lot of work being theoretical.

With this first chapter, I attempted to frame our work in the greater context of the specificity conundrum, a theme which runs through a lot of the modern research on gene regulation in eukaryotes. The general answer to the lack of specificity of transcription factor in eukaryote is cooperation, but as we have seen there isn't yet a unifying framework to think about the exact ways in which this 'collaboration' is implemented. The natural answer, consisting of precise complexes stably bound together, is refuted in the general sense by many example of binding sites with loose arrangement in their binding motifs. The emerging, more complete answer relies heavily on compartmentalization of different nature, and research is underway to determine how exactly the various dynamic compartments affect transcription. Chromatin, as the main scaffold inside the nucleus, obviously plays a large role in this, in part through its effect on the target search mechanism. However, given the scale of these structures, experimental evidence has been lacking for this relationship between transcription factors and chromatin architecture.

The present work aims to target this gap with a novel microscopy approach. During my PhD, I have used a microscope able to combine single-molecule tracking with super-resolution imaging, allowing us to dissect the interplay between chromatin and target search with unprecedented resolution. We chose p53 as our case study for this research, for many reason which we will lay out in the next chapter, but one in particular is its pioneering activity. As I hope the last two chapter have underlined, there is considerable debate around the feedback loop of influence between chromatin architecture and the regulation of transcription. p53, as one of those transcription factor able to dislodge nucleosomes and change the chromatin landscape around its targets, sits right in the midst of these conflicting influences.

Chapter 2

P53, guardian of the genome.

The transcription factor p53, coded by the gene TP53, is important in the DNA damage response. Its crucial role in cancer, with 60% of cancers having some mutation on the TP53 gene (Baugh et al., 2018), has made it the most studied gene of all time, with a staggering average of 2 papers a day published on its biology (Dolgin, 2017). Yet despite the extensive research conducted, we still lack a clear and concise understanding of p53's mechanism of regulation and their functional implication. This is in large part due to the fact that p53 sits in the middle of a uniquely broad regulatory network – although some would take issue with the word uniquely, as it might just be a consequence of the literature bias (Weidemüller et al., 2021).

In this chapter, we will take a look at the biology of p53 and see that it is a good case study to dive more specifically into the concepts outlined in the first chapter. The large number of genes under its direct control and much higher number of potential binding sites provides an example of the specificity conundrum in action. Whether the p53 network of interaction is uniquely broad and flexible, or a simple consequence of literature bias, this situation affords us a wealth of data pertinent to our discussion. We will use this to outline the relationship that p53 has to chromatin in general, and see how its effect on the target search may help shed light on the control of the gene regulatory network of p53.

2.1 p53: Master of Cell Fate, Captain of Apoptosis and much more.

The textbook idea of p53 regulation goes as follows. The TP53 gene is expressed consistently, but p53 levels are kept low by a negative regulator, MDM2, which facilitates p53 degradation. Upon DNA damage, phosphorylation of p53 inhibits its interaction with MDM2 and increases its concentration in the nucleus. This leads to the activation of 3 pathways: cell-cycle arrest, DNA repair, and apoptosis. If DNA damage isn't repaired in

time, the cell dies. In this sense, p53 protects the organism by culling the cells that were too damaged to repair their DNA.

This function earns *TP53* the role of the prototypical tumor suppressor. As mentioned, the majority of cancers have TP53 mutations, suggesting that p53 is one of the major roadblock in the way of a cell getting cancerous, but its association to cancer is seen in other facts. *TP53* mutations are the basis of the hereditary cancer predisposition called the Li-Fraumeni syndrome (Malkin et al., 1990), and experiments with p53 knock-out mice have shown them to be more prone to cancer. The role of p53 as a guardian of the genome can also be seen clearly in cancer once it has developed. Pan-cancer analysis showed that *TP53* mutations are associated with gross changes in the number of chromosomes, resulting in two broad classes of tumors: when p53 is intact, the tumor presents a lot of mutations, when p53 is mutated, the tumor has many copy numbers alterations – i.e. many chromosome parts are either lost or duplicated (Ciriello et al., 2013). This has been attributed to two possible mechanisms. The first is the classical negative one, with culling of the cells after aberrant mitosis ; but in the second p53 takes a more active role and regulates spindle assembly and insures a proper G2/M transition (Vitre and Cleveland, 2012). This active role in genome maintenance is also seen in its binding to LINE and other transposons, participating in their continued repression (Chang et al., 2007). p53's crucial role in genome maintenance has been recognized for a long time, and p53 status is part of the diagnosis tools used to predict therapeutic outcomes (Kastenhuber and Lowe, 2017). But what we see from these fact is also that p53's function cannot be summed up to its pro-apoptotic functions, because it takes a much more active role in DNA maintenance.

What the other functions of p53 are is a surprisingly vast question because p53 sits at the cross-road of many pathways and controls a vast array of 'non-canonical' programs including senescence, inflammation, cellular metabolism, autophagy, pluripotency repression, cellular plasticity, ferroptosis (Kastenhuber and Lowe, 2017). Giving an exhaustive list might be futile. Thanks to Global Run-On Sequencing allowing for the detection of nascent transcripts upon p53 activation, and ChIP-seq studies, the most recent estimates places the number of direct p53 target genes at 350 (Allen et al., 2014; Fischer, 2017), and ultimately up to 1,500 genes whose level of expression are altered by its presence (Hafner, Bulyk, et al., 2019). In this regard, p53 presents a good case for Boyle's omnigenic model (E. A. Boyle et al., 2017): it proposes that all genes contributes to all traits to some degree, and the fact that one single transcription factor alters the levels of a significant fraction of all genes certainly helps in that regard.

Importantly however, not all of the 350 high-confidence, direct target genes are similarly activated in response to p53 activation. Instead, depending on the context, such as the cell-type or the stimulus, p53 engages a qualitatively different set of genes, with only a

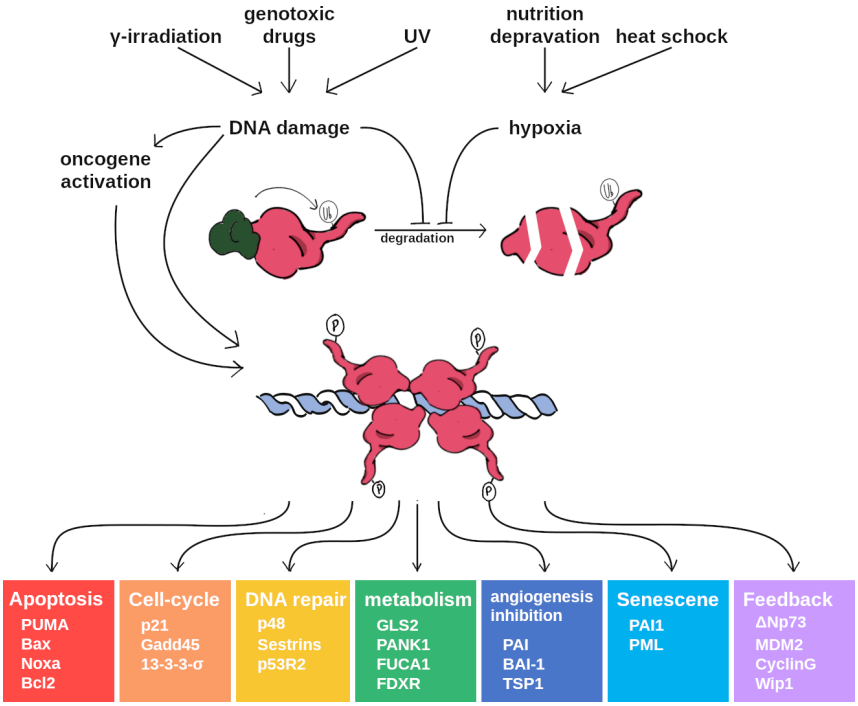

Fig. 2.1 **p53 controls a spectrum of pathways linked to cell-fate after stress.** Under normal condition, MDM2 (in green) binds p53 (in red), which hinders its transcriptional activation, and ubiquitinates it, targeting p53 for degradation. After many kind of stressors, the phosphorylation of p53 prevents this interaction with MDM2, allowing p53 to bind its target genes as a tetramer. Below are a few of the target genes and downstream pathways, including the canonical (apoptosis, cell-cycle arrest, DNA repair) and non-canonical.

minor fraction of about 60 genes which is consistently up-regulated (Kastenhuber and Lowe, 2017). Some of the programs under p53's control have contradictory effects, such as senescence which involves anti-apoptotic factors, but despite the forty years of research, we still lack a clear picture of how p53 transcriptional activation makes a choice towards one cell fate or another. This also makes p53 an interesting illustration of the 'specificity conundrum', as this choice involves preferring some binding over others.

2.2 The DDR's Master Key: A look at p53's structure and binding specificity.

p53 is 393 amino-acid long, with several functional domains (fig.2.2). On the N-terminus are two transactivation domains, both of which are required for full p53 activity, although mutations in either of them leads to different outcomes. After a proline-rich domain comes the DNA-binding domain, and a tetramerisation domain. Interestingly, absence of this tetramerisation domain doesn't prevent p53 from tetramerizing when binding DNA, but it does stabilize the complexes once bound, and is required for tetramerisation in the absence of a DNA template (Veprintsev et al., 2006). Finally, there is the C-terminal domain

which, following the trend in eukaryotic transcription factors, is intrinsically disordered and more than 30 amino-acid long, a feature rarely seen in prokaryotes. In fact, despite the transactivation domains located within it, the N-terminal region also resists crystallization, meaning it is highly disordered (B. Xue et al., 2013).

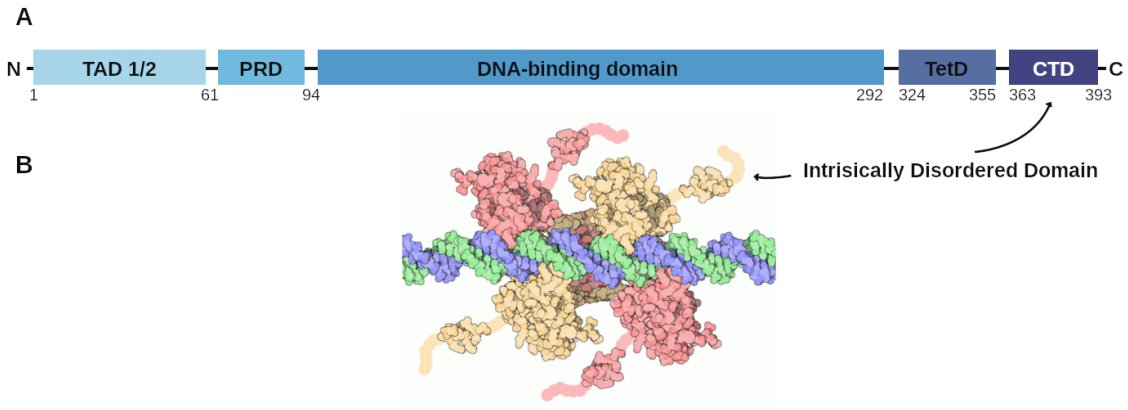

Fig. 2.2 **The structure of p53.** A: Presentation of the different domains of p53. TAD: Transactivation Domain, PRD: Proline rich region, TetD: Tetramerization Domain, CTD: C-terminal domain. B: Structure of p53 binding to DNA as a tetramer, reconstruction from PDB entries (CC-BY David Goodsell). The disordered regions (CTD), which are too flexible for crystallography, are shown in the render as plain colors.

Through its transactivation domains, p53 interact with several components of the pre-initiation complex, including Mediator, TFIIB, TFIID, TFIIH, and promotes their assembly on the gene promoter and transcription initiation. Furthermore, recent evidence has shown that p53 interacts directly with RNA polymerase II, facilitating its recruitment and increasing the efficiency of the elongation phase (Liou et al., 2021). This direct promotion of transcription is in line with evidence showing – after a lengthy debate – that p53 is solely an activator of transcription. While some genes show decreased activity upon p53 activation, mainly cell-cycle genes, this repression is indirect and mediated in large part by the p21-DREAM/RB pathway (Morgan A Sammons et al., 2020). In studies with mice, researchers found that both transactivation domains were needed for p53 to do its job correctly, although each one had slightly different effects on gene activity. For example, when the first one was changed, it had a bigger impact on gene activity than changes to TAD2. Yet, TAD1 could still activate some genes, like Bax, even when both transactivation domains were inactive, suggesting that these two transactivation domains might help p53 choose which genes to target (Brady et al., 2011). Similarly, experiments have shown that p53 lacking the proline-rich domain have a reduced ability to restrict cell growth (Walker and Levine, 1996). The deletion of the C-terminal domain similarly leads to changes in the transcriptional activation of p53, with reduced expression of both pro-apoptosis genes and pro-survival genes (Hamard et al., 2012). Most importantly, the C-terminal domain is also the site of multiple post-translational modifications (PTM) which affect p53

function (Hafner, Bulyk, et al., 2019). Both termini are known to interact with a multitude of proteins, which place p53 at the center of an remarkably large interactome. From an evolutionary perspective, it is noteworthy that both the N and C termini are under much more evolutionary pressure and much less conserved than the rest of the sequence (Lane et al., 2011).

Of course, as a transcription factor, the domain most crucial to the effect of p53 on gene expression is the DNA-binding domain, and the motif it recognizes on the DNA. Fortunately, in this domain, the literature bias affords us a wealth of data for no other transcription factors, including at least 132 different ChIP-seq studies. This allows us to put accurate number of the challenges faced by p53 in finding the correct targets in the genome.

As mentioned previously, p53 binds as a tetramer. As such, its canonical is composed to two half-sites, two decamers of RRRCWWGYYY (where R= A/G, W=A/T, Y=C/T), with sometimes a few nucleotides between the two decamers as spacers (T.-A. T. Nguyen et al., 2018). With two decamers, the canonical motif of p53 is 20bp long, which is unusually long for an eukaryotic transcription factor. If we look at it from the information theory perspective, this amounts to 20.6 bits of information, which is significantly higher than the average transcription factor motif – 12.1 bits – but falls significantly short of the 30 bits required to specify unique address (Wunderlich and Mirny, 2009). High information content is typical of TFs involved in cell-autonomous transcription programs, but as we can see in this case it is still too little to operate under the prokaryote paradigm. The consequence of this lack of information is reflected in the high number of perfect half-sites and full sites expected probabilistically – 732,400 and 179 respectively – which matches roughly the number actually seen in the human genome – 732,000 and 550. For full sites, the three-fold increase compared to expectation suggests positive selection. Still, in silico motif searches have turned out nearly 100,000 canonical p53 responses elements, less than 8% of which have been found bound in any dataset (Martin Fischer, 2019). p53 binding is also observed in non-canonical sites, such as sites with a single mismatch, or one or two bases between the decamers, although such binding sites are less consistently and less strongly bound (T.-A. T. Nguyen et al., 2018). Interestingly, a lot of genes that have been individually reported as repressed by p53 have non-canonical sites with spacers, or no sites at all – despite p53 binding – which is also the case for the Alu elements which p53 helps repress, although no mechanism has been proposed to explain this peculiar correlation (Hafner, Bulyk, et al., 2019).

T.-A. T. Nguyen et al. (2018) performed an enviable meta-analysis on a total of 41 ChIP-seq datasets in activated condition and 17 datasets in basal condition, in various cell types and treatment, and provides essential insights about the specificity of p53 binding. Their analysis shows a total of 54 947 individual peaks of p53 binding, representing only

1.6% of the total amount of p53-like sequence in the genome. Of these, >1000 loci are bound in at least 20 of the activated datasets, and might be considered part of a core, cell-autonomous cistrome of p53. Interestingly, the remaining variance in binding site strongly correlates with cell-type, suggesting that the least frequently bound loci hint at the more specific aspects of p53 regulation. For example, cell-cycle related genes typically contain high-affinity motifs in their nearby p53 response element, while the p53 response elements near apoptosis related genes featured a broader range of affinities (Hafner, Bulyk, et al., 2019).

These observations – and some before (Inga, Storici, et al., 2002) – first led researchers to propose an affinity model, in which the decision over cell fate would be decided by the different affinities of the binding sites. With lower affinity binding sites, apoptosis genes would require higher concentrations of p53 before being effectively activated, whereas cell-cycle genes were triggered at lower doses, leading to apoptosis only in the case of important DNA damage. However, this idea has been since falsified. Some key regulators of apoptosis, *PUMA* and *NOXA1*, have high-affinity motifs (Kaeser and Iggo, 2002). Furthermore, there is no consistent correlation between p53 levels and binding at promoters of apoptosis genes (Smeenk et al., 2011), and some motifs near apoptosis genes are bound even at low p53 concentrations (Kracikova et al., 2013). Moreover, only 11% of peaks were associated with differential gene expression in at least one data set, and some of these include peaks with non-canonical motifs. Conversely, 25% of the perfect motifs are bound in none of the activated datasets.

This data aptly illustrate that transcription factor specificity in eukaryotes is not simply explained by the presence of DNA motifs, nor the affinity of said motifs, which shouldn't come at a surprise given our discussion so far. Clearly, p53 binding and subsequent activity relies on several additional mechanisms to discriminate between binding sites, which we will explore in the rest of this chapter.

2.3 Regulation and cooperation: who pulls the strings of p53?

The solution most often discussed in response to the apparent lack of specificity of transcription is cooperation with other factors. These protein-protein interactions generally come in two variety: signaling, in the form of post-translational modifications (PTMs), or the formation of complexes. We will explore both of those.

Thanks to mass spectrometry, more than 300 PTMs have been described for p53, and they are considered to crucial part in understanding the p53 gene regulatory network

(Hafner, Bulyk, et al., 2019). The most crucial of these PTMs are involved in p53's interaction with its primary negative regulator MDM2. In normal circumstance, MDM2 deposits several ubiquitins on p53 which targets it for proteasome degradation – additionally, MDM2 binds p53 at the N-terminal domain and hides it from other potential partners (Kussie et al., 1996). Conversely, upon DNA damage, the phosphorylation of serine and threonine residues prevent the MDM2-p53 interaction, stabilizing p53 thus leading to increasing p53 levels. These PTMs form the basis of p53 regulation and lead to a crucial quirk of p53 biology: MDM2 is a direct target gene of p53, meaning p53 activation leads to up-regulation of its negative regulator, which creates an oscillator.

A number of kinases like ATR & ATM or HIPK2 are known to phosphorylate serines at the N-terminus of p53, within the transactivation domains. Distinct phosphorylation events occur in response to different cues ; gamma-irradiation leads to a rapid increase in Ser6 and Ser15 phosphorylation, while UV exposure leads to a delayed increase in Ser392 phosphorylation. P53 response after gamma-irradiation and UV exposure is in turn different, although if and how the different phosphorylation mediate this difference in response in not known. ATM/ATR, in contrast, mediate a phosphorylation at Ser18 whose effects are better understood, as it has been shown to be required for cell-cycle arrest.

The C-terminus, in turn, is strongly regulated, but more so by methyltransferases and acetyltransferases, which respectively methylate and acetylate the lysine residues. Lysine is a charged residue, while their acetylation neutralize this charge. In turn, this leads to a change in the electrostatic interactions of the CTD – which is an Intrinsically Disordered Region (IDR) – with chromatin: accordingly it has been shown that the acetylated inhibits the CTD's ability to bind DNA (Friedler et al., 2005). Acetylation of p53 has thus often being studied in bulk, looking at the overall effect of increases in acetylation levels of p53, with mixed results. Preventing the loss of the charge from acetylation via lysine-to-arginine substitutions had only effect on gene expression only in some cell-type and, while acetylated p53 is enriched at the promoter of *CDKN1A* – the main cell-cycle arrest target gene of p53 – but it had equal or even greater effects on apoptosis than on cell-cycle arrest. These reports were reconciled by the finding that the CTD enhances p53 binding at non-canonical sites, and that this function is inhibited by acetylation (Laptenko et al., 2015), which may give acetylation its ability to fine-tune the expression of genes down-stream of p53. Methylation of the same lysine residues can probably be understood within the same framework, as mono- and dimethylation do not change the charge of the residue, but do shield it from acetylation. Lys382 mono- and dimethylation leads to less binding at both the promoter of *PUMA* and *CDKN1A* (West et al., 2010) – the main effectors of apoptosis and cell-cycle arrest respectively, both possessing high-affinity motifs in their enhancers – while abrogating the modification as Lys382 leads to premature apoptosis (Shi et al., 2007).

In this context, methylation may be useful to keep p53 away from some of the high-affinity sites that would trigger apoptosis, an irreversible program.

Finally, these PTM show significant cross-talk, with multiple interdependencies observed. For example, certain phosphorylation events promoted binding of key acetyltransferases such as p300 and/or CBP, and subsequent acetylation of the CTD (Sakaguchi et al., 1998). Similarly, there is an interplay between acetylation and ubiquitylation which is crucial for regulating all p53 functions by controlling its stability (M. Li et al., 2002).

In summary, we see that some PTM are specific to some type of DNA damage and play a role not only in regulating p53 stability against the ever-present inhibitor MDM2, but also in fine-tuning the target gene selection. The cross-talk between all of these PTM and how the influence cellular response is an active field of research, and as we saw they involve many important partner proteins such as p300. Cooperation with co-factors is another oft proposed way of increasing the specificity of transcription factor response but, as we will see p53 differs from most in this regard.

Indeed, p53 seems to work in isolation – although it does bind as a tetramer. This is seen in the fact that p53 binding sites are not associated with the binding of other transcription factors (Verfaillie et al., 2016) and it is able to function as a transcriptional activator in yeast, where it able to activate transcription even though its potential cooperative partners are presumably absent (Inga, Monti, et al., 2001). A small caveat to this has been shown during embryonic stem cell differentiation, during which p53 binding sites were found alongside the binding sites of pluripotency factors such as OCT4 and SOX2, potentially hinting at the fact that p53 might function differently in development and during the DDR (Akdemir et al., 2014). Another factor, SP1, seems to be involved in the outcome of, p53 activation: when SP1 is overexpressed, p53 activation leads to apoptosis, whereas SP1 depletion protects cells from p53-mediated apoptosis (H. Li et al., 2014). Interestingly, 115 of the target genes that this study identifies as common targets for p53 and SP1 contain response elements for both within 500bp of each other, suggesting some co-binding. This was not associated with the canonical genes of apoptosis such as *PUMA*, *BAX* or *NOXA1*, but rather genes associated with the repression of alternative pathways. This seemingly contradicts the findings of (Verfaillie et al., 2016), however a subsequent study by Catizone et al. (2020) using the same massively parallel reporter assay confirmed this instance of co-regulation, as well as others like ATF3, a difference in results that might result from differences in the exact protocol. They overall come to the same conclusion however, that p53 is both necessary and sufficient for upregulation of target genes near p53 response elements, although some co-factors may play a role in regulating the strength of the response. Both (H. Li et al., 2014) and (Catizone et al., 2020) find that the relative position of the motif of p53 and the co-regulator are highly variable, arguing against the

formation of a complex, and rather a 'billboard' model similar to what has been described in Drosophila enhancers (Spitz and Furlong, 2012). This flexible model of gene regulation proposes that different combinations of factors can bind to a response element and produce the same transcriptional outputs, effectively buffering the gene against the loss of one of the factors. In this model, the binding of the different factor might be independent, but in the case of p53 an interesting fact might suggest some degree of cooperative binding, although indirect : the binding of a transcription factor often leads to the production of short enhancer RNA (eRNA) (Morgan A. Sammons et al., 2015) which is correlated with enhanced p53 binding (Allen et al., 2014).

Lastly, like many other eukaryotic transcription factors, p53 has paralogs in the genome, p63 and p73. Like other family of transcription factors, those three paralogs bind very similar motifs, and as such p63 and p73 bind the majority of p53 motifs in the genome, especially the high-affinity ones (Brandt et al., 2009). From an evolutionary perspective, *TP63* and *TP73* are the ancestral form of the gene, but they are more lineage specific than *TP53*, they are largely restricted to basal and ciliated epithelial cells respectively (Morgan A Sammons et al., 2020). Of these two paralogs, p63 is the most studied. It was initially described as competing with p53, for example in the case of the *CDKN1A* locus, which (Westfall et al., 2003) showed to be bound by both but less activated in the presence of p63 compared to when only p53 was present. This competitive binding made p63 a dominant negative repressor of p53. However, further studies have shown that p63 functions primarily as an activator, with only a small overlap between the genes it regulates compared to those of the 350 direct target genes of p53 ; of the 180 high-confidence target genes of p63, only 19 are up-regulated by both factors (Woodstock et al., 2021), casting doubt on the importance of this sibling's rivalry. At the endogenous *CDKN1A* locus, ChIP-seq studies find p63 to be bound alongside p53, but with contradictory effect on the gene expression (Fischer, 2017). More study are needed to dissect the cross-talk between these three siblings. One question particularly apropos given the subject of this work is how is it possible for the p53 and p63 to have so little overlap in target genes when they have such similar DNA-binding domains.

Thus we see a little tension in the literature about whether or not p53 works as a 'single-factor' model. All agree that p53 on its own is able to induce transcription by binding at p53 response elements, there are evidences that some co-factors may help fine-tune the response in a context-dependent manner. These co-factors come in addition to the complex network of acetyltransferases and other protein modifiers, and helps provide some mechanism for how the specificity gap is bridged. Those protein modifiers are also extremely important for a crucial part of the p53 response: its pioneer activity and ability to not only react to, but also change the chromatin landscape.

2.4 No hiding place for response elements: p53 pioneering activity in chromatin.

For a number of transcription factors, the response to the question 'How to choose between many binding sites?' lies in a large part in chromatin accessibility. The classical compartmentalization of the nucleus into euchromatin and heterochromatin, or more holistically assembled by (T. Cremer et al., 2020) into the inter-, active and inactive chromatin compartments, segments the nucleus into a regulatory switchboard, selectively exposing or obscuring binding sites. When it comes to p53, the story gets significantly more complicated. p53 is a member of the restricted club of pioneer factors ; factors that have the ability to bind closed chromatin and open it up to activate the transcription of the genes within. As such, chromatin structure affects p53 differently than other TFs, and in this section we will explore this two-way relationship.

The presence of nucleosomes inhibits the binding ability of most TFs, but the reverse is true for p53, which was found early on to bind chromatin with higher affinity than naked DNA (Espinosa and Emerson, 2001). In vitro, it has further been shown that p53 can bind response elements directly wrapped on this histone (Yu and Buck, 2019). The consequences in vivo have been measured at the genome-scale by Su et al. (2015), an interesting study combining ChIP-seq and DNase1-seq to look at chromatin accessibility, which found 64% of p53 binding peaks to be within repressed regions. Interestingly, they find an inverse correlation between the affinity of the motif and chromatin accessibility: the high-affinity motifs tend to be inside regions of repressed chromatin – and specifically associated with transposable elements – while response elements with lower-affinity motifs are located in more accessible regions. This ability of p53 overall means that it is less affected by the chromatin landscape than many TFs. The main predictor of p53 binding remains the presence of a p53 motif: regardless of the surrounding chromatin and histones, p53 finds its response elements. This translates into the observation that p53 binding sites are largely conserved across cell-type and treatment (Verfaillie et al., 2016). Despite this, an analysis by (Morgan A Sammons et al., 2020) showed that the variation in binding which does exist correlates with cell-type. Genome-wide studies have also found no influence of DNA methylation. Additionally, it has been found that DNA hypomethylation and CpG islands are associated with more p53 binding, but only in normal, non-cancerous cells (Botcheva et al., 2011), which could be one way to fine-tune which response elements are more bound than others. Additionally, this highlights the fact that conclusions made in cancer cells may be misleading for p53 function in normal cells.

Crucially, pioneer factors are typically not only able to find targets inside nucleosome-rich heterochromatin, but also fish them out and transform the region into active chromatin.

Accordingly, it has been shown that p53 binding can lead to the eviction of the nucleosome, possibly paving the way for the recruitment of the rest of the transcription machinery (Morgan A. Sammons et al., 2015). But the eviction of nucleosome is only the first step in making a region into active chromatin ; the next involve the recruitment of histone modifiers to depose the appropriate epigenetic marks in the surrounding region. This activity has also been shown for p53, with several of the co-factors mentioned in the previous section. For example, p300, responsible for the acetylation of several residues in p53's own CTD, is also to acetylate histones H4 (Barlev et al., 2001), while SETD1A, recruited by both p300 and p53, methylates Lys4 of the histone H3, another gene-activation histone mark (Tang et al., 2013). Lastly, p53 is able to recruit the energy-dependent SWI/SNF chromatin remodeling complex, which promotes chromatin accessibility (Oh et al., 2008).

Despite this pioneering activity, a genome-wide study in fibroblast using ATAC-seq before and after p53 activation found that the majority of p53 binding sites do not lead to increases in accessibility of the chromatin (Younger and Rinn, 2017). Either the sites occurred in accessible chromatin which remained accessible (565 binding sites out of 2,638) or in inaccessible chromatin which remained so (1,630/2,638). Only a total of 275 binding sites were associated with an increase in chromatin, with some known p53 target genes found at those loci. This confirms p53's role as a pioneer transcription factor, but also shows that this activity is limited, with no clear evidence as to why some loci see their chromatin opened up but not others.

In summary we see that hiding p53 response elements inside inaccessible chromatin regions doesn't bias the response of p53 in the same way it would for a classical transcription factor. Given that motifs in heterochromatin are also generally those with the highest affinity, it also means that those unresponsive sites still sequester p53 for significant amounts of time, which means inaccessibility is not directly a way to speed up p53's search for the right response elements in the chromatin. Motifs inside heterochromatin are still found and bound, yet only some of this bound sites lead to activation of transcription, with unclear mechanism to explain what guides p53 and its target selection. This pushes us to explore different ways in which chromatin and other molecules may act as a scaffold for p53's functionality.

2.5 Sliding into your DNA Motifs: the guides which help p53's search for its targets.

Some of p53's perplexing behavior when it comes to target selection and the effect which follows p53 binding may be planted in the weak multivalent interactions and associated

subnuclear compartments which we began to explore in Chapter 1. In particular we have seen two – not incompatible – ways in which condensates and sliding on non-specific DNA can reduce the dimension of the search, helping the protein to find its targets faster. In this section we will go over the evidence for both when it comes to p53, and how these may bias the target selection.

The concept of phase separation and condensates has begun to reshape our understanding of cellular compartmentalization and function. As I laid out in 1.3, they can act as hubs, raising to local concentration of some molecules, but they have also been proposed as a way to speed up the target search of transcription factors (Hansen et al., 2020). According to this idea, since molecules are less likely to leave the condensate once entered, they have a higher chance of backtracking and exploring the space within the condensate more fully, meaning condensate help make the exploration compact. A recent study provided evidence that p53 may concentrate in condensates of 53BP1 molecules following DNA-damage (Kilic et al., 2019). The studies shows that perturbing those condensates changes the expression of p53 target genes, however it must be said that in this case, the condensates form at the site of DNA damage, and not around p53 response elements, meaning the mechanism by which this impacts gene expression is different. Moreover, it must be said that fluorescence microscopy – both in live and fixed cells – doesn't show any visible condensates of p53 in vivo, even after DNA damage. As IDR are thought to be be one of the main vectors for those weak multivalent interactions necessary for phase separation, we tried in a recent work from our lab to swap p53's IDR with that of FUS, a protein known to form condensates, which resulted in the formation of p53 condensates, which interestingly this lead to a decrease in the transcription of p53's main target *CDKN1A* (Mazzocca, Loffreda, et al., 2022). Thus it seems unlikely that p53 forms condensates with its native IDR, although it may interact to some degree with the DDR-related condensates formed by 53BP1.

Besides IDR, RNA molecules are also another type of molecules able to mediate weak multivalent interactions. Here it is interesting to note that p53 has been shown to bind with a long non-coding RNA called DINO, which stabilizes the protein in return and leads to increased binding at the *CDKN1A* promoter (Marney et al., 2021). Another point of evidence is the presence of eRNA at some p53 response elements, produced in a p53-independent manner but whose production increases upon p53 binding, which has been found to lead to enhanced p53 binding and transcriptional activation of multiple p53 target genes (Allen et al., 2014). Using 4C, Melo et al. (2013) found these eRNA-rich response elements to interact with multiple distant genes, although these interactions were found to preexist p53 activation. We might try to understand these findings in the light of the mechanism proposed by Henninger et al. (2021): the local concentration of negatively charged eRNA allows tends to attract the positively charged p53 IDR and retain it longer

at the site. This funnel-like behavior, similar to what the phase condensates have been proposed to do, may be crucial given that the very short binding time of p53 measured in live-cell imaging, a fact too often not taken into consideration when discussing chromatin looping to distal sites. Data from our lab suggests that p53 binds for an average of 3.1s in untreated conditions, and 6s in activated conditions (Loffreda et al., 2017). In this context, it may be crucial to ensure p53 is retained in the vicinity of the target to bind several times, or is otherwise able to find the target again quickly.

It is for this very purpose that p53's IDR, despite its seeming inability to form condensate, may have an important role in shaping the response of p53.

The impact of the IDR of p53 is first seen with studies using p53 mutants lacking the last 30 amino-acids of its IDR. This mutant version, $p53\Delta30$, has been shown in ChIP-seq studies to have significantly different binding patterns, with effects on transcriptional activation: $p53\Delta30$ fails to bind two thirds of its known response elements, and has a decreased association with the remaining third (Laptenko et al., 2015). In this study, most of the elements not recognized by $p53\Delta30$ were non-canonical motifs, suggesting a role in particular for the binding of these low-affinity sites. Importantly, $p53\Delta30$ doesn't bind any response elements that isn't bound by wild-type p53, indicating that the IDR has a positive effect on binding. On the other hand, the lack of IDR changes both the expression of pro-apoptosis and pro-survival genes, making it difficult to sum up its effects in a simple model (Hamard et al., 2012). Nevertheless, it is noteworthy that removing p53's ability to interact non-specifically with DNA changes the occupancy patterns of the specific sequences inside its response elements. The importance of p53's IDR shows also in its role in allowing it to bind nucleosomal DNA, making it an essential domain for its pioneering activity (Morgan A. Sammons et al., 2015). Seemingly confirming this is data from our own lab using single-molecule tracking, which suggests p53 lacking its IDR is unable to penetrate into chromatin dense regions (Mazzocca, Loffreda, et al., 2022).

In vitro, p53 has been shown to be able to slide on DNA with the help of its IDR (Tafvizi et al., 2011). This sliding, allowing for temporary 1D diffusion, was the other mechanism besides phase condensates proposed to allow for facilitated diffusion, this type of diffusion with periodic reduction in the dimensionality of the search which speeds up the target search which I described in 1.4. This model of 3D+1D diffusion encounters some problems in eukaryotic cells where nucleosomes are present and could act as obstacles. However a recent work has shown in vitro the ability of p53's IDR to promote intersegmental transfer (Graha Subekti and Kamagata, 2021): short jumps from one segment of DNA to another nearby one could be a way to bypass nucleosomes when p53 encounters one. In promoting this type of facilitated diffusion, the IDR of p53 may help it reach some motifs that it

would otherwise miss.

In this section we have seen the evidence for the main ways by which the target search can be sped up in eukaryotes, something which can help bring back some of the missing specificity in the mechanisms by which transcription factors enable transcription. On top of this specificity, the target search mechanism also explains the observation that TFs reach their targets faster than the diffusion limit. For this purpose, condensates have been described in a funnel-like model, both making the target bigger by having a blob of molecule around it, and by sequestering the TFs inside to increase the chances of it finding the target. But while some tentative evidence may go in this direction, it is clear that p53 condensates if they exist are of a different nature than those classically described. Baring this explanation, facilitated diffusion in the form of sliding on DNA remains as the main model. Strong evidence exists in vitro for this facilitated diffusion, but the functional implications of this IDR mediated non-specific binding and sliding is lacking, and requires further study.

It is in this gap of knowledge that the present thesis stands. Previous work in our lab have looked at the effect of changing p53's IDR on the diffusive behavior of p53, its ability to explore chromatin and the resulting changes in gene expression. This work was started on the inverse premise, how does transforming the chromatin landscape affect p53's ability to find and regulates its targets ? By doing so, we aim to further our understanding of the interplay between p53 and its nuclear environment, recognizing that both the dancer and the stage are integral components in this cellular ballet.

Our interest was specifically piqued by a process in which p53 is deeply involved: senescence. This peculiar state of stable cell cycle arrest that has profound implications for aging and cancer. In particular, the phenomenon of oncogene-induced senescence (OIS) provides a unique model for studying such transformations, since it is marked by a dramatic reorganization of the chromatin landscape, including the formation of senescence-associated heterochromatin foci (SAHF). We propose that understanding the impact of these high-density foci of heterochromatin on p53 target search and binding might offer valuable insights into the broader mechanisms regulating p53 behavior.

Chapter 3

Living on the Edge: The Dual Role of Senescence.

Injuries to the DNA trigger the DNA-damage response, a regulatory cascade in which p53 and many others cooperate to detect and repair DNA lesions. If the damage is reversed the cell can reenter the cell-cycle, however the damage can be irreversible, which can lead the cell to activate death mechanisms in order to preserve the integrity of the tissue in which it is embedded. Such cells can also preserve tissue integrity by entering another cell state, complementary to programmed cell death : cellular senescence. Senescence is a cell state characterized by two main features, irreversible cell-cycle arrest and the senescence-associated secretory phenotype, the latter of which signals to the immune system which will, in due course, remove the senescent cell.

Cellular senescence has long been viewed as a barrier to cancer, in particular when it is triggered in reaction to the activation of an oncogene (oncogene-induced senesence, or OIS). However, a number of studies have shown senescent cells to promote tumor progression and invasion. This makes senescence an interesting example of antagonistic pleiotropy, a double-edge sword with both beneficial and detrimental effects (Schosserer et al., 2017), the study of which is essential to understand the implication of the proposed therapies targeting senescent cells.

Interestingly for our study, senescence also manifests in a large scale reorganization of chromatin. In particular, cells in OIS show senescence-associated heterochromatin foci (SAHF), regions of dense, repressive chromatin which are prominent in images with DNA staining. Research has progressed on their formation, but mechanistic insights into their function is lacking. SAHF are generally thought to be a protective structure, preventing senescence bypass, however they are not a universal feature of senescence, casting doubt on their function (Kosar et al., 2011). These elements make SAHF an interesting target of inquiry for the relationship between transcription factors and chromatin, particularly for p53, a pioneer factor shown to be essential for SAHF formation. In this chapter, we will

look at the senescence phenotype and its consequences for human health, before focusing on the changes in chromatin architecture.

3.1 The paradox of senescence: protector or perpetrator?

Cellular senescence was originally described by Hayflick and Moorhead in 1961, who noticed fibroblasts in culture had a finite number of cell divisions they could go through. This so-called Hayflick limit was immediately linked to aging, from the obvious implication that a cap on the number of cellular division would have for the maintenance of tissues. However, the heterogeneity of the phenotype, a running theme of the literature, made it hard to detect senescent cells, and thus difficult to confirm the phenotype even existed in vivo, let alone to identify a supposed causal link between senescence and aging.

Today, the International Cell Senescence Association (V. Gorgoulis et al., 2019) defines senescence as a state triggered by stressful insults characterized by (1) prolonged and generally irreversible cell-cycle arrest, (2) macromolecular damage, (3) altered metabolism and (4) with secretory features (fig.3.1).

The inclusion of the qualifier in 'generally irreversible' hints at a shift in our understanding of senescence, which until recently was thought of as a static endpoint. In this view, cellular senescence was understood chiefly as a barrier against cancer (Campisi, 1997; M. Serrano et al., 1997). But this idea encounters an immediate objection: at first glance, it is hard to understand what advantages would senescence have other direct cell death. Even though the phenotype is stable, and cells are rarely able to bypass it, it would seem like senescence only introduces more opportunities for errors, and thus be a strictly inferior tumor-suppressive mechanism compared to apoptosis (Childs et al., 2014). This makes senescence an interesting example of pleiotropism, with senescent cells involved in different antagonistic function, and this has important consequences for human health. Indeed, senescence has important antitumorigenic effects, but also protumorigenic effects, and nowhere is it more aptly illustrated than in the senescence-associated secretory phenotype (SASP).

A defining characteristic of senescence, the SASP consists of various molecules, including inflammatory cytokines, growth modulators, angiogenic factors and matrix remodeling factors (Juan C. Acosta et al., 2008; Coppé et al., 2008) although the exact makeup varies according to the cell type and senescence trigger, with considerable cell-to-cell variation even within the same model of senescence (Wiley et al., 2017). The molecular mechanism of its production relies on DNA Damage Response (DDR) pathway components, such as ATM, CHK2 and NBS1 (Faget et al., 2019), and should be viewed as an actively and continuously regenerated pathway associated with the persistent DNA

damage. Other pathways, including the cGAS-STING – which detects the accumulating cytosolic DNA (Dou et al., 2017) – and JAK-STAT pathways, influence the SASP through the key transcription factor NF-κB. The mTOR regulator modulates the SASP post-transcriptionally, impacting SASP without altering senescence growth arrest (Laberge et al., 2015). A notable component is NOTCH1, which controls pro-inflammatory cytokines in the SASP and mediates juxtacrine communication. Crucially, NOTCH1 seems involved in the transition between an early, reversible phase of senescence in which the growth arrest is mediated via p21, to a later mostly irreversible phase driven by $p16^{INK4a}$ (hereafter p16), which is associated with a noticeable change in the composition of the SASP (Hoare et al., 2016). This phase shift may be important in understanding the differences in consequence between an acute onset of senescence compared to a chronic presence of senescent cells.

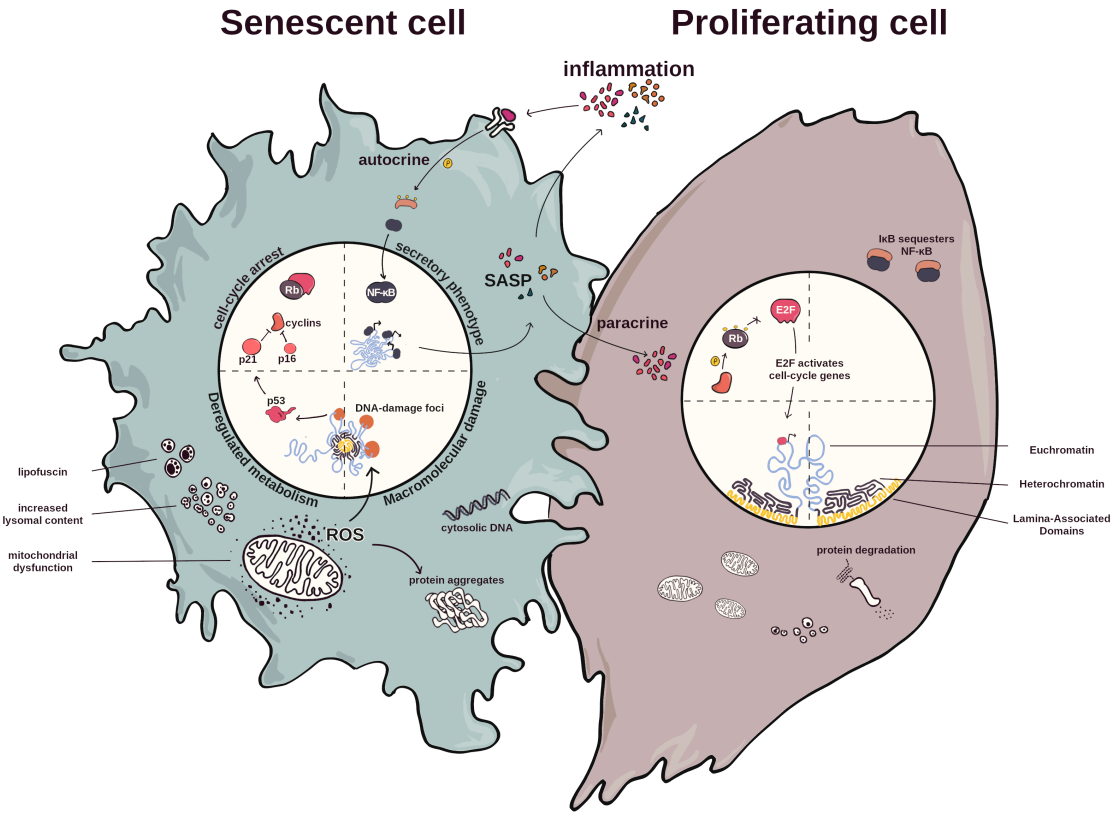

Fig. 3.1 **The 4 hallmarks are senescence: cell-cycle arrest, secretory phenotype, macromolecular damage and deregulated metabolism.** The cell-cycle arrest is largely mediated by p21 and p16, which inhibit the cyclin-kinase E & D. The hypophosphorylated Rb is then able to block E2F activity. The Senescence-Associated Secretory Phenotype (SASP) consists of several types of molecules. The inflammatory molecules help to recruit the immune system, but the SASP also as autocrine and paracrine effects. The mitochondrial dysfunction plays a central role through the creation of Reactive Oxygen Species (ROS) which attack macromolecules such as proteins and the DNA. This in turn leads to an increase in lysomal function.

In its acute, controlled and scheduled version, cellular senescence has been associated with a number of important physiological roles. In response to injuries or stress, cells turn senescent and the cytokines contained in the SASP help to recruit immune cells to the site of damages (Demaria et al., 2014), a process critical for the clearing of cellular debris. Furthermore, still in the context of wound healing, the ability of the SASP to induce stemness and the angiogenic factors contained within it accelerate the healing process, while limiting excessive fibrosis at the site of the injury (Jun and Lau, 2010). The SASP also has physiological role during development. In several locations such as the inner ear, there is evidence of a developmentally programmed senescence, which is followed by the clearing of these cells by macrophages. In the developmental context, blocking senescence is partially compensated by apoptosis, but leads to noticeable difference in tissue patterning (Muñoz-Espín, Cañamero, et al., 2013). On top of these physiological roles, a recent study provided a mechanism as to how the phase-transition between early and full senescence relates to the antitumorigenic role of senescence (Sturmlechner, C. Zhang, et al., 2021). They find that the early phase of p21-driven cell-cycle arrest places the cells under immunosurveillance, while the NOTCH1 mediated transition towards a p16-driven cell cycle arrest – and associated changes in the composition of the SASP – polarize the macrophages keeping the cell under surveillance, leading its clearance. In this sense, those two phases provide a grace period to the cell while keeping it under surveillance: if the DNA damage cannot be resolved in time, the cells is cleared away.

It is possibly when this clearance fails, and thus the SASP persists, that the chronic effects of it start to transform this acute, scheduled senescence more problematic. Since the activation of an oncogene is one of the main ways by which senescence can occur, senescence cells are often premalignant cells located in the tumor niche, and there several SASP factors promotes tumor growth through chronic inflammation, mitogenic signaling, stemness and angiogenesis (Schmitt et al., 2022). These effects have been evidenced by xenograft mouse models, in which grafting of senescent cells boosted malignancy and metastatic ability, an effect which was abolished after inhibition of NF-κB (Ohanna et al., 2011). Ironically for a secretome whose primary function is to attract immune cells, the matrix remodelling factors of the SASP have been shown to help shed the ligand NKG2D, which allows the cells to evade immune surveillance (Muñoz et al., 2019), with some proposing senescence as a major mechanism for cancer recurrence (Saleh and Gewirtz, 2022). This effect on cancer relapse might also be mediated by the so-called paracrine senescence: the SASP has the ability to induce senescence in neighboring cells (Juan Carlos Acosta et al., 2013), and here the ability of the SASP to promote stemness can drive the formation of new cancer stem cells (Schmitt et al., 2022).

Moreover, chronic exposure to the SASP may be crucial in aging and the onset of age-related dysfunctions. Observations have shown that senescence cells accumulate in the

body with age, although the reasons are probably multiple: shortening of telomeres with increasing divisions, accumulated protein damages, or a less effective immune system in old age (Van Deursen, 2014). Combined with paracrine senescence, the ability of senescent cells to induce senescence in their neighbors, this leads to a threshold theory of aging (Chaib et al., 2022), whereby once the burden of senescent cells passes a certain threshold, the self-amplifying loop isn't able to be resolved by the immune system anymore. Beyond this point, the individual is permanently exposed to chronic inflammation in a state known as 'inflammaging' (Franceschi and Campisi, 2014). This sterile inflammatory state can cause dysfunction in nearby and distant non-senescent cells, impairing tissue function and reducing regenerative capacity (Chaib et al., 2022).

This mechanism makes senescent cells an attractive target for therapies. Naturally, senescence attracted renewed interest when a team led by van Deursen demonstrated the use of senolytics, molecules targeting senescent cells specifically, and extended the lifespan of mice (Baker et al., 2011). But leaving aside ethical concerns over life-extension therapies – although uses for the treatment of diabetes is also under investigation – the dual nature of senescent cells invites caution in their use. It also invites further study in the exact role of senescence around cancer, literally as in its effect on the cancer niche, but also upon cancer onset and after therapies, to tease out the beneficial aspect from those we might want to suppress. But the complex, pleiotropic nature of senescence is unfortunately mirrored in the extreme heterogeneity of the phenotype.

3.2 The 50 shades of senescence depending on the trigger.

Heterogeneity has been a running theme in the literature about senescence since the beginning, to the point that the existence of one mechanism common to all types of senescence, a somewhat existential question for the field, is still unresolved (Hernandez-Segura, Nehme, et al., 2018). In this section go over the sources of this heterogeneity, and the different triggers for senescence.

The heterogeneity of senescence is first felt in the lack of specific biomarkers. Identifying those biomarkers has long posed a challenge due to their specificity and sensitivity. The development of the senescence-associated β-galactosidase (SA-β-GAL) assay was a milestone (Dimri et al., 1995), but it is associated with other cellular processes and states (Muñoz-Espín and Manuel Serrano, 2014), and the same is true for the more recently proposed lipofuscins, thus both tend to produce false-positives [Evangelou, 2016]. Likewise, gene expression patterns such as the upregulation of p21, p16, downregulation of lamin B1, and cell-cycle arrest assays are indicative but not exclusive to senescence (V.

Gorgoulis et al., 2019). p16, for example, is sometimes said to be the most specific marker (Liu et al., 2019), with some studies in vivo using $p16^{HIGH}$ cells as synonymous with senescent cells. However p16 is also upregulated in other processes such as macrophage polarization (W. Huang et al., 2022), while studies found that not all senescent cells display p16 activation (Hernandez-Segura, Jong, et al., 2017). Additionally, the presence of a secretory phenotype, while a key marker, can vary across different senescence models (Hernandez-Segura, Nehme, et al., 2018). This indicates that no single biomarker is entirely specific or sensitive to senescence, necessitating a more systematic approach using multiple markers. This lack of specific biomarkers for senescence is a significant hurdle in our understanding of senescence: the first evidence that senescent cells accumulate in vivo during aging is from (Dimri et al., 1995), more than 30 years after the discovery of senescence in vitro.

We have already encountered one major source of heterogeneity, namely the temporal one. Senescence, rather than a static endpoint, is a multi-step process, with cells which remain metabolically active. The two main stages differ by the pathway which mediates the cell-cycle arrest. The first stage of early senescence, immediately following DNA damage, is controlled mainly through the p53-p21 pathway. This is correlated with an early secretome that is largely composed of factors under the control of p21 (Sturmlechner, C. Zhang, et al., 2021), as well as downregulation of lamin B1. In this stage, senescence can be reversed upon inactivation of p53 (Beauséjour et al., 2003), allowing the cell to re-enter the cell-cycle. As we have seen, a phase shift after 7 to 10 day, which sees the down-regulation of NOTCH1, and the cell enters full senescence. It is characterized by a rise in the expression level of p16, another critical cycling-dependent kinase inhibitor of the cell cycle encoded by the gene *CDKN2A*. At this stage of full senescence, only cells with low levels of p16 can re-enter the cell cycle upon p53 inactivation (Beauséjour et al., 2003), thus showing p16's role in locking the phenotype. Van Deursen (2014) argues for a third stage of deep senescence, although some changes have been documented to occur when cells are cultured for longer – such as derepression of transposons (De Cecco et al., 2013), and changes in the SASP (Ivanov et al., 2013) – few experiments look at senescent cells over several weeks, thus the characterization of this state is sparse(Fumagalli et al., 2014).

This temporal evolution of senescence is an important consideration when studying senescence, but arguably an even greater source of heterogeneity stems from the source of DNA-damage. Senescence is fondamentally triggered by persistent DNA damage, and as such is intertwined with the DDR. And as we have seen in 2.3, different kind of DNA damage may result in different reactions of the DDR. The triggers of senescence can be broadly classified in three categories : telomere-induced, stress-induced and oncogene-

induced senescence.

Telomere-induced senescence, more commonly called replicative senescence, is caused by telomere attrition. It was the original type of DNA damage associated with senescence, described by Hayflick. Normal telomeres are capped with shelterin complexes which keeps the single-stranded overhang at the end of telomeres from being recognized as a DNA-damage site. When telomeres shorten due to cell division, it prevents the proper capping by shelterin, leading a persistent activation of the DDR machinery (Fagagna et al., 2003; Luke-Glaser et al., 2012; Correia-Melo et al., 2014). Telomere erosion also has a direct effect on gene expression, because it affects the telomere-position effect (K.-H. Lee et al., 2021), which relies essentially on the role that telomeres have in spreading heterochromatin marks. This type of senescence has most readily been associated with aging on a cellular level, as it seems to impose a cap on the number of divisions cell can go through. It is important to note in this context that individuals with constitutively long telomere are more prone to cancer in old age (Willeit et al., 2010).

Stress-induced senescence – sometimes termed premature – is a loose category regrouping any senescent state triggered by either acute or chronic stressors. In this case, the activation of the DDR results generally from exposition to cytotoxic compounds or endogenous stressors, both of which generally produce DNA-damage foci through the creation of reactive oxygen species (ROS) – although some drugs like Cisplatin damage DNA directly. ROS are a crucial pin in the mechanism of senescence in general, and stress-induced senescence in particular. It has been shown that the continued production of ROS is necessary for the maintenance of senescence (Correia-Melo et al., 2014). In fact, contrarily to the image that 'persistent DNA damage' might bring about in the reader's mind, the majority of DNA-damage foci are short-lived. They are instead constantly produced and resolved, a process which keeps the DDR engaged and thus leads to senescence (Passos et al., 2010). Stress-induced senescence includes a crucially important type, therapy-induced senescence, since the mechanism of action of both chemotherapy and radiation therapy is often the induction of genotoxic stress and it can induce senescence. The creation of senescent cells during cancer treatment is increasingly recognized as one of the major factors in therapeutic outcomes, with both anti- and pro-cancer effects (Schmitt et al., 2022).

Last but not least is **oncogene-induced senescence** (OIS). OIS was first described after the ectopic expression of an oncogenic version of RAS (M. Serrano et al., 1997). Finding that the expression of an oncogene triggered cell-cycle arrest was surprising, but further studies clarified this paradox. Senescent cells appear at early stages of tumorigenesis, but not in later, invasive tumors (Collado and Manuel Serrano, 2006; M. Serrano et al.,

1997), which cemented OIS as a barrier that tumors had to overcome before becoming malignant. Since the original discovery, more than 50 oncogenes have been shown to induce senescence (V. G. Gorgoulis and Halazonetis, 2010). Of all these oncogenes, the most studied models for OIS are the Ras proteins (H-RAS, N-RAS, K-RAS) and the kinase BRAF. The HRAS mutation G12V, and the BRAF mutation V600E have received particular attention, with the later shown to be present in human naevi (Michaloglou et al., 2005), and we have used both these models in the present work. For OIS, the DNA-damage comes from replication stress, as oncogene activation tends to involve error-prone DNA synthesis, double-strand breaks and accumulation of collapsed replication forks (Zhu et al., 2020), thus engaging the DDR which triggers senescence in the long run(Di Micco et al., 2006).

This represents only a brief overview of the heterogeneity of senescence, which in some very fundamental ways mirrors that of cancer. The type of damage, the ensuing DDR and the temporal evolution form a multilayered phenotype which is challenging to tease apart, requiring single-cell technology and microscopy experiments to take into account the cell-to-cell variability. Understanding it is necessary however ; OIS, for example, is likely to be a common milestone in the development of oncogene-driven cancers, thus a potential therapeutic target. As we have seen, the DDR is center-stage in this phenotype, and p53 is among its most important conductor, thus we need to spend some time looking into p53's role in senescence.

3.3 At the crossroads of fate: p53's role in the regulation of senescence.

With the complexity of the p53 gene regulatory network which we touched upon in the last chapter, it is perhaps unsurprising that its role in senescence is complex to sum up. However, the same wide gene regulatory network may explain in a large part the pleiotropy which characterizes senescence, and its contradictory characteristics. Given its centrality, p53 has long been a drug target for cancer, as well as many interactions that may keep p53 from functioning in cancer, such as its interaction with MDM2, its main inhibitor. However, harnessing the p53 gene regulatory network is made difficult in no small part because of its overlap with senescence, and the complex role of the latter in cancer. In this section we will go over some of the subtleties of p53 regulation happening during senescence, expanding from the crucial p53-p21 axis into its roles affecting the SASP, and those all evolve as senescence progresses.

Of course, the role of p53 starts with the cell-cycle arrest. P53 induces the expression of the cycling-dependent kinase (CDK) inhibitor p21 encoded by the gene CDKN1A (V. Gorgoulis et al., 2019). In turn, p21 leads to the inactivation of the E2F family of TFs (Aksoy et al., 2012), which are responsible for the cell-cycle, thus blocking it, generally in the G1 phase (Johmura et al., 2014). Importantly, p21's role isn't just to block the cell-cycle. It also participates in the repression of apoptosis, as the knock-down of p21 in senescent cells was shown to lead to 30% of cells going into apoptosis, although this wasn't the case in OIS (Yosef et al., 2017). This makes the p53-p21 axis one of the central pathway of senescence, with many trying to target this relationship for either cancer treatment or senolytics (Martin et al., 2020). Although, as we have seen, there is a phase shift after about a week, after which the cell-cycle arrest role is taken over by another CDK inhibitor, p16, encoded by CDKN2A, and its downstream effector pRb. Once the p16-pRb axis is online, downregulation of p53 will no longer lead to senescence escape on its own, instead demanding the simultaneous inhibition of p16 (Beauséjour et al., 2003). This is mainly due to the gradual heterochromatinization of the E2F target genes, such that inhibition of p53 and thus reactivation of E2F factors has no effects (Salama et al., 2014). Despite the more permanent role of p16, there is evidence that the p53-p21 axis is more responsive since the knockdown of p53, but not p16, inhibits the establishment of senescence after various triggers (Childs et al., 2014). Thus the p53-p21 axis is necessary for the establishment of senescence, but not for its maintenance. Accordingly, after a peak in the second day, the levels of p53 go down gradually, with one study finding it to go down almost to initial levels (Sturmlechner, Sine, et al., 2022). Interestingly, in this experiment, knock-down of p53 using shRNA at day 10 had no effects on p21 levels. The researchers propose that, in order to remain viable in the long run, senescent cells need to limit p53 activity, in particular its pro-apoptotic functions.

The changing p53 activity is also reflected in a changing binding landscape. In a ChIP-seq experiment, (Kirschner et al., 2015) measured the binding profile of p53 after a single day of exposure to the DNA-damaging agent etoposide, compared to two different 'chronic' conditions, one five days after the transfection of the oncogene H-RAS, the other expressing H-RAS and E1A simultaneously, where E1A keeps the cells proliferating and sensible to apoptosis. Interestingly, both chronic conditions were similar with each other, and different from the acute condition. After an acute up-regulation, p53 showed sharp peaks at promoters, while in chronic up-regulation, there was an over-representation of CpG-island linked genes, seemingly associated with genes involved in fatty-acid synthesis and RNA processing.

These changes during the development of senescence mirror the general distinction between acute and chronic exposure to the SASP that we have touched upon in the first section. It is thus fitting that p53's role in senescence extends beyond the cell-cycle arrest,

and influences the SASP.

While p53 is a key factor activated upon DNA damage, it is not at the root of the DDR. The two kinases ATM and ATR, on the other hand, are among the first responders, and act upstream of p53. ATM phosphorylates p53 directly at Serine 15, and indirectly at Serine 20, which inhibits its degradation by MDM2. But p53 is far from their only substrate, and they phosphorylate many other proteins which will ultimately lead to the activation of the key transcription factors of the SASP, including NF-κB (Sheekey and Narita, 2023; Kang et al., 2015) . This role places ATM/ATR upstream of both the cell-cycle arrest via p53 and the SASP via NF-κB, but it doesn't mean those two phenotypes are entirely decoupled. As a start, p53 inactivation increases the inflammatory elements and level of the SASP (Uehara and Tanaka, 2018), giving p53 an inhibitory function against the SASP. As a result, p53 loss of function in some cancers can lead not only to excessive inflammation (Uehara and Tanaka, 2018), but also to the wrong kind of inflammation, as the SASP in p53-null cells fail to activate macrophages (Lujambio et al., 2013). Thus the presence of p53 seems to have a important effect on the cell-to-cell communication aspect of the SASP, with important consequences: restoration of p53 in tumor leads to senescence, which not only stops proliferation but through this effect on the SASP triggers immune recognition (W. Xue et al., 2007; H.-A. Chen et al., 2023). At this time, we are still far from understanding the complete molecular mechanisms of this inhibition/modulation of the SASP, however we know that NF-κB and p53 are engaged in a complex cross-talk, in which they generally mutually antagonize each other, although they can cooperate, for example in the case of apoptosis and senescence via p300 (Carrà et al., 2020). Although it is interesting to note, not all p53 effect on the SASP are though to come through its direct interaction with the NF-κB pathways. p21 by itself has been shown to lead to the expression of a number of factors, through its effect on pRb (Sturmlechner, C. Zhang, et al., 2021). Surprisingly, despite the fact that both CDK inhibitors p21 and p16 act through Rb, p21 overexpression produced a secretome different from that of p16 overexpression, which led the p21-driven secretome only to contain the chemokine CXCL14, which the researchers found to be essential in the process of macrophage activation. While these show that p53 has an influence on the SASP and its functions, p53 is not at the core of this phenotype.

It is interesting to read this in light of the previously mentioned fact that p53 levels and p21 levels become progressively decoupled through senescence. Not a lot is known about why the p53 gene regulatory network changes over time during senescence, and why the same protein abundance doesn't produce the same effects in terms of transcriptional activation. Nevertheless, some mechanisms have been proposed. An interesting one gives center-stage to two different kind of nuclear bodies associated with senescence, namely PML bodies and DNA-SCARS – DNA Segments with Chromatin Alterations Reinforcing

Senescence (Sheekey and Narita, 2023). p53 is known to localize inside DNA-SCARS, while FOXO4 localizes in PML bodies, and the two bodies fuse together during senescence (Rodier et al., 2011). Recently, Baar et al. (2017) showed that disrupting the known interaction between FOXO4 and p53 releases the latter from the fused nuclear bodies, and lead to apoptosis. This has been proposed to lead to the p53-dependent activation of p21, and in this context it is important to note that the promoter of p21 contains two p53 binding sites flanking a FOXO4 binding site (Bourgeois and Madl, 2018). Cooperation of FOXO4 and p53 might extend beyond this co-binding, as both share the acetyltransferase p300 and are both able to recruit it to their binding sites, and both share the sirtuin SIRT1, which is a deacetylase. This may lead to a complex tug-of-war for acetylation of the p53 C-terminal domain, where acetylated p53 activates apoptosis and ROS production genes more strongly. It is shown that acetyl-mimicking mutation accelerate the entry into senescence, while mutations that abolish acetylation of lysine residues have the opposite effect of preventing entry into OIS (Bourgeois and Madl, 2018). This connection with SIRT1 is interesting on two fronts: sirtuins are NAD+-dependent, which means their activity is dependent on the energy metabolism, which is altered in senescence ; levels of SIRT1 gradually decrease during OIS, while they remain constant in another type of cell-cycle arrest, quiescence (Xu et al., 2020). These complex interactions between p53, FOXO4 and SIRT1 at the level of PTMs may prove to be important drugable targets for treatment (Xu et al., 2020).

As we see, the effects of p53 during senescence extend a lot beyond the cell-cycle arrest, with even non-cell autonomous roles in its modulation of the SASP which affects macrophage activation. Despite the temptation of summing up senescence via a limited number of ‘axis’ such as the p53-p21 axis, one is forced to add more and more interrelated players into those to explain the various dynamics observed as the phenotype progresses. But on top of interacting partners, another thing changes radically during senescence, the chromatin landscape. The remodelling of chromatin, sometimes quite extreme during OIS, likely plays a crucial role in the changing effects of p53 and other players in the nucleus.

3.4 The shifting landscape of chromatin at the onset of senescence.

The onset of senescence is accompanied by profound changes in the chromatin landscape. The reader should not be surprised to hear at this point that these can differ substantially between senescent cells depending on the trigger. The field of the 4D nucleome has been grappling from the beginning with the interdependence between chromatin conformation and transcription, akin to the chicken-or-egg scenario. Fittingly, many have looked at developing embryos to answer this question, but senescence, the end of a cell life, offers

an intriguing alternative viewpoint. This section delves into these aspect, going over the major changes to the 3D chromatin structure, and how they might relate to changes in transcription.

Since DNA damage is a fundamental trigger for senescence, ***DNA-damage foci*** are a hallmark of the chromatin landscape in senescence. DNA-damage foci are condensates of molecules which form around the site of DNA damage, typically associated with the phosphorylated histone variant γH2AX. Despite the term "persistent" DNA damage often used to describe senescence, suggesting persistent foci, live-cell imaging studies have shown that most DNA-damage foci are short lived (<15h) instead being continuously produced and resolved (Passos et al., 2010). This process of constant generation is attributable to the generation of reactive oxygen species within senescent cells, which have the crucial consequence of keeping the DDR active.

Although some are transient, some DNA-damage foci called DNA-SCARS last longer. We have encountered these foci in the previous section, and their structure differs from typical DNA foci, lacking some DNA-repair proteins such as RAD51, but are enriched with DDR mediators such as p53 (Rodier et al., 2011). This study also revealed that frequently coalesce with promyelocytic leukemia (PML) bodies, reinforcing the notion of their distinctiveness from conventional DNA damage foci. Their persistence, and the presence there of the significant interaction between p53 and FOXO4 suggests a role in the senescence phenotype.

Telomeres, the terminal regions of chromosomes, consist of 5'-TTAGGG-3' repeats capped with a G-rich overhang of 50-400 nucleotides. Typically, telomeres are bound by Shelterin complexes, which protects them against being falsely recognized as DSBs (Luke-Glaser et al., 2012). However, with recurrent replication cycles, telomeres begin to de-protect, inducing telomere dysfunction-induced foci and activating the DDR and triggering replicative senescence. Interestingly, a study found that, in replicative senescence, the short-lived DNA-damage foci are typically in the chromosome body, while longer lasting foci are situated at telomeric regions (Fagagna et al., 2003; Hewitt et al., 2012).

Those telomere-associated damage foci aren't the only affect of telomere attrition. The shortening of telomeres affect gene expression via the telomere position effect (K.-H. Lee et al., 2021). Through physical interaction with neighboring regions, they facilitate the spreading of repressive chromatin marks, suppressing the expression of adjacent genes. As telomeres shorten, this repressive effect dwindles, which can lead to increased expression of formerly silenced genes. The reactivation of hTERT in cancer cells is hypothesized to be a potential consequence of telomere shortening (K.-H. Lee et al., 2021).

The detachment of Lamina-Associated Domains (LADs) is another characteristic change in the chromatin landscape. LADs represent chromatin sections bound to the nuclear envelope, found in 1,000-1,500 regions in mouse and human cells, covering over a third of the genome. These domains, ranging in size from 10 kb to 10 Mb (average 0.5 Mb), are mainly transcriptionally inactive and have fewer genes. At the border of LADs, there is an enrichment in H3K9me2&3 and H3K27me3 histone modifications, often embracing human pericentromeric heterochromatin (Steensel and Belmont, 2017).

Lamins, primarily B1, but also B2, A, and C, play significant roles in anchoring chromatin to the nuclear envelope. Although the involvement of Lamin A/C is debatable, the function of Lamin B1 in facilitating LAD attachment to the nuclear envelope, through the Lamina B receptor (LBR), is generally accepted (Steensel and Belmont, 2017). Decreased Lamin-B1 (LMNB1) expression induces LAD coalescence at the nuclear center, and down-regulation of LMNB1 is a marker of senescence (V. Gorgoulis et al., 2019). Interestingly, this link is causal, in the sense that Lamin-B1 depletion can trigger senescence (Lämmerhirt et al., 2022), although it is not known if LMNB1 re-expression could recover reverse senescence.

Lamin-B1 downregulation is further associated with large changes in histone marks, with changes in the distribution of H3K9me3 and H3K27me3, with the appearance of large contiguous stretches of both chromatin marks, but H3K27me3 was also depleted in large contiguous patches, in regions which were previously around LADs (Shah et al., 2013). Given that these domains contain pericentromeric regions rich in LINE and SINE transposons (Steensel and Belmont, 2017), the detachment of LADs may facilitate their reactivation seen in senescence, an interesting fact considering p53 binds LINE elements.

Senescence-Associated Heterochromatin Foci (SAHF) represent a distinct form of chromatin organization that is characteristic of senescent cells, densely compacted chromatin structures (Narita et al., 2003). SAHF are macroscopic domains of densely compacted chromatin that can be seen easily under a microscope, distinguished by exclusive patterns of histone modifications, specifically the overabundance of the repressive mark H3K9me3 and the paucity of the active mark H3K4me3 (Adams, 2007).

The formation of SAHF is thought to come from the detachment of LADs. However, the loss of LMNB1 is near omnipresent in senescence, while SAHF are mainly restricted to OIS, casting some doubts on the precise mechanism. Figure 3.2 tries to illustrate this change in organisation. SAHFs are thought to contribute to the establishment and maintenance of senescence by silencing proliferation-promoting genes. This notion is supported by studies demonstrating that key regulators of the cell cycle, including E2F target genes, are sequestered within SAHF in senescent cells (Narita et al., 2003).

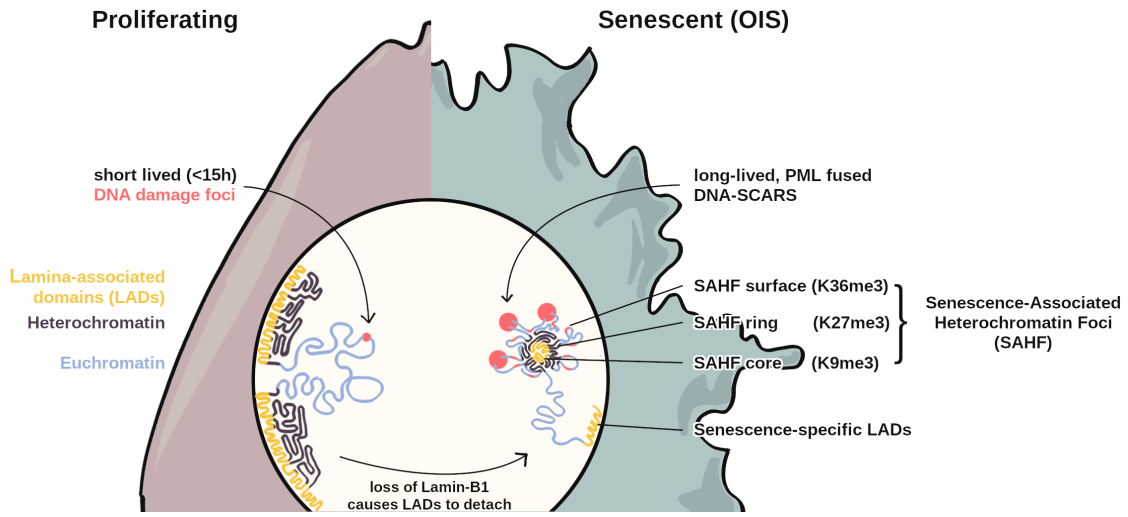

Fig. 3.2 **Some of the chromatin changes featured in oncogene-induced senescence (OIS).** Under normal condition, chromatin tends to be separated into dense heterochromatin compartment tethered to the nucleus enveloppe by lamina-associated domains, and more open compartments of euchromatin. However, senescence auses the down-regulation of Lamin-B1. This impact chromatin organisation in all case, but particularly under OIS the now untethered LADs assemble in a big heterochromatin domain. Histone marks on the domain are organized in concentric rings, with the surface enriched in H3K36me3, a mark associated with DNA repair.

Despite their repressive nature, SAHF exhibit a multi-layered complexity of histone modifications, revealing a more intricate role in gene regulation than initially suspected. The chromatin landscape within these foci is anything but quiescent; instead, it hosts a dynamic interplay of both repressive and active histone marks that govern the transcriptional profile of the senescent cell (Chandra, Ewels, et al., 2015). The repressive histone marks associated with SAHF, such as H3K9me3 and H3K27me3, are essential for maintaining the compact structure of SAHF and silencing pro-proliferative genes (Adams, 2007). Conversely, active histone marks are also present within SAHF. Notably, H3K27ac, typically associated with gene activation and enhancer regions, has been observed at the enhancers of SASP genes (Cruickshanks et al., 2013). This active mark within the seemingly repressive SAHF domain indicates a paradoxical role of SAHF in promoting the expression of certain genes, despite their overall repressive nature. Moreover, distinct patterns of histone marks form concentric layers within SAHF. A study by Chandra, Kirschner, et al. (2012) showed that the core of these foci is enriched in the repressive H3K9me3 mark, surrounded by a ring of the variant histone macroH2A, with an outer layer enriched in H3K27me3. This layered structure suggests a highly organized arrangement of chromatin within SAHF and might be instrumental in fine-tuning the transcriptional output of the senescent cell.

While the presence of SAHF is a widely accepted hallmark of senescence, it is important to note that not all types of senescent cells form these structures. The formation of SAHF appears to be a more common feature of OIS compared to replicative senescence.

Furthermore, the absence of SAHF in some cells displaying senescent traits and a lack of proliferative capacity suggests that these structures are not obligatory for the senescent state (Chandra, Kirschner, et al., 2012). A variety of factors besides LADs detachments have been described in the formation of SAHF, such as pRb (Narita et al., 2003), BRCA1 and BRG1 (Tu et al., 2013) and, interestingly, PML bodies (R. Zhang et al., 2005). This last one is interesting since we know that p53 is enriched in PML bodies, while BRG1 is part of the SWI-SNF chromatin remodeling complex that p53 is known to interact with and recruit (see section 2.4).

As we see, the onset of senescence brings about many changes in chromosome conformation and the whole nuclear landscape in general. These changes are, to a degree, consistent with what can be seen in aging cells, are progeroid syndromes, which also sees perturbation at the nuclear lamina and an associated opening of previously repressed chromatin. The advent of Hi-C method offers an additional way to probe this question and the picture here described is only beginning to be investigated. Investigating the effect of this massive chromatin organization can offer important insights into senescence, but also for the more general question of the influence of chromatin organization on gene expression, especially in the case of OIS where the reorganization happens quickly after the loss of Lamin-B1. Here we find an opportunity to study the multifaceted relationship between p53 and chromatin changes.

Aim of the work.

The large amount of DNA inside eukaryotic cells poses fundamental problems for gene regulation. One of these is the lack of specificity that TFs exhibit compared to what they would require in order to target a restricted number of sites, to the point that information theory predicts tens of thousands of spurious occurrence of their motifs within the DNA. The second related problem is that TFs must swift through a huge amount of non-specific DNA, on top of discriminating between false motifs and actual enhancers, before they can find a suitable target. As evidenced by our presence, however, both these problems are obviously resolved somehow, and TFs bind specific targets, in a context-appropriate manner, in a reasonable amount of time.

The target search mechanism is the name for the strategies evolved for TF to find their targets faster than would be expected. The solution generally accepted in bacteria and yeast turns the foe into a friend, making use of chromatin to slide in one dimension, rapidly scanning the genome in the process called facilitated diffusion. The nucleus of mammals, however, is significantly more complex, and it is not known to what extend all the lessons learned studying yeast can be applied. In vitro studies demonstrated the ability of some TFs like p53 to slide on DNA using its C-terminal domain, similar to the yeast case. However, the presence of a large number of histones complicates the story. For some TFs, alternative ways of periodically reducing the dimension of the search have been proposed, using condensates for example.

However, p53 seemingly doesn't form condensates. At the same time, being a pioneer transcription factor, p53 is not restricted by inaccessible chromatin like some other transcription factors would, which makes heterochromatization is a less compelling explanation for the increased specificity: hiding spurious motifs there does not affect p53's search all that much. This pushes us to find alternative explanations, one of which is chromatin structure. For example, the formation of domains of dense chromatin, as seen in interphase nucleus, creates a surface between chromatin dense and chromatin poor regions. Interestingly for the target search mechanism, diffusion along a surface constitute a reduction in the dimension of the search.

The study of the target search in vivo relies in large part on live-cell imaging, with which we can analyze the diffusive behavior of molecules and detect those periodic reduction in dimensionality characteristic of facilitated diffusion, but correlating these movement with other features can be challenging. Our lab developed a multi-modal microscope combining the ability of single-molecule tracking with super-resolution microscopy, allowing us to image the chromatin landscape with satisfying resolution and probe the interaction between p53 and chromatin in live cells and in real time.

Previous work from our lab looked at the influence of p53's C-terminal domain on this interaction, and for my PhD I take the opposite approach of perturbing chromatin structure instead. I take advantage of oncogene-induced senescence, a process of permanent cell-cycle arrest in which p53 is deeply involved, during which a dramatic reorganization of chromatin occurs. Large senescence-associated heterochromatin foci appear, with a clear delimitation between a dense core and the rest of the nucleus, depleted in chromatin. This model is poised to reveal clearly the function, if it exists, of the surface of chromatin domains.

Chapter 4

Single Molecule Tracking of p53 in untransformed normal cells.

Resolving individual molecules with fluorescence microscopy is not trivial as soon as the concentration of molecules starts rising. The first attempts at obtaining separable, diffraction-limited spots used total internal reflection fluorescence microscopy, a technique which limits the volume of sample illuminated to a thin slice close to the glass. This means that single-molecule tracking (SMT) was long restricted to *in vitro* studies (Tafvizi et al., 2011; Sternberg et al., 2014), studies in small organisms like bacteria or yeast (Elf et al., 2007), and studies of membrane-bound proteins which are naturally confined to a slice (Sako et al., 2000). Recent advances in fluorescent labeling and microscopy illumination schemes have allowed the SMT to expand beyond the membrane and look at proteins freely diffusing in the cell.

Some studies have used SMT to look at p53 dynamics, but most of these are *in vitro* (Tafvizi et al., 2011). When looking at TF dynamics *in vivo*, most studies used cancer cells (Mazza et al., 2012; Gebhardt et al., 2013; Mazzocca, Loffreda, et al., 2022) since those are much easier to work with. However, as we have seen in section 2.4, Botcheva et al. (2011) found significant difference in the binding profile of p53 in normal cells compared to cancer cells, thus we should not necessarily expect the same dynamics of p53 in those two cells.

In addition to being interesting, the use of untransformed cells is required by our study of senescence, which implies that the original cells should be a normal proliferating cell. Since the reaction of p53 might depend on its concentration, we found it necessary, in order to obtain biologically relevant results, to avoid the use of transfection and instead obtain knock-in cells replacing the endogenous p53 loci with a fluorescent tagged version.

This chapter presents the cell lines used in the rest of the thesis and characterizes them, and presents the results of our SMT experiments on untransformed cells. We find

significative difference between the behavior of p53 in normal human fibroblasts compared to cancer cell lines.

4.1 Generation and characterization of the cell lines.

For our study of how the formation of SAHF affects p53 dynamics, we chose two cell lines, IMR90 and BJ cells.

IMR90 cells are fibroblasts derived from normal lung tissue. Our particular cell line has been developed by (Innes and Gil, 2019), and generously gifted to us by Peter Adams (University of Glasgow, UK). It uses a estrogen receptor ligand binding domain, fused to the oncogenic version of the RAS gene, $H-RAS^{G12V}$, one of the most extensively studied oncogene whose activation triggers oncogene-induced senescence (OIS), leading to cells to withdraw from the cell cycle in approximately seven days due to the activation of p53. (Dimauro and David, 2010). These IMR90 ER:RAS cells can thus be easily induced into senescence by being cultured in a medium containing 4-hydroxy-tamoxifen (4-OHT), and the cells become senescent after a week of treatment. The induction of OIS in this cell line is highly reproducible, the senescent cells thus obtained show SAHF in a consistent manner, which are enriched in the heterochromatin mark H3K9me3 as expected (fig 4.1.A). We find the increased presence of DNA-damage foci after the induction of OIS (fig 4.1.B) and 87% of cells are β-galactosidase positive (fig 4.1.C-D). Finally, we measure by qPCR the up-regulation of both *CDKN1A* (p21) and the SASP genes (fig 4.1.E-F).

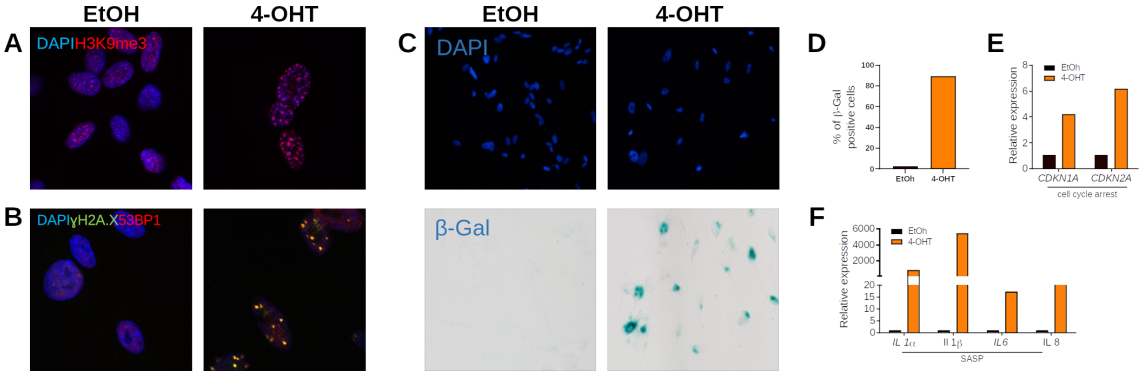

Fig. 4.1 **IMR90 ER:RAS cells induced into senescence display hallmarks of OIS. A:** Immunofluorescence of H3K9me3 (red) shows reorganization of heterochromatin into large foci. **B:** Immunofluorescence of γH2AX and 53BP1 shows the apparition of DNA-damage foci. **C:** Cells are β-galactosidase positive after 8 days of treatment. **D:** Quantification of the β-galactosidase assay. **E and F:** qPCR for hallmark genes of senescence, cell-cycle arrest and SASP respectively.

The advantage of IMR90 is that it is a well characterized cell line of fetal human lung fibroblasts. For our particular purpose of studying the relationship between chromatin and p53 behavior, several teams have performed ChIP-seq of p53 on these cells (Botcheva

et al., 2011; Kirschner et al., 2015; Morgan A. Sammons et al., 2015; Aksoy et al., 2012). In addition, there are RNA-seq data after p53 activation and in the context of senescence (Hernandez-Segura, Nehme, et al., 2018), that can be used to relate p53 binding to gene expression.

However, despite these advantages, the IMR90 ER:RAS unfortunately have a resistance to geneticin, which made the antibiotic selection impossible with the p53 vectors for generating fluorescent p53 knock-ins impossible, since they also rely on geneticin resistance. As such, replacing the endogenous loci with a fluorescent tagged version proved challenging. We briefly explored the possibility of ectopic expression with transfection, but being non-transformed cell lines, IMR90 are very hard to transfect and the transfection efficiency stubbornly remained <1% even after protocol optimization. This is in addition to the major drawback of over-expression when using transfection, which bring us well outside of physiological conditions, limiting the interpretability of our results. For these reasons, we decided to use another cell line for the knock-in, while we restrict the use of IMR90 ER:RAS cells to experiments besides SMT, such as characterizing the gene expression levels of key components of the senescence phenotype.

For the knock-in of a fluorescent tagged version of p53, we used a line of BJ cells, also fibroblasts but this time obtained from the foreskin of newborn males. For SMT, the preferred tag is not GFP, but HaloTag, a 33 kDa peptide that is fused to the protein of interest (England et al., 2015). The reason HaloTag is favored for single-molecule tracking is favored rather than GFP for example is two fold. First, HaloTag is not fluorescent by itself, but can form a covalent bound with ligands containing organic dyes. This crucially allows us to use limiting concentration of the ligand to label only a small fraction of the molecules, achieving the low labeling limit (tens of molecules per nucleus) required for SMT. Second, the organic dyes are more stable than GFP, allowing for a longer tracking.

For the purpose of the knock-in, we opted for the CRISPR-Cas9 nickase technology (Ran et al., 2013), which substitute the DSB induced by Cas9 with two separate single strand breaks or nicks, in close proximity. This limits off-target effects by effectively using two separate 20 nucleotide guide RNA. This double-nick induces homologous recombination with high efficiency – that is to say, up to 10% – which allows for the insertion of a template, in theory at least. In practice, normal cells that are not overly proliferating remain very hard to knock-in, even more so when the insertion relates to sensible or lethal genes (Murakami et al., 2020), as is the case for p53, compounded by anoikis, a form of programmed cell death fibroblasts are sensible to that occurs when the density of cells is too low to allow for the cell-to-cell interactions necessary for them to survive. As a result, obtaining this BJ cell line took well over 2 years of protocol

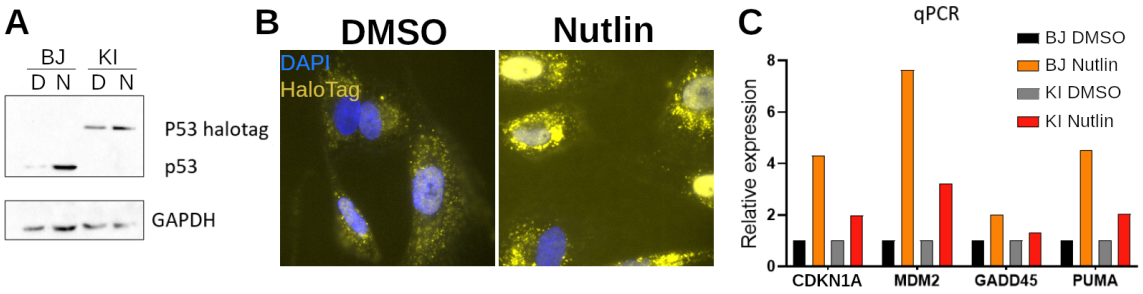

Fig. 4.2 **Characterization of pool p53-HaloTag knock-in (KI) A:** Western-blot of the parental BJ cell KI pool show no endogenous p53 in the pool. **B:** Fluorescent microscopy after HaloTag labeling shows p53-HaloTag correctly localized in the nucleus. Note the basal activation of p53 in control conditions. **C:** qPCR of some p53 target genes shows p53-HaloTag capable of transcriptional activation after treatment with nutlin.

optimization and repeated attempts. Much thanks are due to Daniela Gnani, who persevered in this endeavor for much longer than I did and finally devised a successful protocol.

After many repeated attempts, we validated a pool of halotag-p53 knock-in (thereafter, KI) cells. Western-blot showed no presence of then endogenous p53 (fig 4.2.A). Upon nutlin treatment for 3h, the pool showed up-regulation of both p53 and its target genes, both at the protein and mRNA level (fig 4.2.C), although both the Western blot and the microscopy showed that p53-HaloTag is expressed at a basal level even in untreated conditions (fig 4.2.B). Microscopy showed p53 correctly located in the nucleus upon activation. Some clusters of bright signal seem to appear in the cytoplasm. These clusters, absent in typical immunofluorescence against p53 in normal cells – in which p53 signal appears smooth, and typically brighter in the nucleus – appear at different wavelength than the 561nm laser used for the illumination of HaloTag (see Methods), and they are typically immobile when looked in live-cells, behaving differently than typical p53 molecules. Finally, while the qPCR shows that the induced expression of p53 target genes is lower across the board, their relative proportion are consistent between the parental cells and the KI pool.

Moreover, we characterized the onset of senescence in both the parental cell line and the pool. These primary cells do not possess an inducible oncogene, and thus the protocol for is different than for IMR90 ER:RAS. We rely here on the lentiviral transduction of $BRAF^{V600E}$, a well-studied oncogenic mutation of the B-Raf kinase. After the transduction, the cells are cultured for 10 days and progressively settle into OIS (fig 4.3.A). The SAHF appear progressively over this period, and at day 10 they form clear heterochromatin foci, and DNA-damage foci appear at the same time.

Unfortunately, we had some difficulties while characterizing OIS in the KI pool. The β-galactosidase assay resulted in the cells dying, while the qPCR produced discordant and inconsistent results, both compared to the parental cells and in between runs. A large part of the discrepancy in the activation of hallmark genes of senescence compared to the

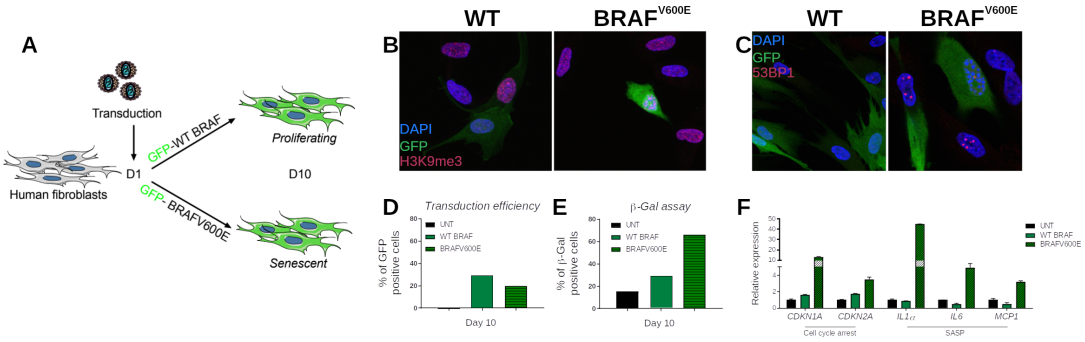

Fig. 4.3 **BJ cells transduced with BRAF^{V600E} shows the hallmarks of senescence.** **A:** Protocol of senescence induction. The vector contains a GFP reporter to identify transduced cells. **B:** Immunofluorescence of H3K9me3 (red) shows reorganization of heterochromatin into large foci. **C:** Immunofluorescence of γH2AX and 53BP1 shows the apparition of DNA-damage foci. **D:** Quantification of the efficiency of transduction via cell-sorting. **E:** Quantification of β-galactosidase assay shows above 60% of positive cells. Note the proportion is higher than the efficiency of transduction thanks to paracrine senescence. **F:** qPCR shows upregulation of hallmarks genes for cell-cycle arrest and SASP.

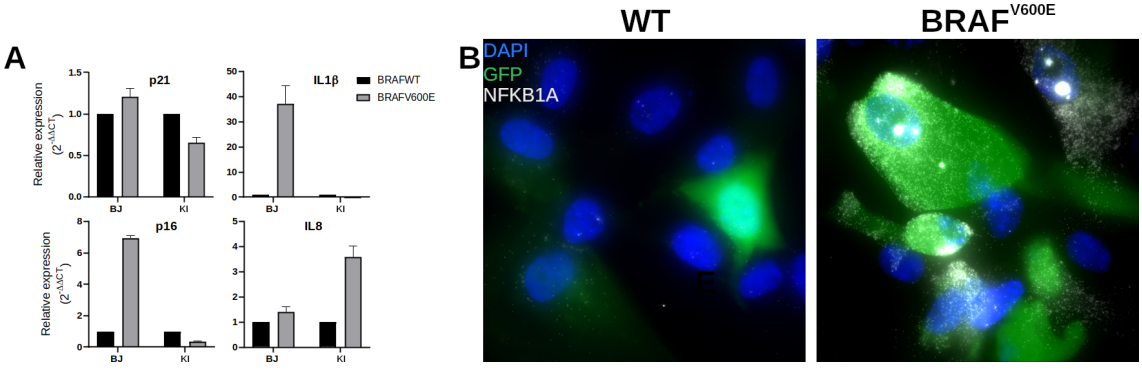

Fig. 4.4 **Transduction with oncogenic BRAF^{V600E} in the p53-HaloTag KI pool shows conflicting results.** **A:** qPCR of hallmark genes of senescence shows downregulation of cell-cycle genes in the KI pool after transduction. **B:** smFISH for the p53 target gene *NFKB1A* shows upregulation after transduction.

parental cell line is probably due to the basal activation of p53 in the KI pool (fig.4.4.A). Thus even as the expression of p21 inside the KI pool seems to decrease upon senescence induction, it also starts from a very expressed state. Some SASP genes are upregulated while others are down-regulated, while smFISH against *NFKB1A*, a gene regulated by boh NF-κB and p53 (Fischer, 2017), shows upregulation (fig.4.4.B).

We attribute these results to the high basal activation of p53, but more importantly to the heterogeneity inherent to working with a pool of cells. As such, we held off further characterization until a clone could be isolated, but the task proved challenging. Given the anoikis fibroblasts are subject to, they do not take kindly to being isolated into different wells. For this reason, the experiments presented in the rest of the thesis have been performed on the KI pool. This represents a limit to their interpretability, since we cannot guarantee the cells are isogenic. Fortunately the use of microscopy alleviate some of these concerns: if cells in this pool exist that do not possess p53-HaloTag, they will simply remain dark and thus not be imaged.

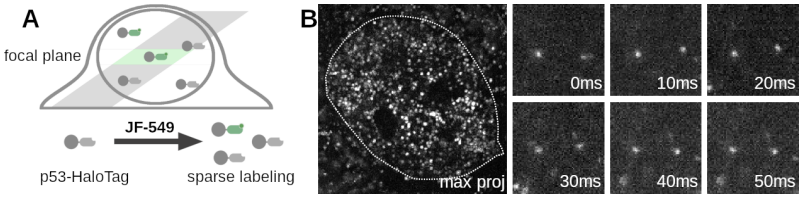

Fig. 4.5 **Single-Molecule tracking resolves single molecules of p53-Halotag inside live human fibroblasts. A:** SMT protocol. A combination of inclined illumination and sparse labeling limits the fluorescence given off by molecules outside the focal plane. **B:** Maximum projection of an SMT acquisition of p53 inside a nucleus, with the nucleus outlined. Details on the right shows a sequences of images instead, with single-molecules diffusing. Scale bar: 5µm

4.2 The dynamical behavior of p53 in normal fibroblasts differs from cancer cells.

Having obtained a pool of cells with HaloTag-p53, we move on to measure the dynamical behavior of p53 inside those using SMT. In a first experiment, we want to compare the behavior of p53 in basal and activated conditions. To do this, we use a treatment with Nutlin-3a (thereafter, nutlin), a small molecule which disrupts the interaction between p53 and its inhibitor MDM2, leading to an increase in p53 levels.

Briefly, the cells are seeded in LabTek chambers suitable for live-cell microscopy on the day prior to the experiment. For the treated condition, I change the medium with one containing 10nM of nutlin for 4h. The medium is then replaced with one containing 1nM of the HaloTag ligand JF-549, a limiting concentration.

The combination of sparse labeling and Highly Inclined and Laminated Optical sheet (HILO), allows us to fit individual diffraction-limited spots in the acquired images, after only bleaching the molecules for a few frames. The frame interval is 10ms, of which only 5ms are illuminated to reduce motion-blur and bleaching, for 20s of total acquisition – although some crowded cells are imaged twice). After acquisition, the individual spots in each frames are connected into tracks using the TrackMate plugin of ImageJ (Tinevez et al., 2017), with only the frames where molecules are sufficiently distant from one another kept for the analysis, to minimize errors in tracking (see Methods).

In a first attempt to analyze the movement of molecules, we look at the distance traveled by all molecules from frame to frame, and plot them in an histogram. For a single population of molecules diffusing with diffusion coefficient D, we would expect the probability of moving by a distance r between time points separated by $\Delta T = 10ms$ to be of the form $P(r,\Delta T) = \frac{r}{2D\Delta T} \cdot e^{\frac{-r^2}{4D\Delta T}}$. Thus fitting the histogram of displacement is a way to measure the coefficient of diffusion. This method is often preferable for SMT, as the MSD curve, which requires us to look at longer trajectories, is subject to biases when analyzing the short trajectories caused by fast diffusing molecules exiting the limited thickness of

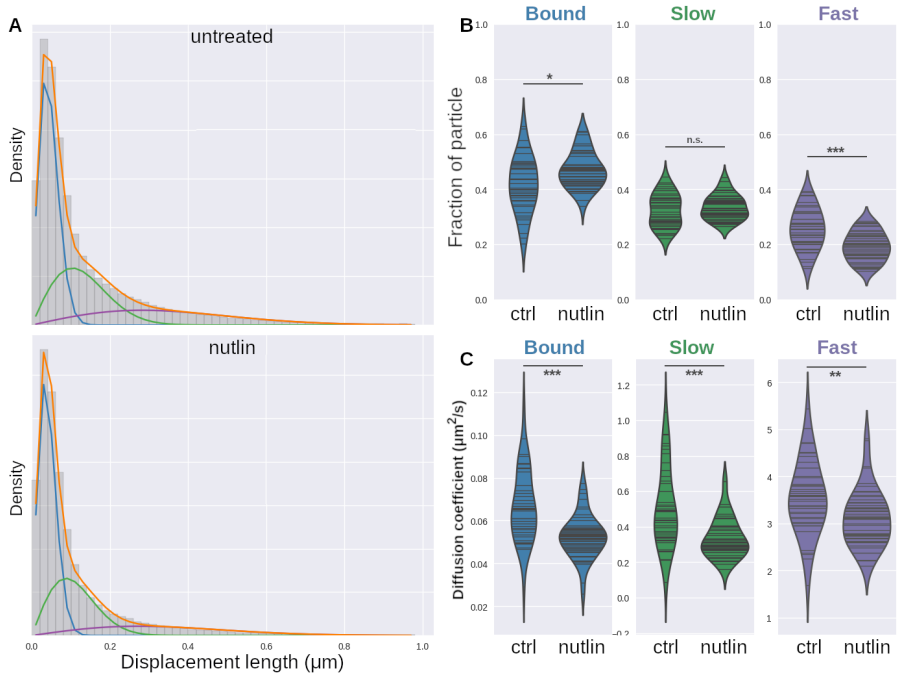

Fig. 4.6 **Histogram of displacement shows evidence for three sub-populations of p53 molecules which differ by their kinetics. A:** Histogram of the displacement length observed for p53-HaloTag diffusing in cells treated with DMSO (top) or nutlin (bottom) **B** Fraction of cells in each population at the single-cell level. **C:** Diffusion coefficient for each population at the single-cell level. N_{ctrl}: 14, N_{nutlin}: 11, each line in the violin plot is an acquisition.

our observation slice. When plotting and fitting such an histogram, we find it to be the sum of distinct sub-populations of p53 molecules, each with their own diffusion coefficient and thus a differing histogram (fig. 4.6.A), which remains true before and after treatment with nutlin. To objectively determine the optimal number of sub-populations, we employed the Bayesian Information Criterion (BIC). By utilizing this criterion, which penalizes model complexity while rewarding goodness of fit, the data indicated a preference for a model with three distinct sub-populations of p53 molecules. This result is consistent with previous reports (Mazza et al., 2012; Loffreda et al., 2017).

The kinetic parameters and relative fraction of each population are extracted with a model taking into account the bias of higher defocalization for fast molecules, such that we would under-count long displacements because they have a higher chance of ending out of focus. This gives us diffusion coefficients of 0.05, 0.38 and 3.7$\mu m^2/s$ for each sub-population, which I will call bound, slow and fast respectively from now on. While one might expect a null diffusion coefficient for bound molecules, there are two reasons to expect a non-zero value: (1) even with zero movement, the localization error would give an apparent movement back and forth around the real position of the molecule and (2) the chromatin itself moves at around this speed, which we can see by looking at histone H2B with SMT, which in our lab we have measured at 0.06$\mu m^2/s$ for the same 10ms frame interval.

The treatment with nutlin has an effect on the relative proportion of the different populations (fig.4.6.B) of reports the relative fraction of each sub-population on a single-cell basis. As expected, the bound fraction is higher after the nutlin treatment, although this fraction of bound molecules is significantly higher than previously reported figures, including from our own lab. In those cells, which were from cancer cell lines, the fraction of bound molecules was typically around 20% in basal conditions, and consistently under 30% even after irradiation of the cells, while we find more than 40% of p53 molecules bound even in basal conditions. A part of this difference is surely due to the basal activation of p53 which we observed in our pool of cells (fig.4.2). However, this is a less compelling explanation for the activated condition: p53 levels are indeed higher upon nutlin treatment, and the bound fraction remains higher than any previously reported figure.

The diffusion coefficient of p53 is also consistently lower in activated conditions, regardless of the sub-population (fig.4.6.C), however this is very hard to interpret. Indeed, many different biological mechanisms could be responsible for it, from an increase in the tetramerisation rates to any of the co-factors which can interact with p53 (see chapter 2.3). The same difficulties apply to the slow moving molecules.

While the histogram fitting approach provides us with essential insights into the dynamical behavior of p53 and its variations across different conditions, it fundamentally represents a bulk analysis method. However, without the ability of tell which molecule belongs to which population, we cannot analyze their behavior separately, nor can we correlate it with a reference image, which is our ultimate goal. Thus I need a way to segment the tracks obtained with TrackMate into the different sub-population. For this purpose, I will use a Hidden Markov Model (HMM), and try to highlight the difference in diffusive behavior between the slow and fast moving p53 molecules.

4.3 Slow p53 molecules show signs of facilitated diffusion.

Many properties of the different p53 sub-population can only be analyzed if we assign particular state to the observed molecule's position and jumps. We chose to use vbSPT, a solution developed by Persson et al. (2013). It tries to fit a HMM to the data, with different states characterized by their distribution of displacement, and a probability to switch between states at any time point.

Contrarily to some method, it is thus able to segment within a single track – such that we can observe the transition between a diffusive state and a bound state – however it uses the Markov assumption of a memory-less process, which is somewhat inaccurate for bound molecules. We use vbSPT to segment the tracks into three separate populations.

The diffusion coefficient of the fitted HMM are in line with the estimation with the histograms of displacements, although slightly lower which is probably due to the

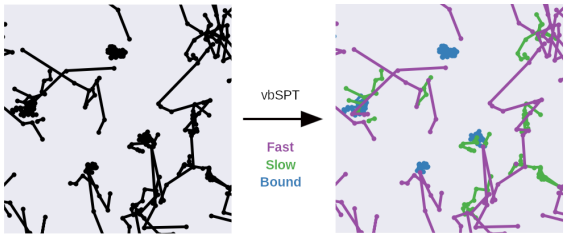

Fig. 4.7 **vbSPT allows for the separation of the 3 sub-populations according to their diffusion coefficient.**

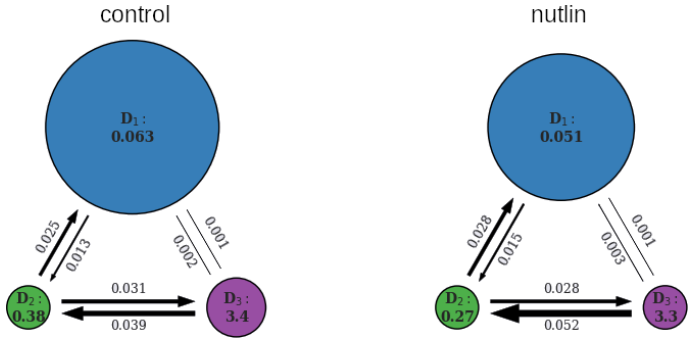

Fig. 4.8 Fitting the vbSPT model reveals asymmetric transition rates between the different state of diffusion. **Left:** vbSPT fit on control cells. Right: vbSPT fit on nutlin treated cells.

defocalization bias not being taken into account. The transition rates give an interesting first observation: fast molecules overwhelmingly go through a period of slow diffusion before binding ; binding directly from the fast state is strongly disfavored.

Once the tracks have been separated according to vbSPT's best estimation, we can plot the distribution of angles between consecutive displacement. This is done because free diffusion is expected to be isotropic, while a departure from isotropic diffusion can underlie anomalous diffusion (see chapter 1.4). For example, the over-representation of near 180° angles is evidence of the frequent backtracking typical of compact searches (Hansen et al., 2020).

From this analysis we see that, while the sub-population of fast p53 diffuses mostly isotropically, the slow diffusing molecules are anisotropic, diffusing back on their track about twice as often than they continue forward. This preference doesn't change in intensity upon nutlin treatment ($F_{180/0}$ = 1.85 for the control, $F_{180/0}$ = 1.91 after nutlin treatment).

Since anisotropic diffusion could be the result of a number of underlying phenomenons, we need to go further. One way to differentiate between different molecular dynamics is to look at the interdependency between this distribution of angles and the distribution of jumps. The existence of a relationship between the two, as in the case where anisotropic behavior peaks for a certain displacement length, is characteristic of facilitated diffusion (fig.4.10.A). Furthermore, (Hansen et al., 2020) demonstrated in the case of CTCF that in

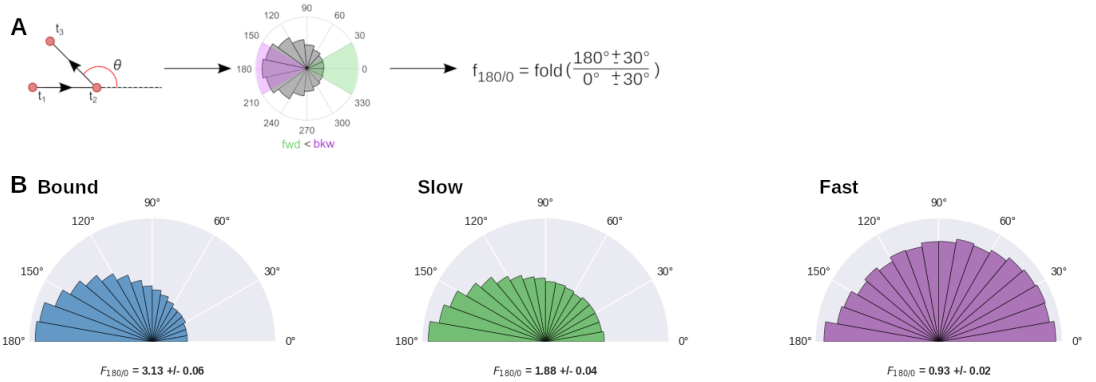

Fig. 4.9 **The distribution of angles of displacement shows that slow p53 molecules diffuse anisotropically, but not fast p53 molecules. A:** Method: The angles traced between consecutive displacements of a molecule are recorded into a histogram. The metric called fold-anisotropy (or $_{180/0}$) is the ratio between forward jumps and backwards jumps. **B:** Distribution of angles between consecutive jumps for bound, slow and fast molecules of p53-HaloTag.

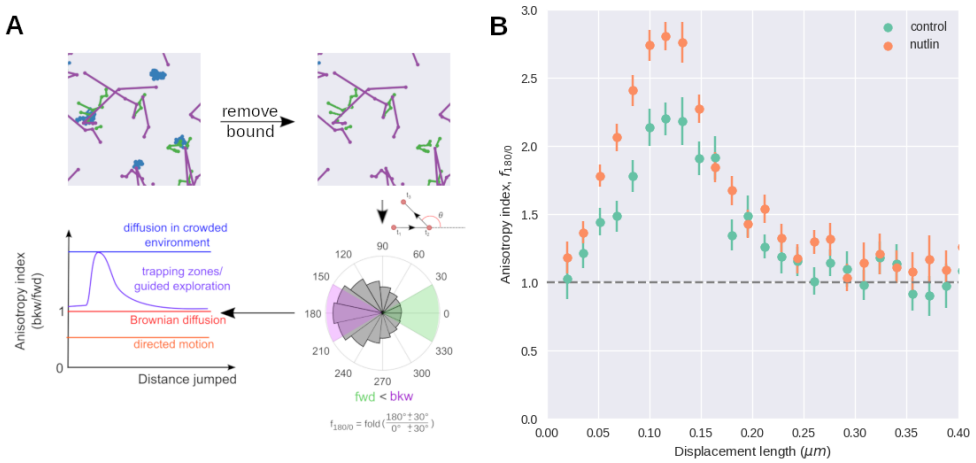

Fig. 4.10 **The anisotropy of slow p53 molecules is distance-dependent, a sign of guided diffusion. A:** Method: Bound molecules are filtered out, then the fold-anisotropy is calculated as before. Cartoon shows the expected curves for different type of diffusion. **B:** Fold-anisotropy curves with respect to the displacement length, for control and nutlin treated cells.

the case of the presence of a condensate acting as a 'trapping zone', this peak corresponds to the typical size of such trapping zones. After removing the bound molecules from the analysis, we perform such an analysis for p53 in fibroblasts shows this characteristic peak, with anisotropy increasing to a maximum at around 130nm before tapering off (fig.4.10.B), possibly a sign of local confinement. Interestingly, this behavior is separable between the sub-population: only the slow sub-population shows a peak in anisotropy, while fast p53 is isotropic at all length scales.

From this analysis we see that the peak in anisotropy increases upon treatment with after nutlin treatment, while the length at which it peaks remains the same. This results contrast with results obtained for p53 in cancer cells, in which the maximum anisotropy $F_{180/0}$ reached was around 1.5 after irradiation, when we measure a $F_{180/0} > 2$ even in

untreated condition, rising to almost 3 after nutlin treatment. The same caution we applied to the higher bound fraction could be applied here, p53 levels are higher in basal conditions, but the difference persists in activated conditions when p53 levels in the cancer cell line were high as well.

In conclusion, the diffusing population of p53 can be separated into two different sub-population, a slow and a fast diffusing one, which exhibit different dynamical behavior. With an analysis of the angle distributions in consecutive displacements, we showed that the slow diffusing p53 diffuse anisotropically, and we a correlation between this anisotropy and the displacement form frame to frame, such that the slow p53 molecules which move 130nm are almost three time more likely to go back on their tracks. This backward motion is compatible with local confinement, a type of guided exploration which help to periodically reduce the dimensionality of p53's search for its targets.

These results are in line with the reports of (Hansen et al., 2020) for the case of CTCF, and is not shared with all transcription factors – for example p65, a sub-unit of $NF-\kappa B$, shows no sign of guided exploration. Contrarily to CTCF however, p53 does not form condensates, even if it is found enriched in some nuclear bodies, and thus the proposed explanation for this confinement does not seem to hold. Given p53's interaction with chromatin both specifically and non-specifically (see chapter 2.5), we speculate that chromatin domains may be a driving force behind the observed anisotropic diffusion and instances of local confinement.

Therefore we move on to probing this interaction by combining SMT with a super-resolved map of chromatin to reveal the spatial correlation between p53 dynamics and chromatin architecture.

4.4 Probing the interplay between p53 and chromatin architecture.

Some previous work have looked at the interplay of DNA-binding proteins and chromatin (Miron et al., 2020) , but most looked at fixed cells to take advantage to super-resolution techniques to get detailed images, but depriving them of the dynamical aspect of this relationship, and making them subject to fixation artifact . On the other hand, some works used SMT techniques (Teves et al., 2016), but here the lack of resolution limits the ability to precisely measure this relationship. The microscope built in our lab solve this problem by allowing the use of multifocal Structured Illumination Microscopy (SIM), a super-resolution technique.

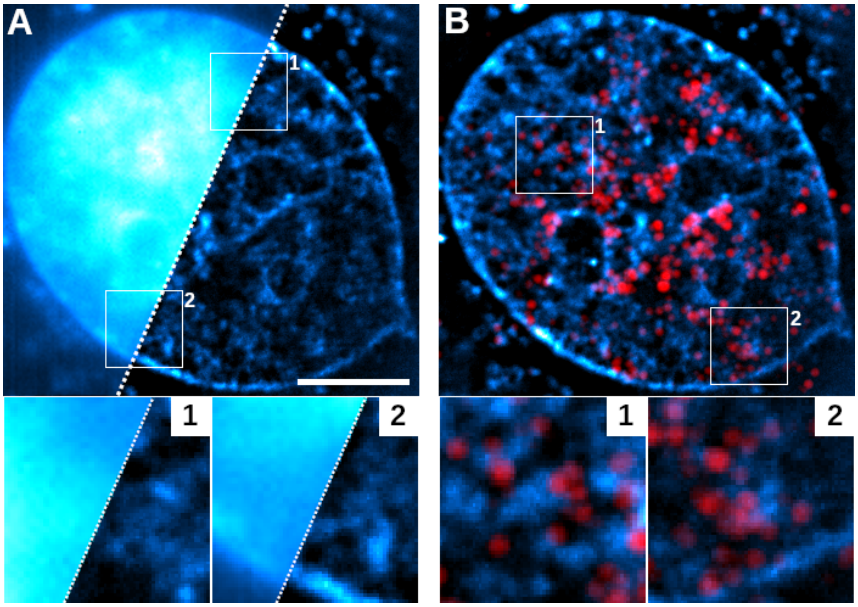

Fig. 4.11 **SIM improves the resolution of the reference image of chromatin dramatically. A:** Comparison between wide-field and SIM image of chromatin stained with Hoechst in the same nucleus. **B:** Maximum projection of p53-HaloTag imaged with SMT overlayed on the super-resolved chromatin. Details show that the resolution obtained allows us to see how p53 interacts to chromatin domains. Scale bar: 5µm

SIM uses a controllable array of mirror to scan a pattern of separable illuminated spots on the sample (see methods). Deconvolution of the spots provides a resolution improvement of $\sqrt{2}$, and the pinhole effect provides optical sectioning and an additional gain in resolution of $\sqrt{2}$, adding up to a theoretical gain of 2. The super-resolved images of chromatin structure are taken at 405nm, after labeling with Hoechst, a DNA-binder agent compatible with live-cell imaging, and shown in fig4.11.

The comparison between widefield and SIM demonstrates not only a much more detailed image, but also illustrate how much improvement we gain from from the optical sectioning (fig4.11), reducing the out-of-focus fluorescence dramatically. The resolution and contrast becomes sufficient to observe the sponge-like pattern described for the interphase nucleus by others using electron microscopy (Ou et al., 2017; Miron et al., 2020) , with chromatin domains separated with branching channels of chromatin poor regions in between. Although the typical size of the chromatin domains of 300nm in diameter is larger than the figures reported in those paper, a difference potentially attributable to our weaker resolution, the reference image is sufficiently fine-grained for us to correlate with p53 movement *in situ* (fig4.11.B).

To do so, we take a reference image of the chromatin (which takes 20s) before and after each SMT acquisition, which also lasts 20s. The references are visually inspected to check if they have changed too much in the meantime, the cell is excluded from the analysis, although this is rarely the case at these timescales (see Video 1). After segmentation

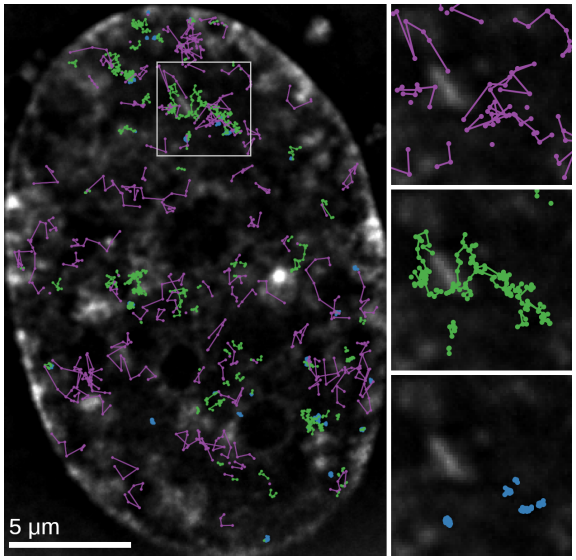

Fig. 4.12 **vbSPT classified p53-molecules seem to interact differently with chromatin.** Example of a SIM/SMT acquisition, with classified trajectories overlaid on the chromatin. Details on the right separates the trajectories according to their classes. Purple: Fast, Green: Slow, Bound: Blue.

with vbSPT, we can first plot the movement of the p53 molecules belonging to each sub-population in relationship with chromatin (fig4.12).

To do so, we take inspiration from the analysis presented in (Miron et al., 2020). It consist in segmenting the nucleus into 7 classes depending on the chromatin density – as measured by the intensity of the Hoechst signal – and record how many molecules of p53 are present in each classes. However, in addition to the localization of molecules, given that we have access to the speed of molecules, we can also separate the molecules of p53 according to their speed. As seen in fig.4.13, the dynamic nature of our acquisitions means we can look at the statistical distribution of velocities. For example, we can ask if the slow p53 are enriched in different compartments than fast moving molecules of p53.

This allows us to reveal the general preferences of p53 for the different domains of chromatin density (fig.4.14). First, it seems that in general, fast and slow molecules of p53 preferentially occupy regions of different chromatin densities, with fast molecules being mostly excluded from the regions of highest chromatin density, while slow molecules are enriched there. Second, the preferential enrichment changes upon treatment, with a striking enrichment of p53 molecules in class 3 (perichromatin) upon nutlin. In both conditions, it seems like only slow moving molecules are able to penetrate in chromatin dense regions and bind there. Interestingly, p53 is seemingly depleted from the interchromatin compartment, although the fast moving molecules in untreated cells are found there more frequently.

Here we see that considering the dynamical aspect of the interplay between p53 and chromatin reveals a non-trivial relationship. In which slow molecules penetrate deeper into dense chromatin domain, with the majority of binding happening inside denser chromatin regions. These slow molecules exhibit anisotropic diffusion, contrary to the sub-

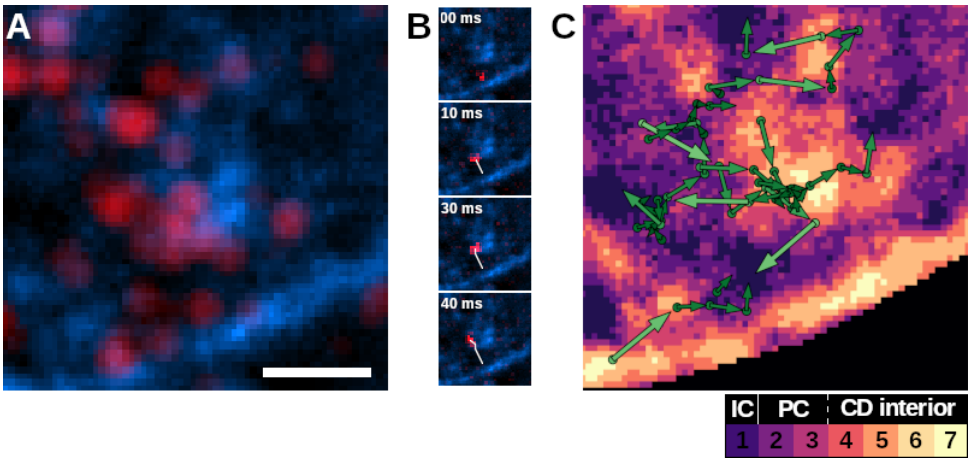

Fig. 4.13 **Combining SIM & SMT in live-cells reveals the dynamic relationship between p53 and chromatin. A:** Detail of an SIM/SMT acquisition with a maximum projection of p53-HaloTag (red) and chromatin (blue) **B:** Sequential SMT images show the movement of one p53 molecule in context. **C:** Result of the segmentation of chromatin into 7 classes according to their chromatin density. Both the position and velocity vectors of recorded p53-HaloTag molecules are shown in context with green points and arrows respectively. The chromatin classes are characteristic of three compartments. IC: Interchromatin, PC: Perichromatin, CD: Chromatin Domain. Scale bar: 1µm

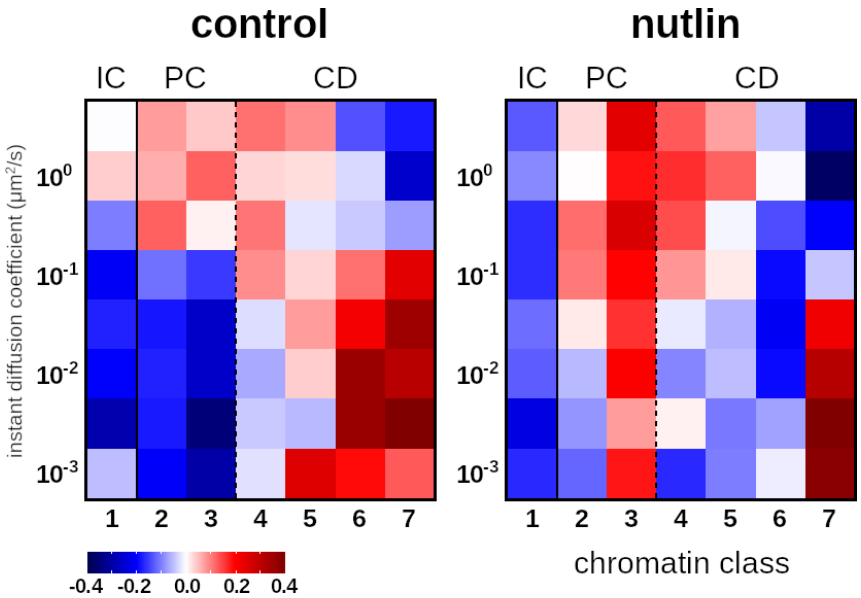

Fig. 4.14 **Dynamical enrichment of p53 in chromatin compartment of different density shows a different behavior of slow p53 with respect to chromatin.** The plot shows the enrichment of molecules in different classes of chromatin density (horizontal axis) in relationship to their instant diffusion coefficient (vertical axis), compared to a uniform repartition inside the nucleus. The data is normalized row-wise. **Reading:** p53-Halotag molecules diffusing at $10^{-2}\mu m^2/s$ are found 40% more than expected by chance in the 6^{th} highest chromatin density regions. IC: Interchromatin, PC: PeriChromatin, CD: Chromatin Domain.

population of fast moving p53 molecules which are mostly diffusing in the interchromatin compartment.

Striking as it is, this data remains mostly qualitative. In order to better understand this behavior, we sought to model it using physics-based models of diffusion, such that we could hopefully reproduce the behavior that we observed, such as the volume exclusion on one hand, and the penetration of slower molecules inside denser chromatin regions before binding there. Such a model would also allow us to quantify important parameters of the relationship between p53 and chromatin.

Chapter 5

Modeling the volume exclusion exerted by chromatin onto transcription factors.

The diffusion of molecule in a fractal media is a well-known phenomenon which can produce anomalous diffusion and speed up the target search. Given that we observe fast diffusing p53 to be excluded from the chain of dense chromatin regions, it is logical to ask if this is responsible for the anisotropic diffusion we observed, and if this could help the target search mechanism. That is to say, could volume exclusion lead to decreased time for p53 to reach its targets? This effect of volume exclusion here would be different from the process of facilitated diffusion where p53 slides along the chromatin fiber. To answer this question in the case of p53 or other molecules, it is however necessary to measure the degree to which those transcription factors are excluded from chromatin dense region, and what target search speed increase we could expect from such value.

Beyond this, modeling can bring deeper understanding of complex phenomenon. As Gunawardena (2014) puts it in his review, models are logical machines for converting assumptions into conclusions. By grounding our model in physical assumptions about the diffusion of molecules in a medium, and seeing those succeed or fail to reproduce the data, we learn what physical interactions drive the behavior of the system. Not only does that further our understanding, it also proves extremely valuable for comparing the behavior of different systems be they, in our case, different transcription factors, or different spatial organization of chromatin dense and chromatin poor domains.

In this chapter we make use of the work of (S. A. Isaacson et al., 2011), which models the volume exclusion from chromatin dense regions as molecules diffusing in a repulsive potential. While this model in itself predicts an increase in the speed of the target search, we show that the diffusion of molecules simulated in this system behave differently from the actual molecules movement recorded by SMT. We then extend the model by taking into account short-lived DNA interactions, which we show is necessary to explain how p53 and other transcription factors penetrate inside dense chromatin domains.

5.1 Mathematical description and stochastic simulation of diffusion in a volume exclusion potential.

Following the work of Samuel A. Isaacson et al. (2013), we modeled the diffusion of molecules in the nucleus as a diffusion inside a repulsive potential whose strength is proportional to the local chromatin density. This model approximates the effect of volume exclusion by making it more difficult for molecules to diffuse towards regions of higher chromatin density. An advantage of this model is the single constant determining the strength of the volume exclusion, which could be used to easily compare the influence of chromatin on different type of molecules beyond the single case of p53.

We model diffusion in a bounded domain Ω, with a potential U(x) whose value is directly proportional to the chromatin density at x. With a protein diffusing with a diffusion coefficient D, k_B the Boltzmann constant and T the temperature in Kelvin, the evolution of the system can described with the Fokker-Planck equation.

$$\frac{\partial}{\partial t}(x,t) = D\nabla \cdot \left(\nabla p(x,t) + \frac{1}{k_B T} p(x,t) \nabla U(x) \right) \tag{5.1}$$

This system has a Neumann boundary condition, which imposes that molecules cannot diffuse through the membrane. This equation can be used to derive the fluxes of molecules diffusing from any point x to the neighboring region. Given the nature of our data, chromatin images in the form of an array of pixels, we need to discretize the Fokker-Plank equation to obtain a discrete diffusion master equation. The resulting form of the master equation becomes (S. A. Isaacson et al., 2011) :

$$\frac{dP}{dt}(\overrightarrow{i},t) = \sum_j [\alpha_{i \leftarrow j} P(\overrightarrow{j},t) - \alpha_{j \leftarrow i} P(\overrightarrow{i},t)] \tag{5.2}$$

which in plain english is easily read as a relationship saying the change in density of molecules at a given pixel i is equal to the amount of molecules that diffused inside the pixel during the amount of time dt, minus the amount of molecules that diffused out. Importantly, this amount of molecule diffusing to neighboring pixels depends on the value of the difference in the value potential between those two pixels, and hence depends on the difference in chromatin density. Thus this difference in intensity between neighboring pixels i and j determines the rate of diffusion $\alpha_{i \leftarrow j}$ which has the form:

$$\alpha_{i \leftarrow j} = \frac{2D}{h_d^2} \frac{1}{1 + exp((U(x_i) - (U(x_j)/k_b T)} \tag{5.3}$$

where h_d is the pixel size. Here we depart slightly from the naming convention in S. A. Isaacson et al. (2011). Since the value of the potential is proportional to the chromatin

density at any pixel, we can rewrite $U(x_i)$ as $\overline{U}I_i$, with I the intensity of the chromatin signal at the pixel i, and $\overline{U}$ a scaling constant. This gives us the simplified equation :

$$\alpha_{i\leftarrow j} = \frac{2D}{h_d^2}\frac{1}{1+exp(R\cdot\Delta I)} \tag{5.4}$$

with the factor $R = \overline{U}/k_bT$, which I term the 'repulsiveness' of chromatin. Given that we can measure the diffusion coefficient D, the repulsiveness is the only free parameter in the system. It determines the strength of the repulsion felt by diffusing molecules: higher values of R make it harder for molecules to diffuse in regions of higher chromatin density. This system can be studied by solving numerically the systems of ordinary differential equations derived from the master equation. However, to study the dynamical features that a molecule diffusing in such a potential would exhibit, we can simulate the stochastic process by using the Gillespie method. This models diffusion as a continuous time random walk, which we can simulate directly on the lattice provided by the chromatin maps that we acquire in SMT experiments. Briefly, the simulation according to the Gillespie method works as follows :

1. we pre-compute all the jump rates $\alpha_{i\leftarrow j}$.
2. we initialize the molecule at a random pixel within the nucleus with probabilities according to the steady-state distribution.
3. for each time, we draw pseudo-times dTj for diffusion to each neighboring pixel j, from exponential distributions such that $dT_j \sim exp(\frac{1}{\alpha_{i,j}})$. The pixel j to which the molecule diffuses is chosen as $\min_j(dT_j)$, the direction with the smallest pseudo-time.
4. we update the position of the molecule to pixel j, and the time to $T + dT_j$.
5. Repeat from 3 until the desired time as elapsed.

This method is shown to be an exact simulation of stochastic processes (Gillespie, 1977). The simulation is performed on the chromatin map reconstructed with SIM, coming from 5 different nuclei. The images are up-scaled by a factor 2 with cubic interpolation, such that the distance h between two pixel centers is halved, from 54nm to 27nm. This alleviates artifacts due to the discrete nature of our images. We note that this diffusion process takes place in the 2 dimensions of our microscopy images.

As a first way to apprehend the behavior of molecules diffusing in such a system, I simulated a thousand molecules diffusing for 5 seconds each at increasing values of the repulsiveness R. In order to better compare with experimental data, I convert the continuous time simulation by recording the positions of molecules at 10ms intervals, the same interval used for SMT. In order to avoid biases, the starting point of each

trajectory is chosen randomly from the distribution given by the steady-state. The resulting trajectories are analyzed for their anisotropy, in a similar way than described previously. Interestingly, the simulated trajectory are more anisotropic as the repulsiveness coefficient R increases (figure 5.1), with the fold-anisotropy ratio increasing to 1.28 for a value of 20 for the repulsiveness. This result is expected: with stronger volume exclusion, the system resembles more and more the case of diffusion in a fractal medium, which is known to give rise to anisotropic diffusion.

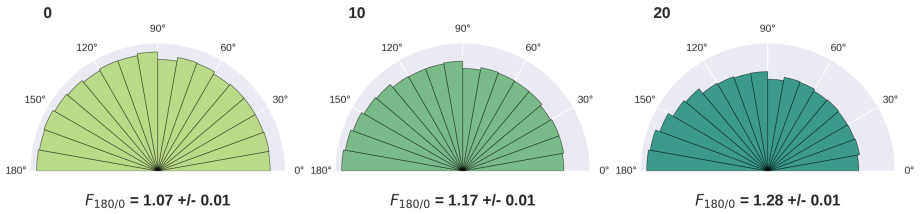

Fig. 5.1 **Diffusion in a volume exclusion potential produces anisotropy.** Distribution of the angles between two consecutive frames, for increasing values of repulsiveness, shows an increasing fraction of backward displacements. From left to right, R is 0, 10 and 20. Fold-anisotropy metric is reported below the histograms.

It has been noted by other teams that this could explain the difference observed in the diffusing behavior of molecules, as different proteins could by excluded from chromatin with different strength (Woringer and Darzacq, 2018). However, as we see from figure 5.1, for volume exclusion to be a significant factor, the value of the repulsiveness R must be significantly high. Similarly, (Samuel A. Isaacson et al., 2013) found using similar Gillespie simulations that there is a significant decreases in the search time for values of R approaching 20. However their work was limited by the fact that they possessed no experimental data of the trajectory nor of the positions of transcription factors, and thus could not conclude if real transcription factors made use of volume exclusion as they predicted. Since we produced such data, we set out to compare the model to it.

5.2 Volume exclusion fails to reproduce the behavior of various transcription factors.

A good scientific model should give out more insights than the data we put in. While fitting a model to p53 behavior can be done, it is much more interesting to see how the volume exclusion mechanism factors in the behavior of a range of different molecules in general, and transcription in particular. To do so, I made use of the results of SMT experiments performed by Matteo Mazzocca in our lab on 3 different transcription factors, p53, RelA (a sub-unit of the NF-κ B factor) and CTCF, as well as the histone H2B and HaloTag on its own, we can provide a reference for a non DNA-interacting molecule. I reanalyzed the data obtained in those experiments to fit a model of the influence of chromatin on the movement of those factors.

The SMT acquisitions were performed in a similar manner as previously described, with the two main differences being the cellular model and the use of transfection – of note, here p53-HaloTag is thus diffusing in cancer cells. The cells used in these experiments were graciously gifted to us by the Legube lab (DIvA cells used for p53, p65, H2B and HaloTag), and the Tijan and Darzacq group (U2OS-derived cells used for CTCF). All cells were transiently transfected with plasmids expression HaloTag-conjugated version of the proteins 24h before the experiments, barring the U2OS cells which stably express CTCF 8.

In a first exploratory approach, we can analyze the repartition of all these different factors with relationship to chromatin in a manner similar to fig 4.12. The reference chromatin image is segmented into 7 different classes according to the chromatin density, and we record where the different molecules fall with respect to those (fig 5.2.A). As expected, the plotted enrichment reveals the histone H2B is enriched in chromatin dense regions, meaning it is found there more than expected by chance (fig 5.2.B). This is of course unsurprising, as H2B co-localizes with chromatin for obvious reasons, to the point of being used as a marker for chromatin. More interesting are the behavior of the other three DNA-binding molecules. All show some degree of volume exclusion from chromatin dense regions, with interesting features however. Like HaloTag, RelA seems to be more simply excluded from chromatin dense regions, whereas p53 is depleted in the inter-chromatin compartment and, while excluded from the densest regions, is enriched more specifically in regions of intermediate densities. CTCF seems to follow a similar trend than p53, except for a specific enrichment in the 6th most dense region.

The behavior of CTCF is especially interesting because its repartition in relationship to chromatin has been measured before in Miron et al. (2020), which does not report this specific enrichment. One possible explanation for this departure is an artifact coming from the fixation in that study. Formaldehyde fixation, which the researchers use for this study, has been shown to disrupt the interaction between TFs and chromatin and produce misleading results especially for factors with shorter binding times (Teves et al., 2016; Raccaud and Suter, 2018), and this might explain why we see this population of bound CTCF in live-cell imaging which doesn't appear in fixed cells.

Similarly to the previous chapter, we can make use of the dynamical aspect of our data to separate this enrichment in different chromatin regions with respect to the instant diffusion coefficient of the molecule at the time.

We see that, with respect to this dynamical enrichment, p53 behaves similarly in this cancer cell model than in the fibroblast presented in the previous chapter. The dynamical separation also highlights the difference in behavior between p53 and RelA in a striking manner. Whatever the diffusion coefficient, RelA molecules stay outside of the core of chromatin domains, while both p53 and CTCF penetrate as slower moving molecules. But there is also a striking difference between CTCF and p53 in that all very slow molecules of CTCF are inside chromatin bound regions.

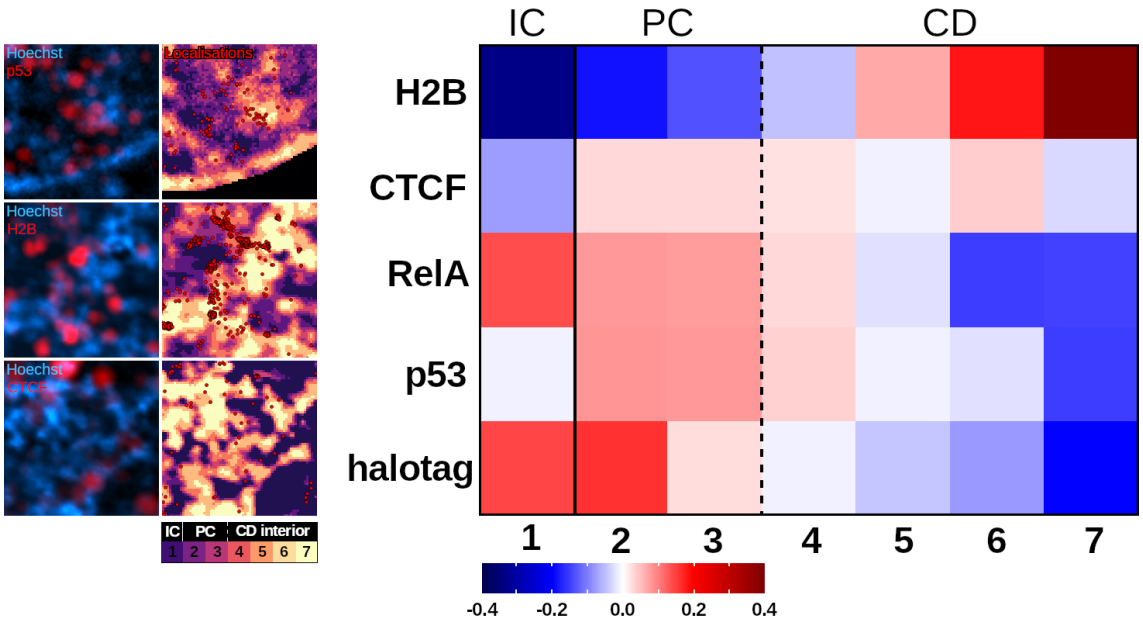

Fig. 5.2 **5 different nuclear factors distribute differently with respect to chromatin density. Left:** Representative images of molecules in the nuclei of cells. The left image shows a maximum projection of the SMT acquisition (red) in the context of chromatin (blue). The corresponding right image shows the segmentation of the chromatin into 7 classes of chromatin density, with the localized molecules overlaid. **Right:** Enrichment of the factors with respect to chromatin density, compared to a uniform repartition inside the nucleus. The data is normalized row-wise. **Reading:** p53-RelA molecules are found 15% more than expected by chance in the least dense chromatin region. IC: Interchromatin, PC: PeriChromatin, CD: Chromatin Domain. Number of cells: N_{H2B}32,N_{CTCF}=31,N_{RelA}=32,N_{p53}=29,$N_{HaloTag}$=31.

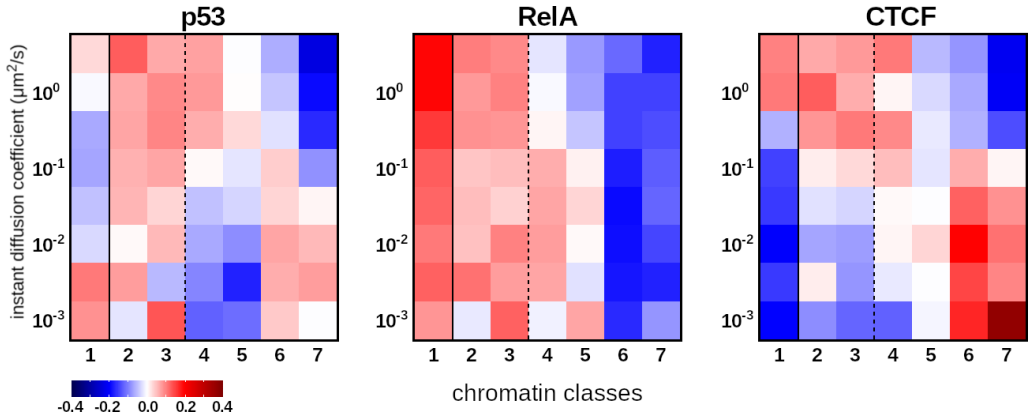

Fig. 5.3 **The dynamic enrichment of p53, RelA and CTCF distinguishes between different behaviors.**

We can now make use the volume exclusion model described in the previous section. Our objective is two fold: (1) use the experimental data to fit the free parameter of the model and measure the strength of the repulsive potential felt by different TFs and (2) see if volume exclusion manages to capture the dynamical behavior of these molecules.

One way to fit the model is to use the steady-state distribution that it predicts. The form of the equation 5.3 was chosen so that it is the Boltzmann distribution. This ensures that the steady-state is :

$$P_i^{SS} = \frac{e^{-RI_i}}{\sum_j e^{-RI_j}} \tag{5.5}$$

From this equation, we can trivially derive a log-likelihood (see annex) of seeing the observed distribution of molecules given a certain value of R. We consider the repartition of molecules seen in the SMT to reflect the steady-state distribution, and find the value of R which maximizes the log-likelihood. We present the results in figure 5.4. Additionally, following the same method than for fig 5.1, I simulated 500 thousand trajectories of 5 seconds each for each molecule considered, and used the simulated data to plot the enrichment in different chromatin compartment that we would expect using the fitted values for R.

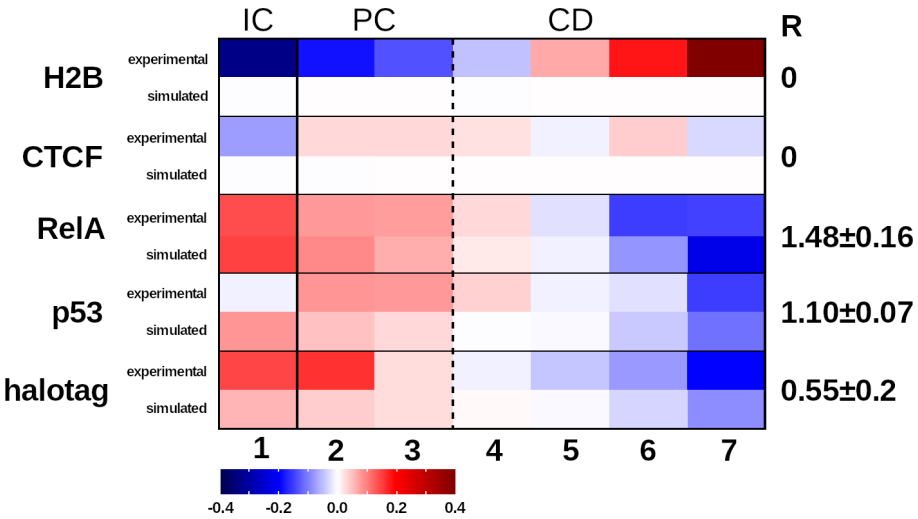

Fig. 5.4 **The experimental data suggests very low values of repulsiveness and fits the data poorly.** Comparison between the total enrichment presented in fig.5.3 (top half of each row, labeled 'experimental') and the predicted total enrichment based on the fitted R value (bottom half of each row, labeled 'simulated'. Fitted R values are reported for each factor on the right, with 95% confidence intervals.

The resulting value of the repulsiveness R are surprisingly low. Unsurprisingly, the value for H2B is 0, an expected value for a factor which co-localizes with chromatin. As a result, the expected enrichment is flat, indicative of a molecule which is not affected by chromatin density and diffuses freely. For HaloTag alone and RelA, volume exclusion results in an enrichment with respect to chromatin which is similar to the experimental one, although in this case also R is low. Finally, for CTCF and p53, the expected enrichment is clearly discordant from the experiment.

The explanation for this appears clearly when we plot the expected dynamical enrichment of simulated molecules for different values of R. As we can see in FIG 5.5, increases in the repulsiveness result in an exclusion from chromatin dense compartments regardless of the velocity of the molecules considered.

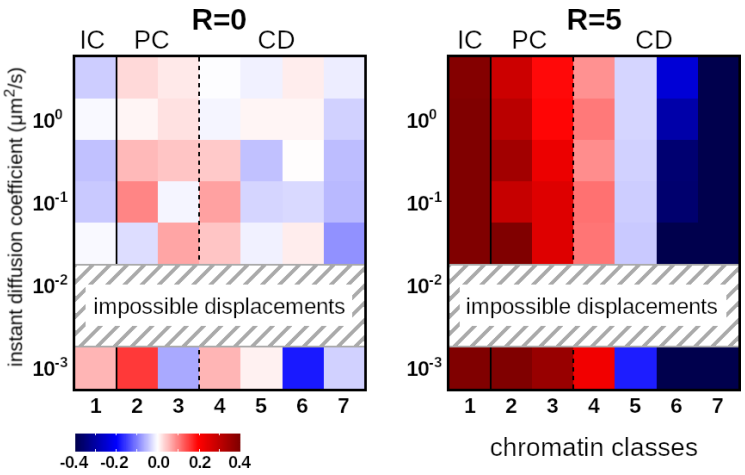

Fig. 5.5 T**he predicted dynamical enrichment of molecules diffusing inside a volume exclusion potential does not show a dependence on local chromatin density.** Note: due to the discreetization, the minimal displacement length is $0.027\mu m$, hence some values of the instant diffusion coefficient are impossible to measure in the simulated data.

This is a logical result from pure volume exclusion, whatever the speed molecules react to the same potential. When trying to fit the purely repulsive model to the real world data of p53 and CTCF, the only conclusion from the model which can explain the presence of molecules in chromatin dense regions is that they are not excluded at all. We can compare this kind of flat dynamical enrichment to the observed ones presented in 5.3, with p53 and CTCF showing that slow or bound molecules (those moving with instant diffusion coefficient below $10^{-2}\mu m$) are enrichment in dense chromatin compartment. In contrast, we see in this figure that the RelA sub-unit fits much better this idea of a simple volume exclusion, with a 'flat' enrichment or RelA in chromatin poor regions (equivalently, exclusion from chromatin dense regions), regardless of its speed.

Thus we showed that the volume exclusion model developed by S. A. Isaacson et al. (2011) fails to account for key features of the dynamical behavior of real transcription factors. It can fit to the behavior of some molecules, such as RelA and the non DNA-interacting HaloTag, but even for these the fitted value of R falls well short of the values predicted by S. A. Isaacson et al. (2011) to significantly reduce the search time (which hovered around 20). Moreover, there exists clearly an important category of transcription factors which are subject to forces beyond volume exclusion, allowing slow moving molecules to penetrate inside dense chromatin domains.

5.3 Non-specific DNA interactions are necessary to reproduce TFs dynamics in relationship to chromatin.

One obvious ways in which DNA-interacting molecules differ from simple particles diffusing in a repulsive potential is their ability to bind DNA. Although the bind it specifically for longer periods of times, most TFs, among which p53, are known to interact with DNA non-specifically in a variety of ways. This formed the basis of an extension for the model proposed in Isaacson, 2011, to take into account short-lived binding interaction with the chromatin scaffold.

For the extended model, we consider two states that the diffusing molecule can have, bound or unbound. We model the unbound/diffusing state in the same manner as before when it comes to diffusion, however an additional event is available to it, namely it can bind with a rate of k_i^{on}. We let kon be proportional to the local value of chromatin density at a given pixel i, reasoning that the higher amount of DNA in the volume translates in more binding sites available to encounter. On the other end, once bound, the particle detaches from the scaffold with a constant rate k^{off}. k^{off} is chosen small enough that the typical binding time is below the frame time of 10ms. This way, this non-specific binding models non-specific interactions, which aren't captured in our experiment at 10ms of frame time, but would translate in an overall slow-down of the diffusion of the molecule, possibly creating this sub-population of slow-diffusing molecules that we have observed in section 4.3. Denoting the two states unbound and bound with the exponent u and b respectively, the addition of this new state results in the following master equation :

$$\begin{cases} \frac{dP_i^u}{dt} &= \sum_j \left[\alpha_{i \leftarrow j} P_j^u - \alpha_{j \leftarrow i} P_i^u \right] - k_i^{on} P_i^u + k^{off} P_i^b \\ \frac{dP_i^b}{dt} &= k_i^{on} P_i^u - k^{off} P_i^b \end{cases} \tag{5.6}$$

The derivation of the steady-state for this new system is significantly more involved than the previous one, and I thank Nacho Molina for his help on it. The full derivation is in the annex. In the end, the steady-state probability of being at pixel i ends up as:

$$P_i^{SS} = \left((1 - F_b) + F_b * \frac{I_i}{\langle S \rangle} \right) \cdot \frac{e^{-RI_i}}{\sum_j e^{-RI_j}}, with \langle S \rangle \equiv \frac{\sum_i I_i e^{-RI_i}}{\sum_i e^{-RI_i}} \tag{5.7}$$

In which F_b is the fraction of molecules engaged in the bound state. Of note, however, this bound fraction is not the same as the bound fraction generally used in the literature, which measures the fraction of molecules bound specifically. Here, it represents short-lived non-specific interactions, and is unobservable directly by microscopy. Instead, it becomes a second free parameter. This steady-state distribution can be used to fit the model to the observations, using the same method of log-likelihood maximization as described previously.

However, there is a little difficulty to consider before running Gillespie simulation of the molecules with the fitted parameters R and F_b. In the framework of this model, chromatin has a retarding effect on diffusion due to the short-lived non-specific interactions. We thus need to consider that we can't use the observed diffusion coefficient measured for each molecules in SMT, as these would incorporate such retarding effects. We thus found a heuristic to estimate this 'ideal' diffusion coefficient. We start from the diffusion coefficient of the presumably non DNA-interaction molecule HaloTag, as measured from SMT experiments with the histogram of displacement. Because the diffusion coefficient is inversely proportional to the cube of the radius of the particle, we can use the difference in size between HaloTag and the molecule, for example p53, to estimate the expected diffusion coefficient of a non DNA-interacting molecules the size of p53. This expected diffusion coefficient can be used for the Gillespie simulation, and compared to the observed diffusion coefficient, which gives an indication for the value of Fb we should expect. To estimate the radius of each molecules, we use their size in amino-acids and the scaling law which Hong and Lei (2009) estimated using the Protein Data Bank, where the radius scales with $R_n(N) = 2.24 * N^{0.392}$. Of note, the fraction bound in our model depends on the ratio between the two parameters k^{off} and k^{on}. Both on a theoretical and pratical level, the exact value of these parameters do not impact the simulation, and indeed the simulated trajectories are similar when we half or double the values of those parameters. The absolute value of k^{off} only starts having an impact when it results in binding times of the order of the duration of the simulation.

The value of the fitted parameters R and F_b are presented in fig.5.6. Of note, in this model, the values of the repulsiveness is significantly higher than in the previous model. This is explained by the fact that, when only non-specific interactions are considered, the molecules in this system are over-represented in regions of high DNA density because they are more like to bind there. Thus this over-representation needs to be balanced by a higher degree of repulsiveness to counter-act it, although the values are still much lower than the values predicted by S. A. Isaacson et al. (2011) to be significantly beneficial for the search time. While the bound fraction may seem quite high, it is important to remember that this should represent non-specific interactions. Figures of a similar magnitude have been reported, for example in bacteria, TFs have been speculated to spend 90% of their time bound transiently at non-specific sites (Elf et al., 2007).

I can then use these parameters to simulate trajectories for all of these molecules inside chromatin, according to the same procedure as before. From the resulting trajectories, we plot the dynamical enrichment plots presented before. Immediately, we see that in contrast to the model taking only repulsiveness into account, this model correctly reproduces an enrichment of slow molecules inside dense chromatin domains. I quantify the improvement added by taking binding into account by taking the ratio of the residual sum of squares compared to experimental data with and without binding (fig.5.7), and we see, as expected

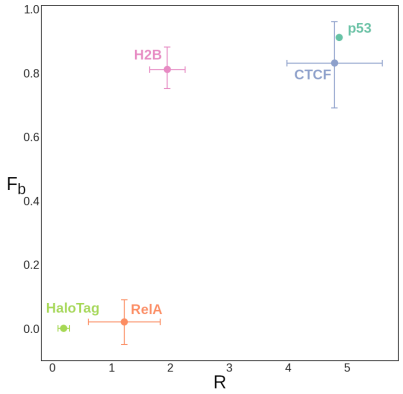

Fig. 5.6 **Resulting parameters from fitting the model to observation for each molecule. Error bar are the 95% confidence intervals from resampling the data for multiple fitting (see Methods).**

from the previous section, that for RelA and the inert tracer HaloTag, volume exclusion alone explains the data. But, varying the values of the parameters, we can reproduce either the enrichment of H2B inside chromatin dense regions by increasing the binding component, and the exclusion of the RelA sub-unit by letting the repulsiveness dominate. In this model, the two forces act in a sort of tug-of-war to explain the repartition of molecules in relationship to chromatin. It can also, in the case of CTCF, explain strong shift between slow/bound CTCF inside chromatin dense regions and the faster moving population in the chromatin poor regions. However, it struggles to capture the enrichment of p53 in the perichromatin region. One potential explanation for this discrepancy is the existence of two distinct populations of p53 molecules. These could behave differently due to various factors, possibly influenced by specific co-factors. As a preliminary solution, one could consider adjusting the model to simulate both populations separately — using different sets of parameters — and subsequently combining the results. However, further experiment are needed to motivate this combination, such as measuring the proportion of p53 molecules with specific PTM, which could be a clue for their different state and potential binding partners. Importantly, this model also fails to reproduce the anisotropic behavior of slow molecules (fig 5.7), which isn't surprising as the bound state has no effect on diffusion beyond keeping the molecule in place longer.

Despite these limitations which invite caution in the explanatory power of the model, we consider it to say something valuable. In their study, Miron et al. (2020) observed a zonation of chromatin marks around chromatin domains. To explain this zonation, they relied on a different kind of physics based model developed by Maeshima, Kaizu, et al. (2015), which looked at the effect of a molecule's size on its capacity to penetrate chromatin dense regions. Thus they reason, the zonation simply results from the volume exclusion of dense chromatin, which results in the inability of some larger complexes to diffuse deeper inside the domains, and restricts their activity to the outer surface of chromatin domains. Because we have access to data on the repartition and dynamics of a certain kind

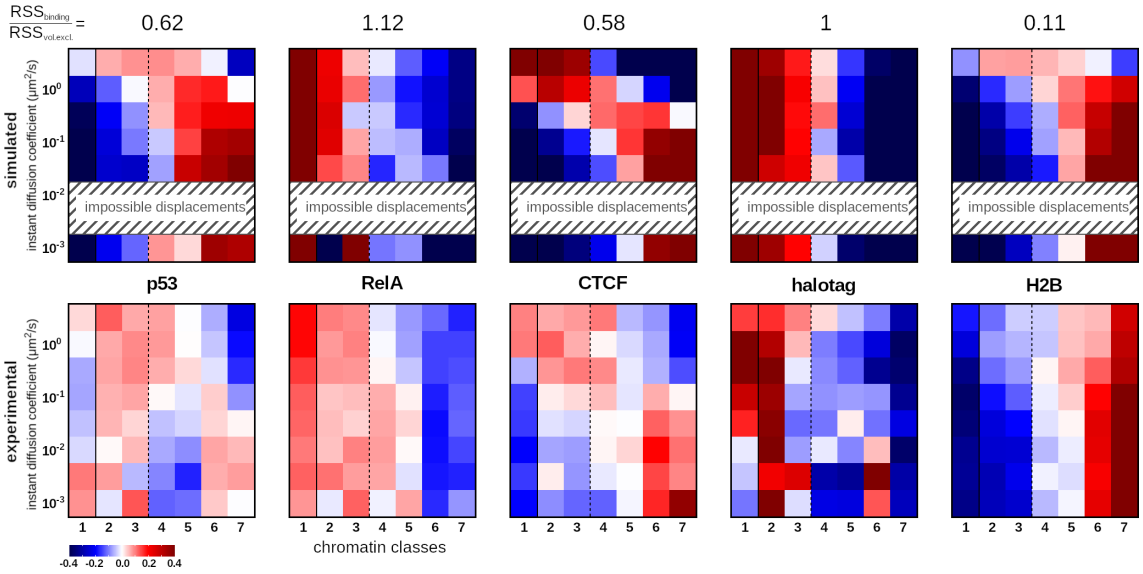

Fig. 5.7 **Accounting for binding recapitulates the experimentally observed dynamical enrichment of DNA-interacting molecules. Top row:** Dynamical enrichment of simulated molecules with the fitted parameters. **Bottom row:** Dynamical enrichment of the molecules experimentally observed with SIM/SMT. Above the figure is a quantification of how much better to the new simulated data fit the experiments, as measured by the ratio of the residual sum of squares between the data produced by pure repulsion, or the data produced by repulsion and non-specific binding.

of chromatin factor, namely transcription factor, we were able to test this mechanism in this context, and we provide evidence that volume exclusion is an imperfect explanation for their behavior. Furthermore, we provide evidence that their ability to interact and bind DNA can give these factors the ability to penetrate inside chromatin dense regions, which results in a dynamical behavior which matches more closely the experimental data.

Many caveats exist in our analysis, as well as in the base data this was derived from. Despite the improvement allowed by the use of SIM, the chromatin domains are at the edge of our resolution limit. This significantly hinders our ability to study the interplay between transcription factor and the surface of these domains specifically. While it was enough for use to look at the overall repartition of molecules in relationship with chromatin in a broader sense in this chapter, studying the interface between chromatin poor and chromatin dense regions is impossible to do with any degree of accuracy in this model. This is why we move on to cells induced into senescence, where the heterochromatin foci are much bigger and their surface easier to study.

Chapter 6

The interplay between p53 and Senescence-Associated Heterochromatin Foci.

In this chapter, we want to build on the foundations established by our previous exploration of p53 dynamics and their interplay with chromatin. While some works have tried to perturb this relationship by using mutant versions of p53 (Loffreda et al., 2017; Mazzocca, Loffreda, et al., 2022; F. Huang et al., 2009) , we make use of our fibroblasts cell-lines to change the chromatin context instead. With the onset of oncogene-induced senescence (OIS), the nuclear landscape changes dramatically and large Senescence-Associated Heterochromatin Foci (SAHF) appear.

The 4D nucleome community, especially the Cremer group (T. Cremer et al., 2020), has put forward the idea that the nucleus is divided in three main compartment ; the inactive compartment formed by chromatin dense regions, the branching interchromatin compartment essentially devoid of DNA and, at the interface between the two, the active compartment. This framework gives an important role to the interchromatin channels in which many molecules such as RNA would circulate, but also points out that what we typically call euchromatin forms a surface around chromatin dense compartments.

These foci provide us with larger-scaled features that are more easily identified and delineated than the chromatin domains of proliferating cells, which will allow us to study with less ambiguity the behavior of p53 – or any other molecules – at these interface between chromatin poor and chromatin dense regions in the nucleus. We can thus try to study the characteristics of p53 dynamical behavior around SAHF, and compare it its behavior in proliferating cells without SAHF.

6.1 Characterization of p53 dynamics in SAHF presenting cells.

In these experiments, we use the knock-in pool of BJ cells obtained thanks to the efforts of Daniela Gnani and described in section 4.1. As described there, it presented several issues, first among which a basal activation of p53 even in untreated cells. Further, the BJ cells sensitivity to anoikis makes the isolation of clones significantly harder, and we haven't yet succeeded in this task. This heterogeneity compounds on the heterogeneity of senescence in the first place (see section 3.2). Fortunately microscopy is uniquely placed to mitigate these problems, at least to an extent, given the inherently single-cell level of detail allowed by the method.

In this system, the induction of OIS is done with transduction of the oncogenic mutant of the BRAF gene, $BRAF^{V600E}$, along with a GFP reporter. After 10 days, >10% of the cells are GFP positive, but twice as many cells present SAHFs since, with paracrine senescence, senescent cells are able to induce senescence in their neighbors. Fig.6.1.A shows a representative image of a cell with SAHF, and demonstrate that, even with the more obvious foci, the addition of SIM in our microscopy allows for a much sharper image with less light coming from out-of-focus planes.

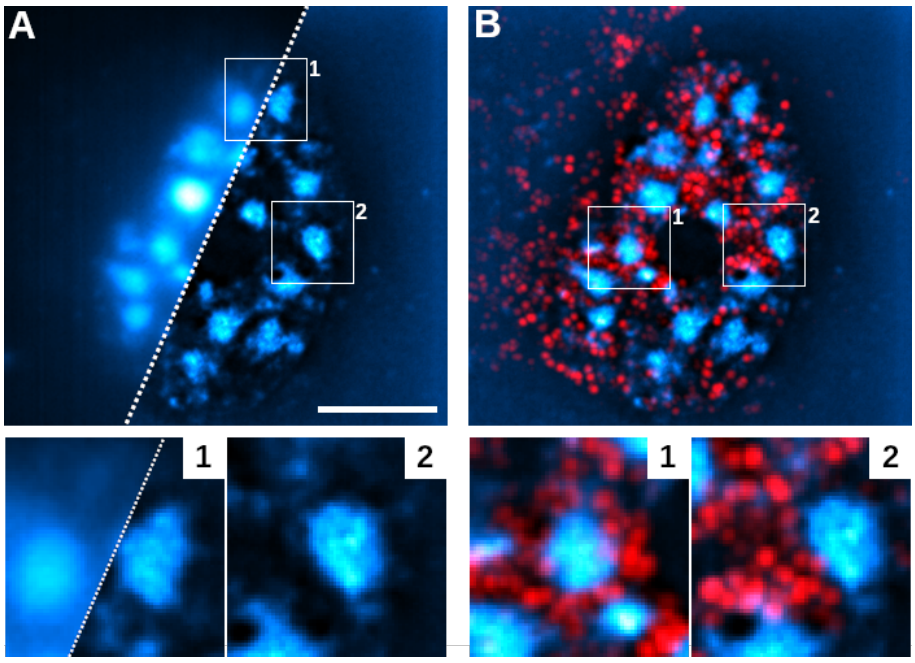

Fig. 6.1 **The dual use of SIM & SMT allows for much more precise reference image of chromatin. A:** Comparison between a wide-field image against the SIM image of the same nucleus. **B:** Maximum projection of the SMT acquisition showing p53-HaloTag (red) in the context of chromatin (blue). Scale bar: $5\mu m$

In each such cells, I acquire 40 seconds of SMT, with a time resolution of 10ms, while taking 3 references images of chromatin (at t=0, t=20 and t=40s). The GFP signal allows

us to distinguish between primary senescent cells, who received the oncogenic protein, and secondary senescent cells, who got induced into senescence by the former. While we intend to make use of this in further research, in the present thesis I focus on the distinction between SAHF-presenting cells compared to those without SAHF. As such, cells presented in the $BRAF^{V600E}$ group in subsequent figures all have SAHFs.

Upon visual inspection of the movies, the interaction between p53 molecules and the surface of SAHF is evident to our eyes. We can see that some p53 molecules appear to be exploring the surface of SAHF for short periods of time, and diffuse slower as they do so. This behavior might be indicative of p53 sliding on the surface of the SAHFs, which is interesting since it would constitute an instant of facilitated diffusion ; diffusion on a surface, in 2D, allows for the periodic reduction in the dimensionality of the search described in chapter 1.4.

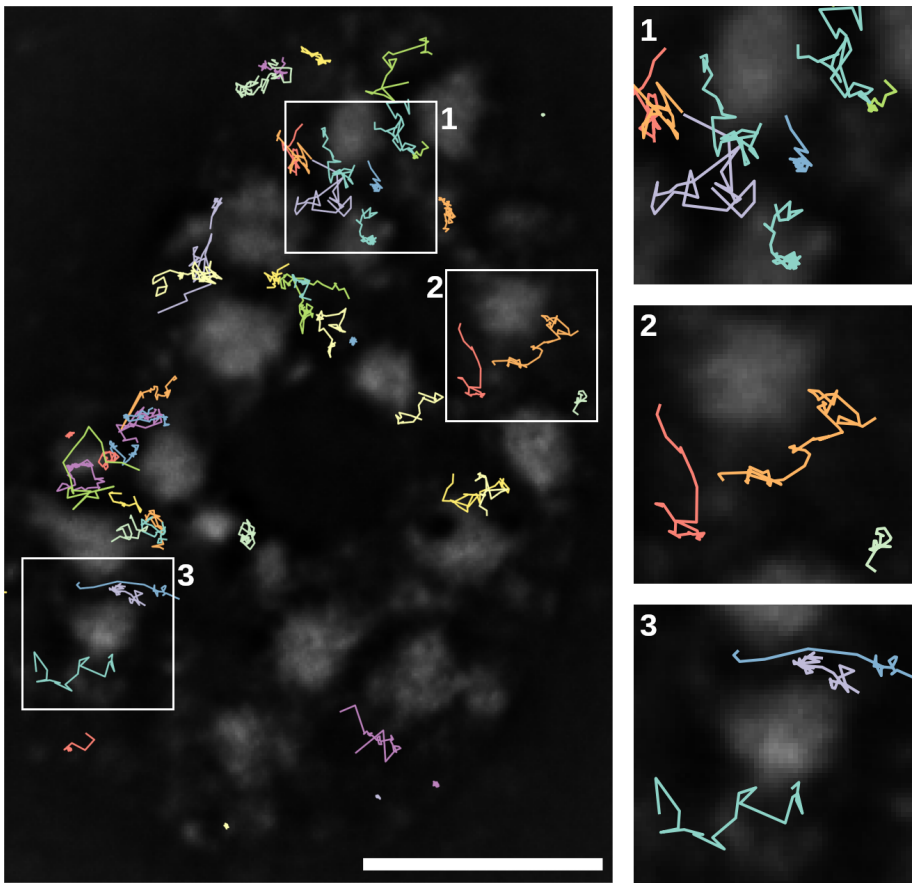

Fig. 6.2 **Recorded trajectories of p53-HaloTag in their chromatin context.** Tracked trajectories of 200ms or more plotted on top of the reference image of chromatin. Detail show p53-HaloTag interacting with SAHF. Scale bar: $5\mu m$

In order to start the quantification of this visual impression, we proceeded with the segmentation of tracks with vbSPT, in the same manner as described in previous sections. Consistent with the other results, we find that p53 separates in 3 sub-populations.

Furthermore, the slow moving molecules similarly display a diffusive behavior compatible with guided exploration, showing an anisotropic diffusion (fig 6.3.A). Moreover, once bound molecules are filtered out, we see that this anisotropy peaks when the displacement of the molecules between two frames is approximately equal to 130nm (fig 6.3.B).

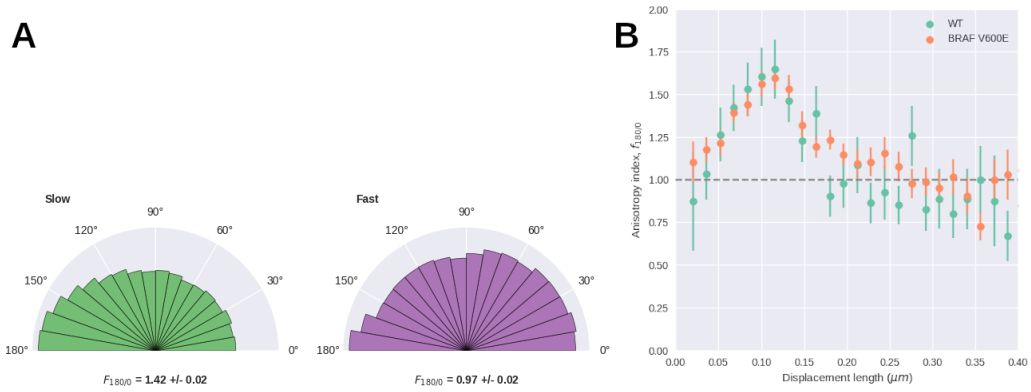

Fig. 6.3 **P53-HaloTag separates into 3 sub-populations according to their diffusion coefficient, with evidence for guided exploration.** A: Distribution of angles between consecutive displacements, for the slow and fast p53 populations. B: Fold-anistropy plotted against displacement length.

It is interesting to note that this 130nm figure was seen in all cases, from the control and nutlin treated cells shown in chapter 4.2, to the cells here transfected with either wild-type BRAF or its oncogenic version. It has been suggested in the case of CTCF, with the help of stochastic simulation, that the distance for which anisotropy peaks can be correlated with the size of the compartment which exerts the transient confinement of guided exploration [Hansen, 2020]. For CTCF, the researchers argue that these zones of transient confinement are phase separated compartments. However, our results are hard to make sense of in the light of this idea. Microscopy experiments, including ours with either in live-cell with p53-HaloTag or in fixed cells with immuno-fluorescence, show that p53 seemingly doesn't form condensates. Moreover, our results here would suggest that whatever phenomenon – or phenomena – is causing this transient confinement remains pretty much unchanged even after the induction of OIS and the appearance of SAHF. Of note, whereas the peak anisotropy was higher upon nutlin treatment in proliferating cells, here both proliferating and senescent cells have the same maximal fold-anisotropy value of 1.6. This may be reflective of the basal activation even in proliferating cells.

To see how the relationship between p53 dynamics and chromatin changes after the appearance of SAHFs in the nucleus, we then plot the dynamical enrichment plots used in the previous parts. Here we see a striking difference between the proliferating and senescent cells.

While in cells transduced with wild-type BRAF, p53 seemingly displays a behavior similar to what was observed in untreated proliferating cells, the relationship between p53 dynamics and chromatin density changes dramatically inside cells with SAHF. It is

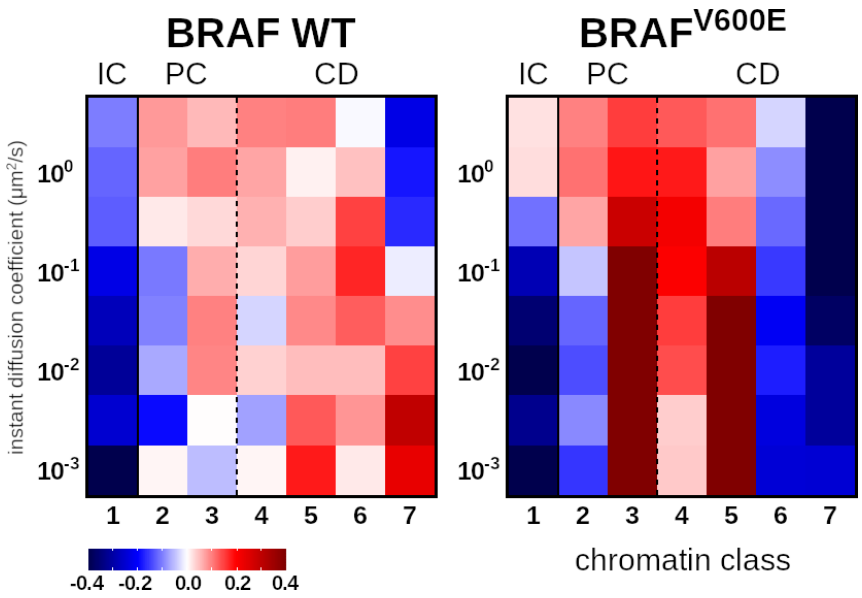

Fig. 6.4 **P53 shows a striking differences in dynamical enrichment in SAHF-presenting cells.** N_{BRAF} = 29, $N_{BRAFV600E}$ = 55.

only the fastest moving p53 molecules for which this enrichment starts to taper off, with a slight enrichment in the 2nd class. Conversely, only those molecules are not depleted in the interchromatin compartment. The behavior of p53 in SAHF-presenting cells is interesting in that it seemingly magnifies the effect we have been observing in cells which much smaller dense chromatin compartment. Slower molecules are the ones enriched in dense chromatin compartments, while the depletion of p53 from the inter-chromatin compartment is more obvious. This depletion is interesting to us because of the results of our volume exclusion model: if volume exclusion was the only phenomenon influencing p53's repartition in the nucleus, we would expect it to be flatly enriched in regions with less chromatin (see fig.5.5). With SAHF, however, p53 remain excluded from class 6 and 7, those of highest chromatin density. We hypothesize two potential explanations. Firstly, the core chromatin of SAHF, which originates from constitutive heterochromatin present along the nuclear envelope in proliferating conditions, might interact differently with transcription factors when agglomerated into highly dense foci. Such interaction may be less visible when the constitutive heterochromatin is lining the envelope, meaning we image very little of it in each microscopy. Alternatively, the exclusion of p53 could be attributed to localization errors prevalent in proliferating cells, where the smaller size of the chromatin domain might have had a greater impact on our results. The two hypothesis are not mutually exclusive, and we find likely that studying p53 dynamics around SAHF is less prone to errors owing to their larger size.

6.2 p53 in senescent cells work more efficiently than in proliferating cells.

The enrichment of p53 at the surface of the SAHF is in line with the idea of the active chromatin compartment as presented in (T. Cremer et al., 2020), and the fact that this compartment, essentially coating the exterior of inactive chromatin compartment, has some distinct characteristics, such as an enrichment in active chromatin marks ; furthermore, interestingly for our work, the organization of the compartment as a surface might represent a reduction of the dimensionality of the search for targets. This would be an interesting aspect of SAHF, since it would shift the model of their function away from a purely repressive one, and propose a mechanism by which genes exposed on the surface fo SAHF would be found more reliably by TFs.

In a first attempt to connect our analysis so far to the control of gene expression, we used a combination of single-molecule Fluorescence In Situ Hybridization (smFISH) and immuno-fluorescence (IF) to establish a relationship between the level of p53 and the level of expression of one of its pain target gene, *CDKN1A* (p21 in its protein form).

These experiments have been performed not on the BJ knock-in pool, but on our other fibroblast cell-line, the IMR90 ER:RAS cells, because the induction of senescence in those cells is much more reliable and consistent. The protocol for inducing senescence was described in section 4.1, and simply requires treating the cells with 4-OHT (or ethanol for the control) for a week. We compare cells treated this way with cells in which p53 is instead activated by a 4h treatment with nutlin (or DMSO for the control). After which the cells are fixed and hybridized with the smFISH probe against *CDKN1A* first, and the anti-p53 antibodies in a second time.

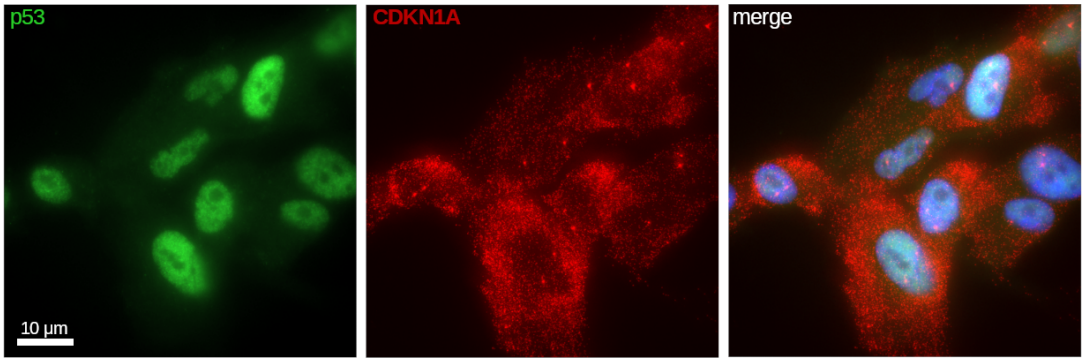

Fig. 6.5 **The combination of smFISH and immunofluorescence reveals the relationship between p53 leves and expression of its target gene *CDKN1A*.** Representative microscopy of the resulting samples after labeling. The use of microscopy allows for an analysis of the relationship between transcription factor concentration in the nucleus and the activation of its targets at the single-cell level. Note the nuclear foci of RNA, which are active transcription sites.

We acquire a z-stack each time, and after the cells are automatically segmented using the machine learning based approach Cellpose (Pachitariu and Stringer, 2022), we use FISHQuant (Imbert et al., 2022) to count the number of mRNA molecules in each cells. For each cell, we also record the fluorescence intensity given by p53. The results are presented in fig.7.

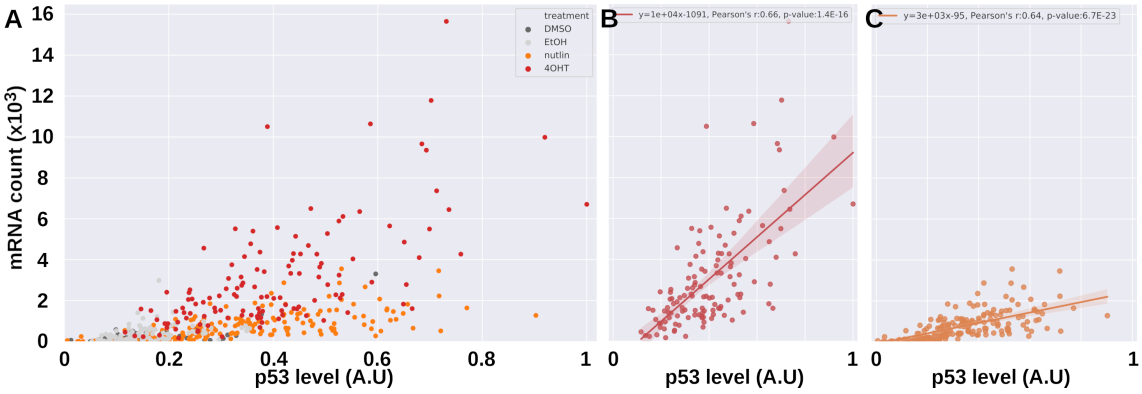

Fig. 6.6 **The relationship between expression of the target gene *CDKN1A* with respect to nuclear concentration of p53 changes in senescence.** A: Scatterplot at the single-cell level of the count of *CDKN1A* mRNA molecules against the p53 fluorescence intensity inside the nucleus. B and C: Plot of the senescent (4-OHT) and proliferating (nutlin) conditions separately, with linear fit.

As we can see, the slope relating the amount of p53 to the amount of mRNA of *CDKN1A* counted differs from the condition in which p53 is activated in proliferating cells, compared to when it is activated in cells induced into OIS, and presenting SAHF. Furthermore, when we quantify the amount of nascent mRNA molecules present in each cells, we notice that the transcriptional activity due to p53 is different in nultin and OIS cells (fig 6.7). When treating with nutlin, the increases in p53 nuclear concentration lead to both an increase in the average number of active transcription sites (TS) per cell (related to the burst frequency) and an increase in the number of RNA produced by each TS (burst size). However, this is not the case in OIS cells, where the burst size is constant with respect to nuclear concentration of p53. In contrast, OIS cells respond with an increase in the number of TS per cell, and it is this increase in burst frequency which is related to the higher count of *CDKN1A* mRNA that we have observed in fig. 6.6).

From the difference in slope between senescent cells and proliferating cells, we come to an interesting conclusion. In senescent cells, it takes less amount of p53 molecules to activate *CDKN1A* (p21) at the same level. The preferential localization of *CDKN1A* at the edge of SAHF is interesting in this regard because it could be compatible with the idea that genes at the edge of the SAHF are found more efficiently than when they are arranged according to the chromosome conformation found in proliferating cells. Of course, this idea relies on the *CDKN1A* locus being situated at the edges of SAHFs. To confirm this fact, we performed some DNA-FISH experiments, locating the *CDKN1A* locus inside OIS.

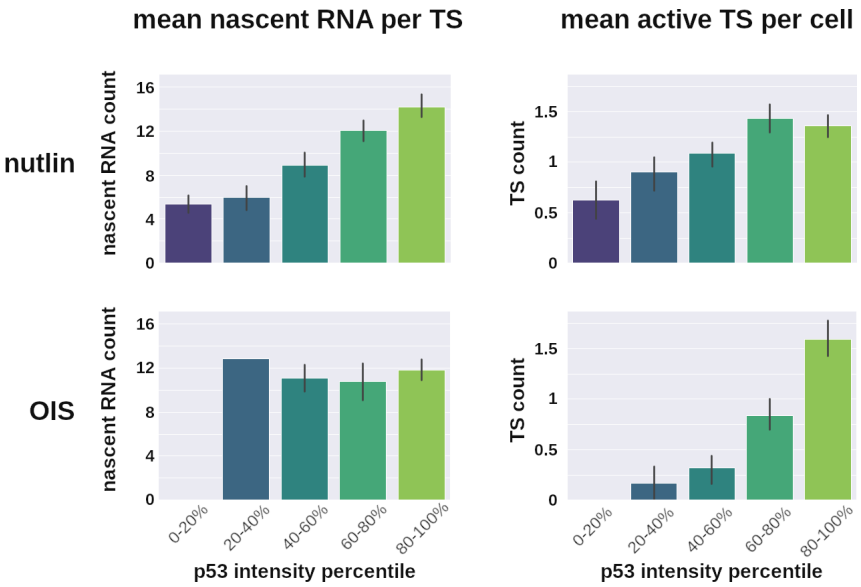

Fig. 6.7 **Increases in p53 concentration in OIS results in an increase in burst frequency but not burst size.** Quantification of the mean number of nascent RNA per transcription site (TS) and mean number of active transcription sites per cell, with respect to the nuclear concentration of p53. The analysis shows that increasing p53 concentrations in OIS cells is correlated with an increased number of transcription sites (burst frequency) while the number of RNA produced in each remains constant (burst size).

Although preliminary, the representative microscopies presented in fig.6.8 seem to indicate that the locus is located preferentially at the edges of SAHFs.

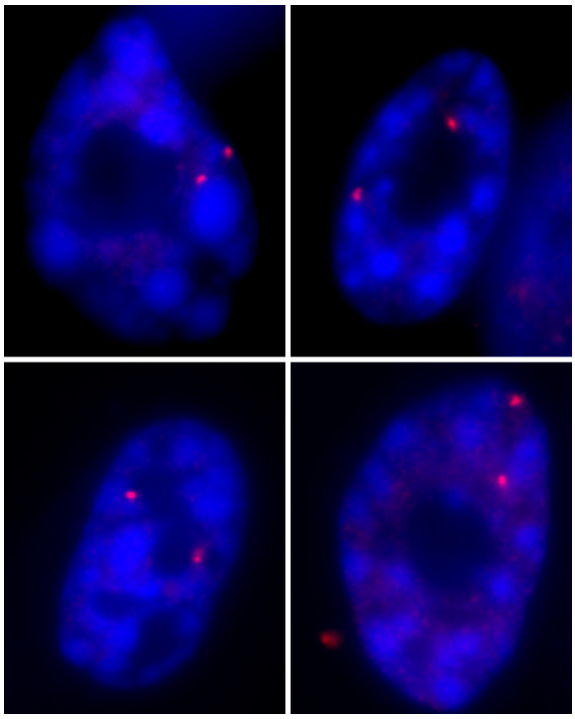

Fig. 6.8 **DNA-FISH of the *CDKN1A* inside OIS cells shows the *CDKN1A* locus at the edge of SAHF.** Preliminary results of a DNA-FISH experiment labeling the *CDKN1A* locus. The representative images presented here show it locating at the SAHF surface.

Although it is compatible, our experiment in this section are far from demonstrating this mechanism, as many other phenomena which we have talked about in section 2.3 could be at play in the difference between proliferating and senescent cells, such as a difference in PTMs. It is interesting to note that even 10 days after senescence induction, we still observe a marked correlation between p53 levels and p21 at the single-cell level. This result somewhat contradicts the idea, presented in section 3.3, that p21 expression gets decoupled from p53 in late senescence (Sheekey and Narita, 2023).

Although it is interesting to recall section 3.4, where we went over the possible influence of DNA-SCARS, the fusion of PML bodies and DNA-damage foci present in OIS. In particular, Baar et al. (2017) had shown that the disruption of the interaction between p53 and these bodies led to the release of p53 and subsequent p53-dependent activation. Here we show that even at similar concentration overall in the nucleus, p53 leads to more activation of p21, thus contrasting a bit with an idea of sequestration of p53. It has been noted before that elevated level of p53 protein lead to an increase in burst frequency, however here we show this relationship to be dependent not just on the number of p53 molecules, but also different between proliferating and cells in OIS (Hafner, Reyes, et al., 2020).

Chapter 7

Discussion.

In order to grow from a single-cell into a full organism, and then maintain homeostasis, eukaryotes need to coordinate the expression of an incredible amount of genes and non-coding RNAs. It is surprising, in the face of more genes and much larger amount of DNA, that TFs in eukaryotes are much less specific than prokaryotic TFs, much less than they ought to be if the strategy for gene regulation was the same. With orders of magnitude more DNA inside the nucleus, our TFs face the challenge of sifting through vast amount of non-specific DNA before finding a target locus that they can actually regulate. There is a long-standing idea that TFs solve this challenge bu transforming non-specific chromatin from an enemy into a guiding scaffold, thanks to non-specific interactions with DNA which allows them to slide back and forth on it.

These non-specific interactions are thought to be mediated by intrinsically-disordered regions (IDR), often a prion-like domain within TFs. Thus a number of studies have looked at mutating the IDR and seeing the effects it has on non-specific binding on one part, and the specific binding at enhancers in the other. In this study, we tried to take advantage of an event during which chromatin is dramatically rearranged, namely oncogene-induced senescence, to gain insights into how chromatin guides its own exploration by the TFs in the nucleus.

We generated cell lines from untransformed fibroblasts and performed single-molecule tracking (SMT) on p53 in those. In line with previous studies (Loffreda et al., 2017; Mazzocca, Loffreda, et al., 2022), we find in that p53 forms distinct sub-populations, which diffuse at different rates inside the nucleus. Slow and fast moving p53 differ in their diffusive behavior, with fast-diffusing p53 diffusing isotropically, while slow moving p53 diffuses anisotropically. Moreover, this anisotropy peaks for displacement of around 130nm, a behavior characteristic of guided exploration, a mechanism which can be used to accelerate the TFs' search for their targets. Some have suggested, on the basis of tracking of CTCF, that this peak of anisotropy is characteristic of the geometric properties of the compartment in which molecules diffuse and are transiently trapped, causing the observed

anisotropy (Hansen et al., 2020). By using an innovative multi-modal microscope allowing us to image chromatin with super-resolution at the same time as SMT, and correlate the diffusive behavior of p53 to its immediate chromatin context, we revealed that p53 molecules diffusing at different speed are enriched in different compartment of chromatin density, with slow and bound molecules being enriched in regions of intermediate and high chromatin density. In contrast to the previous reports on p53 dynamics in human cells, we find a much larger fraction of bound p53 molecules, as high as 40% while this figures over around 20% in cancer cells. One ChIP-seq study (Botcheva et al., 2011) reported binding profile in non-transformed cells that were significantly different than those reported in cancer cells, and this difference might be a similar reflection of the fact that the chromatin landscape differs in cancer and non-cancer cells, with consequence in the binding affinity of p53.

To study the exact nature of the relationship between p53 and chromatin, we turned to physics-based modeling. We use a model developed by (S. A. Isaacson et al., 2011) and consider the diffusion of p53 inside chromatin as the diffusion of a molecule inside a repulsive potential, modeling the idea of volume exclusion which has long been an idea to explain the repressive nature of heterochromatin. Whereas the theoretical exploration of this model by S. A. Isaacson et al. (2011) showed that volume exclusion could reduce the time taken by TFs to find their target in regions of euchromatin, our experiments gave us access to data to test this model. We used our SMT/SIM data to determine the strength of this repulsive potential. Surprisingly, we find that this volume exclusion model is incapable to describe our data on p53. When extending the analysis to other nuclear factors, we find that the volume exclusion model seems compatible with the NF-kB sub unit RelA. Yet, even in this case, the strength of the volume exclusion exerted by chromatin is well below the values predicted to be beneficial for the target search, or those that give rise to diffusion anisotropies similar to the ones measured in experiments.

We next extended the model to account for the ability of NFs to interact non-specifically with DNA for short period of times. Such extended model could reproduce how different NFs - from inert tracers to stronger chromatin binders, like CTCF or Histone H2B, distribute within regions with different chromatin densities and how chromatin density affect their diffusion speed.

The core lesson which we derive from this model is that binding interactions must happen in order for transcription factors to overcome volume exclusion and penetrate inside dense chromatin domains. This is in contrast to previous physics-based models by Maeshima, Kaizu, et al. (2015) which found the size of nuclear factor – thus, presumably, volume exclusion as in their ability to fit in between the chromatin fibers – to be determinant in the ability of molecules to penetrate inside chromatin dense regions. This model was used by (Miron et al., 2020) to explain the zonation of various chromatin marks in relationship to the depth inside the chromatin domain at which they were found. Put simply, smaller

chromatin remodeler can diffuse deeper in chromatin domains, and thus a simple way to limit the spread of active chromatin marks is to have all the associated remodelers be to big to penetrate. Our addition to this debate is to use live-cell imaging and relate not just enrichment in different compartments, but how this affects diffusion speed. In this we find that chromatin interaction is crucial to explain the behavior of transcription factors, even though Maeshima, Kaizu, et al. (2015) argued those molecules small enough to diffuse on their own inside chromatin dense regions. Similar modeling strategies could be employed, using various forces acting on the diffusion of the molecules, such as an attraction to chromatin when in the binding state, or an attraction to area sharp contrast in chromatin signal (i.e. perichromatin), but the model used here has the advantage of making few assumptions on the underlying physics of diffusion.

With this understanding, we come back to the specific case of senescence. During OIS, the chromatin landscape changes dramatically, with notably the appearance of large foci of heterochromatin, SAHF. These SAHF are generally understood as having a repressive function, sequestering away key genes related to apoptosis and the cell-cycle, stabilizing senescence. Using SMT/SIM, we image the motion of p53 in and around those chromatin domains, and find that p53, and in particular the slow-diffusing sub-population of p53 molecules, spends more time than expected on the surface of these domains. These findings may be a much more dramatic visualization of the same kind of dynamical enrichment we observed in proliferating cells, but revealed with more clarity thanks to the size of these chromatin domains. Given that the exploration of a surface, in 2D, reduces the dimensionality of the target search, we would expect this exploration to find targets more efficiently, on the condition that they are present on the surface of the SAHF. We provide tentative evidence going in this direction. First by localizing with DNA-FISH the p53 target gene *CDKN1A* at the surface of SAHFs, then by combining smFISH and immunofluorescence, showing that the relationship between the nuclear concentration of p53 and the activation of its target changes in senescence. In cells presenting SAHF, the same amount of p53 will lead to an increased expression of *CDKN1A*, thanks to an increase in the frequency with which the gene is activated. While still preliminary, the study of p53 dynamics around SAHF revealed much more clearly the depletion of p53 inside regions of poor chromatin density. This fact both confirms to us that SAHF are a good object of study, but is also in line with our modeling results, in that this fact is not explainable by volume exclusion alone, but requires to take the non-specific interactions of p53 with chromatin into account.

While these results are compatible with the idea that p53 finds its targets more frequently at the surface of SAHF, further work would be necessary to prove it conclusively. This is because there are many other reasons why p53 could bind more frequently to the *CDKN1A* promoter, including PTM or co-factors, and other mechanisms such as changing degradation rates could also affect the number of mRNA molecules. Further analysis is

needed to measure important factors such as the changes in PTM or the degradation rates of mRNA. However, the analysis of p53 levels and associated change in the expression of target genes is informative for the elucidation of those factors as well. An interesting extension of this experiment would be to look at many more target genes simultaneously, perhaps using high-throughput, multiplex methods such as MERFISH. We would expect the effect of being located on the surface to affect all genes under p53's control, and the extended analysis would help us outline the real contribution of this mechanism to gene expression. Another avenue we plan to explore in the near future is the transfection of a silencing RNA against DNMT1, which (Sati et al., 2020) has shown to abolish the formation of SAHF. The comparison between cells expressing $BRAF^{V600E}$ with or without DNMT1 may bring us closer to understanding the specific effect of SAHF.

Taken together, our research provides new insights into the behavior of TFs within chromatin. We demonstrated that a minimal model for TFs behavior needs to account for transient binding in order to reproduce the diffusive behavior of TFs and in particular the ability of some to penetrate heterochromatin. We further show that p53 exhibits a particular slow mode of diffusion, correlated with regions of intermediate chromatin density. This behavior is made all the more striking in cells presenting SAHFs, where p53 seemingly scans the surface. These findings further our understanding of the interplay between chromatin architecture and transcription factors, and can be expanded upon in the future.

Chapter 8

Materials and Methods.

Cell culture.

IMR90 ER:RAS were gifted to us by the Peter Adams lab (University of Glasgow, UK). BJ cells were bought from ATCC (ATCC, CRL-2522). All cell lines are grown in DMEM (Thermo-Fisher, cat. 15140122), with 10% heat-inactivated Fetal Bovine Serum (Gibco), 100 units/mL of penicillin-streptomycin (Thermo-Fisher, cat. 25030081). The cells are grown at 37°C and 5% CO2 and passaged every 2 days with trypsine. The cells were tested for mycoplasma contamination once a week via PCR.

Other cell lines ere used in chapter 5 to analyze the behavior of different transcription factors. The DIvA cells were gifted to us by the Legube lab [Aymard, 2014], used for the HaloTag, p65 and H2B condition. U2OS cells stably knocked-in with CTCF-HaloTag were gifted to us by the Darzacq and Tijan group [Hansen, 2020], and used for the CTCF condition. All cells were cultured as described above, barring the addition of 1ug/mL of puromycin (Thermo-Fisher, cat. A1113803) in the case of the DIvA cells, which is used to maintain the expression of the AsiSI-RE enzyme.

Generation of the BJ knock-in cell line.

We used the nickase variant of CRISPR/Cas9 genome editing for the generation of the knock-in cells (Ran et al., 2013).

The transfection uses four plasmids, containing the two guiding RNAs, cas9, and p53-HaloTag flanked by two homology arms for homologous recombination of around 800bp each (fig 8.1). The insertion site delineated by the two nicks is at location chr17:7,669,671-7,669,740 (GRCh38/hg38), near exon 11 of the endogenous TP53.

Cells were seeded in 12-well plates the day prior to the transfection. The transfection was done using jetOPTIMUS reagent (Polyplus, cat. 101000051) according to the manufacturer protocol. The amount of DNA for each plasmid was adjusted depending on

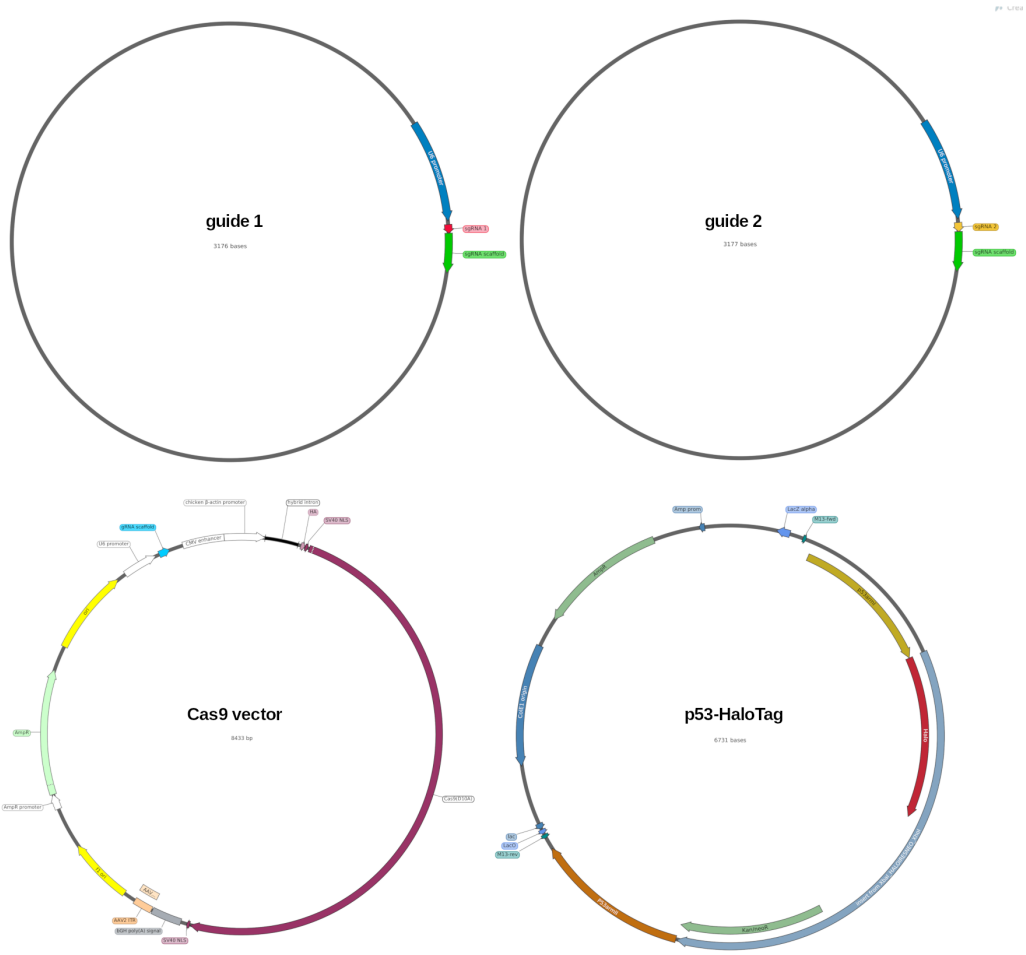

Fig. 8.1 **Map of the four vectors used in the transfection for cas9 knock-in.**

plasmid length, with guide 1, guide 2, cas9 and p53-HaloTag at 730, 730, 1950, 1500ng respectively, 1.5ul of jetOPTIMUS reagent and 100ul of buffer. Cells were incubated for 6h before the transfection medium was removed. After 3 days of recovery, the cells were put under mild antibiotic selection with geneticin (G418 Sulfate, Thermo-Fisher, cat. 10131027) at 300ug/mL for a total of 2 days. As presented in fig.4.2, we screened the pools for the presence of p53-HaloTag and the absence of p53 wild-type using Western blot, after which the functionality of p53-HaloTag was measured with qPCR, and its nuclear localization with microscopy.

Senescence Induction.

Senescence induction

For IMR90 ER:RAS cells, OIS is induced by growing the cells in the regular growth medium with 100nM of 4-hydroxy-tamoxifen (4-OHT). The cells along with a control which receives an equivalent volume of 70% ethanol, are passaged every 2 days and fresh

4-OHT is added. Once fully senescent, after 8 days, cells are stable for a some time, but are generally used right away.

For BJ cells, we use lentiviral infection with the oncogene $BRAF^{V600E}$. The prepared virus were kindly gifted to us by the Eugenio Montini lab (IRCCS Ospedale San Raffaele Scientific Institute, Milano, Italy). Cells are plated 24h beforehand at 70% confluency. The multiplicity of infection used was 4 for the wild-type BRAF and 20 for $BRAF^{V600E}$. The cells were incubated over night and then medium was changed back to growth medium.

Plasmids and transient transfections.

The plasmid used in chapter 5 for transient transfection were synthesized by Genewiz. The vectors use pFN22A backbone (Promega) for p53, pFC15A (Promega) for H2B, pci-Neo-p65-Halotag for p65, and pHTN (Promega) for HaloTag. For p53, the transfection used a p53-KO cell line derived from the DIvA cell line.

For the transfection, cells were plated 2 days before the experiment. Once they reach 60% confluency, the cells were transfected using Lipofectamine 3000 (Thermo-Fisher, cat. L3000008) according to the manufacturer protocol, and imaged on the following day.

Western blot.

Cells cultured for Western blot are washed in PBS and lysed in 300 ul of RIPA buffer (25nM Tris Hcl pH 7.6, 150 nM NaCl, 1% Sodium deoxycholate, 1% Triton X-100, 2mM EDTA dihydrate) with protease inhibitors (Sigma-Aldrich, cat. 4693124001). After 20min incubation at 4degreeC with rotation, the lysate is centrifuged at 12,000 rpm and the supernatant collected. The lysate is quantified by BCA assay (Thermo-Fisher, cat. 23225), loaded on a gel (8% SDS-polyacrilamide) and run at 100V. Proteins are transferred with a cold transfer buffer (25mM Tris, 192mM glycine, 20% methanol) on nitrocellulose membranes which are then blocked in 5% non-fat dried milk in TBS-T solution (0.1% Tween20 in TBS) for 1h at RT with shaking. Primary antibody incubation is performed in the same solution. Afterwards, membranes were washed three times in TBS-T and then incubated for 1h at room temperature with peroxydase-conjugated secondary antibodies (anti-mouse IgG, Cell signaling, cat. 7076, anti-rabbit IgG, Cell Signaling, cat. 7074) diluted in TBS-T. The membranes were developed using ECL substrate (Bio Rad, cat. 1705061) and images were acquired with a CCD camera using ChemiDoc MP imaging system.

The primary antibodies used for fig. 4.2 are rabbit monoclonal anti-GADPH (Abcam, cat. Ab128915) and mouse monoclonal anti-p53 DO1 (Santa-Cruz Biotechnology, cat. sc-126).

RNA extraction and RT-qPCR.

Cells grown on 6-cm dishes were washed once in PBS and lysed in 750 ul of TRIzol reagent (Thermo-Fisher, cat. 15596018). The lysates are purified using silica membranes columns (Machery-Nagel, Nucleospin RNA plus). The RNA thus extracted is checked for purity using NanoDrop fluorometer (Thermo-Fisher). Reverse-transcription was made using High-Capacity cDNA Reverse Transcription Kit (ThermoFisher cat. 4368814) according to the manufacturer's protocol. Real-time qPCR analysis was performed with 5uL of cDNA solution, 150nM of primers and SYBR Green mix (Roche, LightCycler 480 I Master). Along with genes of interest, a triplicate used GAPDH primers to normalize to a housekeeping gene. The relative expression presented use the $\Delta\Delta CT$ method.

smFISH and immunofluorescence combination.

The Single-molecule Fluorescence In Situ Hybridization (smFISH) with immunofluorescence (IF) presented in fig 6.5. was performed on IMR90 ER:RAS cells. First, the ethanol and 4-OHT condition were seeded in 6-well plates a week and cultured in a growth medium with 100nM of 4-hydroxy-tamoxifen. The cells are passaged every 48h. Cells are seeded in 24-well plate with glass covers at the bottom 24h hours before the start of the experiment. For the nutlin condition, the growth medium is replaced with one containing 10 uM of Nutlin-3a (Sigma-Aldrich, SML0580) for 4 hours, while an equivalent volume of DMSO is added to the control condition.

The covers are washed in PBS 3 times and fixed using 4% formaldehyde for 15min at RT. After washing, they are permeabilized using Triton (0.3% Triton X-100, 2mM Vanadyl ribonucleoside complexes in PBS) for 60min. Cells are washed then equilibriated in SSC solution (Thermo-Fisher, Ambion, cat. AM9763), then a solution of SSC and formamide for 10 minute each. RNA probes against CDKN1A (HuluFISH) are in the meantime diluted in a hybridization buffer (2x SSC, 10% formamide, 10% dextran sulfate in RNAse-free water). The covers are placed on top of 50ul drops are the solution and left to incubate at 37°C overnight. The covers are then washed in the SSC/formamide solution for an hour at 37°C and then fixed again with 4% formaldehyde. The covers are then incubated in a solution of primary anti-p53 antibodies for 1h then secondary antibodies for 1h. Finally, they are washed and counter-stained with Hoechst, mounted and imaged.

We acquire a z-stack of the cells at 300nm intervals. Cellpose is used to automatically segment the resulting images, the resulting segmentations are checked and corrected manually, and then FISH-Quant is used for the quantification of mature and nascent RNA in each cells. P53 levels are quantified using the normalized mean intensity inside the segmented nuclei, for a single slice at the z-midpoint of the nuclei.

SIM/SMT

The single-molecule approach used in this thesis relies on a custom-built widefield microscope using an inclined laser beam [tokugana, 2008]. We use a 60x oil-immersion objective (Olympus Life Science, N.A. = 1.49), a Hamamatsu Orca Fusion sCMOS camera (Hamamatsu Photonics, physical pixel size 108.3 nm), and a control system to keep a chamber around the sample in physiological conditions of 37°C, 5% CO2 and humidity. The acquisition presented in this work used a 405nm laser (Coherent, continuous wave current) for UV excitation of the Hoechst label of chromatin, and a 561nm laser (Cobolt 60-DPL, 20mW power). The 561nm excitation followed a stroboscopic scheme, with 5ms of the 10ms frame illuminated to limit motion blur.

The acquisition of SIM images relies on a digital micromirror device placed on the optical path of the UV laser (DMD, Vialux, v7000). The use of these movable mirrors of 13.67µm in size allows the projection of an array of illuminated dots which are scanned across the sample. We record 224 different images which are later reconstructed into a super-resolved image using a custom MATLAB script written by Davide Mazza and Emanuele Colombo in our lab. Concomitantly, another illumination arm reflects the 561 laser used for SMT. On this arm, a moveable mirror in a conjugated plane of the back focal aperture of the objective allows us to incline the light, typically at around 67. The two arms of HILO and SMT are recombined through a dichroic mirror (DM1) transmitting wavelength longer than 490nm (Semrock Inc.). A schematic representation of the microscope is presented in fig. 8.2.

For a SIM/SMT experiment, the cells are seeded in LabTek chambers (Thermo Scientific, cat. 16250671) a day before the experiment. They are incubated for 30min at 37°C in 1nM JF549 ligand in phenol-free medium. They are washed three times in PBS before being incubated in Hoechst 33342 at 2µm/mL for 10min and washed again. Before starting the acquisition the cells are bleached such that the density of fluorescent molecules decreases. The acquisition consist in taking the 224 frames for the SIM, then 20s of SMT (10ms per frame, for a total of 2000 frames), then another SIM stack. In many cases, another SMT and SIM are recorded.

Tracking is done with the ImageJ plug-in Trackmate, detecting spots of 0.8 um and using the simple LAP tracker with no blinking allowed, with a maximum displacement of 1um for one frame to another. Trajectories thus tracked are then combined if more than one movie exists for a single cell. Frames with more than 15 molecules present are ignored from further analysis to avoid tracking errors.

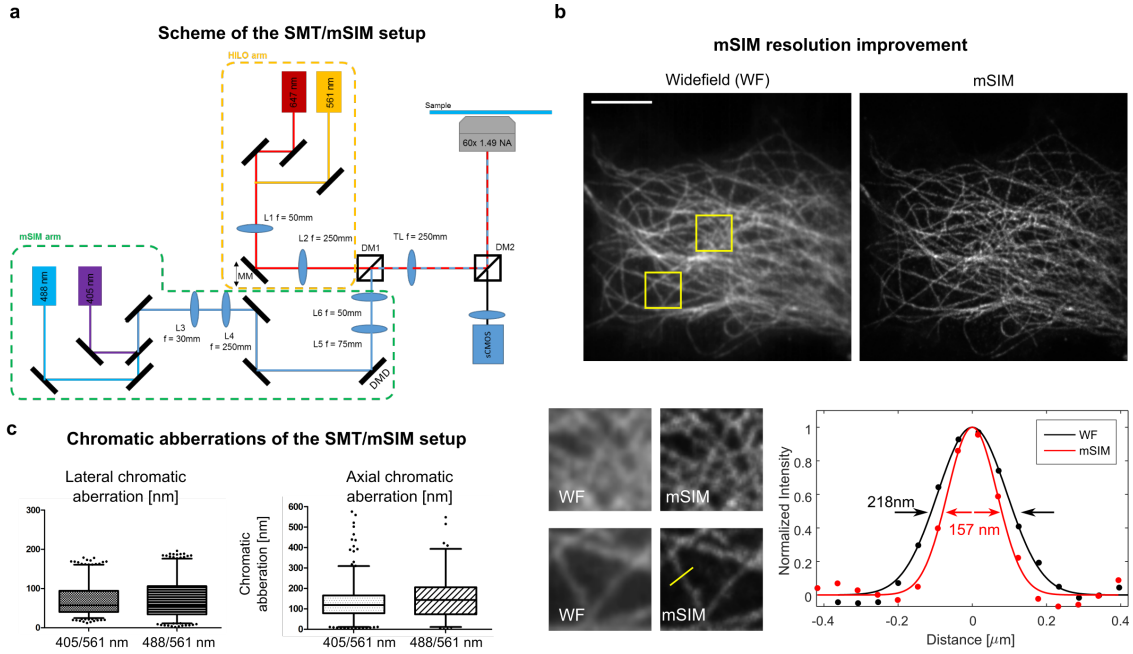

Fig. 8.2 **Scheme of the SIM/SMT microscope. A:** Scheme of the micrscope. **B:** Principle for a SIM acquisition. **C:** Resolution gain measured from microtubule images. **D:** Chromatic aberrations of SIM quantified with 100nm fluorescent beads.

Histogram fitting

The tracks from the experiments are used to generate a histogram of displacement from frame to frame, with a bin size of 20nm. From theory, we can expect the form of the histogram of a mixture of N population of molecules diffusing at their own diffusion coefficient D_i to be:

$$\mathrm{P(r|D)} = \sum_{i}^{N} \left[\frac{n_i'}{N_{tot}} \times \frac{\mathrm{r}}{2Dt} \exp\left(-\frac{\mathrm{r}^2}{4Dt} \right) \right] \times \Delta r \tag{8.1}$$

where:

1. n' is the corrected number of observations.
2. D is the diffusion coefficient.
3. t is the time lag.
4. Δr) is the size of each bin in the histogram.

The correction is necessary to take into account the defocalization bias, which is the fact that fast moving molecules are more likely to diffuse out of the focus slice than slow moving ones, and thus are under-counted (Mazza et al., 2012). The selection between

models with different numbers of sub-populations is done using the Bayesian Information Criteria. Track classification is then handled with vbSPT (Persson et al., 2013). We find that, despite being supposed to converge on the smallest possible number of states, on real data vbSPT tends to fit the maximum number of states specified each time. Thus we impose a number of 3 states guided by the histogram analysis.

Model Fitting

The parameters of the model are fitted using the formula in annex for the steady-state probability. We use the acquisition obtained by Matteo and accepted for publication at the time of writing (Mazzocca, Loffreda, et al., 2022). For each pixel of the nucleus of each acquisition, we record the number of particle spotted in it over the course of the acquisitions, as well as the intensity of chromatin at this pixel. With the formula linking chromatin density and probability of occupancy in the steady-state, we can compute the likelihood of the observation. The minimum of the likelihood is found using either gradient descent (volume exclusion model) or grid search (volume exclusion + non-spécific binding model). To estimate error bars, we resampled 50

References

Acosta, Juan C. et al. (June 2008). "Chemokine Signaling via the CXCR2 Receptor Reinforces Senescence". en. In: *Cell* 133.6, pp. 1006–1018. ISSN: 0092-8674.

Acosta, Juan Carlos et al. (Aug. 2013). "A complex secretory program orchestrated by the inflammasome controls paracrine senescence". en. In: *Nat Cell Biol* 15.8. Number: 8 Publisher: Nature Publishing Group, pp. 978–990. ISSN: 1476-4679.

Adams, Peter D. (Aug. 2007). "Remodeling of chromatin structure in senescent cells and its potential impact on tumor suppression and aging". en. In: *Gene* 397.1, pp. 84–93. ISSN: 0378-1119.

Akdemir, Kadir C. et al. (Jan. 2014). "Genome-wide profiling reveals stimulus-specific functions of p53 during differentiation and DNA damage of human embryonic stem cells". en. In: *Nucleic Acids Research* 42.1, pp. 205–223. ISSN: 1362-4962, 0305-1048.

Aksoy, Ozlem et al. (July 2012). "The atypical E2F family member E2F7 couples the p53 and RB pathways during cellular senescence". en. In: *Genes Dev.* 26.14. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab, pp. 1546–1557. ISSN: 0890-9369, 1549-5477.

Allen, Mary Ann et al. (May 2014). "Global analysis of p53-regulated transcription identifies its direct targets and unexpected regulatory mechanisms". en. In: *eLife* 3, e02200. ISSN: 2050-084X.

Amitai, Assaf (Feb. 2018). "Chromatin Configuration Affects the Dynamics and Distribution of a Transiently Interacting Protein". en. In: *Biophysical Journal* 114.4, pp. 766–771. ISSN: 00063495.

Amitai, Assaf, Andrew Seeber, Susan M. Gasser, and David Holcman (Jan. 2017). "Visualization of Chromatin Decompaction and Break Site Extrusion as Predicted by Statistical Polymer Modeling of Single-Locus Trajectories". en. In: *Cell Reports* 18.5, pp. 1200–1214. ISSN: 22111247.

Axpe, Eneko et al. (Sept. 2019). "A Multiscale Model for Solute Diffusion in Hydrogels". en. In: *Macromolecules* 52.18, pp. 6889–6897. ISSN: 0024-9297, 1520-5835.

Baar, Marjolein P. et al. (Mar. 2017). "Targeted Apoptosis of Senescent Cells Restores Tissue Homeostasis in Response to Chemotoxicity and Aging". English. In: *Cell* 169.1. Publisher: Elsevier, 132–147.e16. ISSN: 0092-8674, 1097-4172.

Baker, Darren J. et al. (Nov. 2011). "Clearance of p16Ink4a-positive senescent cells delays ageing-associated disorders". en. In: *Nature* 479.7372. Number: 7372 Publisher: Nature Publishing Group, pp. 232–236. ISSN: 1476-4687.

Bancaud, Aurélien et al. (Dec. 2009). "Molecular crowding affects diffusion and binding of nuclear proteins in heterochromatin and reveals the fractal organization of chromatin". en. In: *EMBO J* 28.24, pp. 3785–3798. ISSN: 0261-4189, 1460-2075.

Barlev, Nickolai A. et al. (Dec. 2001). "Acetylation of p53 Activates Transcription through Recruitment of Coactivators/Histone Acetyltransferases". en. In: *Molecular Cell* 8.6, pp. 1243–1254. ISSN: 1097-2765.

Baugh, Evan H et al. (Jan. 2018). "Why are there hotspot mutations in the TP53 gene in human cancers?" en. In: *Cell Death Differ* 25.1, pp. 154–160. ISSN: 1350-9047, 1476-5403.

Beauséjour, Christian M. et al. (Aug. 2003). "Reversal of human cellular senescence: roles of the p53 and p16 pathways". eng. In: *EMBO J* 22.16, pp. 4212–4222. ISSN: 0261-4189.

Beck, Martin et al. (Jan. 2011). "The quantitative proteome of a human cell line". en. In: *Mol Syst Biol* 7.1, p. 549. ISSN: 1744-4292, 1744-4292.

Bénichou, O., M. Moreau, P.-H. Suet, and R. Voituriez (June 2007). "Intermittent search process and teleportation". eng. In: *J Chem Phys* 126.23, p. 234109. ISSN: 0021-9606.

Bevilacqua, Philip C., Allison M. Williams, Hong-Li Chou, and Sarah M. Assmann (Jan. 2022). "RNA multimerization as an organizing force for liquid–liquid phase separation". en. In: *RNA* 28.1, pp. 16–26. ISSN: 1355-8382, 1469-9001.

Blainey, Paul C. et al. (Apr. 2006). "A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA". In: *Proceedings of the National Academy of Sciences* 103.15. Publisher: Proceedings of the National Academy of Sciences, pp. 5752–5757.

Boija, Ann et al. (Dec. 2018). "Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains". en. In: *Cell* 175.7, 1842–1855.e16. ISSN: 00928674.

Bolzer, Andreas et al. (Apr. 2005). "Three-Dimensional Maps of All Chromosomes in Human Male Fibroblast Nuclei and Prometaphase Rosettes". en. In: *PLoS Biol* 3.5. Ed. by Tom Misteli, e157. ISSN: 1545-7885.

Botcheva, Krassimira et al. (Dec. 2011). "Distinct p53 genomic binding patterns in normal and cancer-derived human cells". In: *Cell Cycle* 10.24, pp. 4237–4249. ISSN: 1538-4101.

Bourgeois, Benjamin and Tobias Madl (2018). "Regulation of cellular senescence via the FOXO4-p53 axis". en. In: *FEBS Letters* 592.12. _eprint: https://febs.onlinelibrary.wiley.com/doi/pdf/10.1002/1873-3468.13057, pp. 2083–2097. ISSN: 1873-3468.

Boyle, Alan P. et al. (Jan. 2008). "High-Resolution Mapping and Characterization of Open Chromatin across the Genome". en. In: *Cell* 132.2, pp. 311–322. ISSN: 00928674.

Boyle, Evan A., Yang I. Li, and Jonathan K. Pritchard (June 2017). "An Expanded View of Complex Traits: From Polygenic to Omnigenic". en. In: *Cell* 169.7, pp. 1177–1186. ISSN: 00928674.

Brady, Colleen A. et al. (May 2011). "Distinct p53 Transcriptional Programs Dictate Acute DNA-Damage Responses and Tumor Suppression". en. In: *Cell* 145.4, pp. 571–583. ISSN: 00928674.

Brandt, Tobias, Miriana Petrovich, Andreas C. Joerger, and Dmitry B. Veprintsev (Dec. 2009). "Conservation of DNA-binding specificity and oligomerisation properties within the p53 family". In: *BMC Genomics* 10.1, p. 628. ISSN: 1471-2164.

Buenrostro, Jason D et al. (Dec. 2013). "Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position". en. In: *Nat Methods* 10.12, pp. 1213–1218. ISSN: 1548-7091, 1548-7105.

Campisi, Judith (1997). "Aging and Cancer: The Double-Edged Sword of Replicative Senescence". en. In: *Journal of the American Geriatrics Society* 45.4. _eprint: https://agsjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1532-5415.1997.tb05175.x, pp. 482–488. ISSN: 1532-5415.

Carrà, Giovanna et al. (Nov. 2020). "P53 vs NF-κB: the role of nuclear factor-kappa B in the regulation of p53 activity and vice versa". en. In: *Cell. Mol. Life Sci.* 77.22, pp. 4449–4458. ISSN: 1420-682X, 1420-9071.

Catizone, Allison N et al. (May 2020). "Locally acting transcription factors regulate p53-dependent cis-regulatory element activity". en. In: *Nucleic Acids Research* 48.8, pp. 4195–4213. ISSN: 0305-1048, 1362-4962.

Cerase, Andrea, J. Mauro Calabrese, and Gian Gaetano Tartaglia (Mar. 2022). "Phase separation drives X-chromosome inactivation". en. In: *Nat Struct Mol Biol* 29.3, pp. 183–185. ISSN: 1545-9993, 1545-9985.

Chaib, Selim, Tamar Tchkonia, and James L. Kirkland (Aug. 2022). "Cellular senescence and senolytics: the path to the clinic". en. In: *Nat Med* 28.8. Number: 8 Publisher: Nature Publishing Group, pp. 1556–1568. ISSN: 1546-170X.

Chakraborty, Shreeta et al. (Feb. 2023). "Enhancer–promoter interactions can bypass CTCF-mediated boundaries and contribute to phenotypic robustness". en. In: *Nat Genet* 55.2. Number: 2 Publisher: Nature Publishing Group, pp. 280–290. ISSN: 1546-1718.

Chandra, Tamir, Philip Andrew Ewels, et al. (Jan. 2015). "Global Reorganization of the Nuclear Landscape in Senescent Cells". In: *Cell Rep* 10.4, pp. 471–483. ISSN: 2211-1247.

Chandra, Tamir, Kristina Kirschner, et al. (July 2012). "Independence of Repressive Histone Marks and Chromatin Compaction during Senescent Heterochromatic Layer Formation". In: *Mol Cell* 47.2, pp. 203–214. ISSN: 1097-2765.

Chang, Nien-Tzu et al. (Mar. 2007). "The transcriptional activity of HERV-I LTR is negatively regulated by its cis-elements and wild type p53 tumor suppressor protein". en. In: *J Biomed Sci* 14.2, pp. 211–222. ISSN: 1021-7770, 1423-0127.

Charoensawan, Varodom, Derek Wilson, and Sarah A. Teichmann (Nov. 2010). "Genomic repertoires of DNA-binding transcription factors across the tree of life". en. In: *Nucleic Acids Research* 38.21, pp. 7364–7377. ISSN: 1362-4962, 0305-1048.

Chen, Hsuan-An et al. (Feb. 2023). "Senescence Rewires Microenvironment Sensing to Facilitate Antitumor Immunity". eng. In: *Cancer Discov* 13.2, pp. 432–453. ISSN: 2159-8290.

Chen, Yu et al. (Nov. 2022). "Mechanisms governing target search and binding dynamics of hypoxia-inducible factors". en. In: *eLife* 11, e75064. ISSN: 2050-084X.

Childs, Bennett G et al. (Nov. 2014). "Senescence and apoptosis: dueling or complementary cell fates?" In: *EMBO reports* 15.11. Publisher: John Wiley & Sons, Ltd, pp. 1139–1153. ISSN: 1469-221X.

Ciriello, Giovanni et al. (Oct. 2013). "Emerging landscape of oncogenic signatures across human cancers". en. In: *Nat Genet* 45.10, pp. 1127–1133. ISSN: 1061-4036, 1546-1718.

Collado, Manuel and Manuel Serrano (June 2006). "The power and the promise of oncogene-induced senescence markers". en. In: *Nat Rev Cancer* 6.6. Number: 6 Publisher: Nature Publishing Group, pp. 472–476. ISSN: 1474-1768.

Coppé, Jean-Philippe et al. (Dec. 2008). "Senescence-Associated Secretory Phenotypes Reveal Cell-Nonautonomous Functions of Oncogenic RAS and the p53 Tumor Suppressor". en. In: *PLOS Biology* 6.12. Publisher: Public Library of Science, e301. ISSN: 1545-7885.

Correia-Melo, Clara, Graeme Hewitt, and João F Passos (Jan. 2014). "Telomeres, oxidative stress and inflammatory factors: partners in cellular senescence?" In: *Longev Healthspan* 3, p. 1. ISSN: 2046-2395.

Cortini, Ruggero and Guillaume J. Filion (Apr. 2018). "Theoretical principles of transcription factor traffic on folded chromatin". en. In: *Nat Commun* 9.1, p. 1740. ISSN: 2041-1723.

Cravens, Shannen L. et al. (Apr. 2015). "Molecular crowding enhances facilitated diffusion of two human DNA glycosylases". en. In: *Nucleic Acids Research* 43.8, pp. 4087–4097. ISSN: 1362-4962, 0305-1048.

Cremer, Marion and Thomas Cremer (2018). "Nuclear compartmentalization, dynamics, and function of regulatory DNA sequences". In: *Genes, Chromosomes and Cancer* 58.7, pp. 427–436.

Cremer, Thomas et al. (Feb. 2020). "The Interchromatin Compartment Participates in the Structural and Functional Organization of the Cell Nucleus". en. In: *BioEssays* 42.2, p. 1900132. ISSN: 0265-9247, 1521-1878.

Cusanovich, Darren A., Bryan Pavlovic, Jonathan K. Pritchard, and Yoav Gilad (Mar. 2014). "The Functional Consequences of Variation in Transcription Factor Binding". en. In: *PLoS Genet* 10.3. Ed. by Yitzhak Pilpel, e1004226. ISSN: 1553-7404.

De Cecco, Marco et al. (Apr. 2013). "Genomes of replicatively senescent cells undergo global epigenetic changes leading to gene silencing and activation of transposable elements". eng. In: *Aging Cell* 12.2, pp. 247–256. ISSN: 1474-9726.

Demaria, Marco et al. (Dec. 2014). "An Essential Role for Senescent Cells in Optimal Wound Healing through Secretion of PDGF-AA". In: *Dev Cell* 31.6, pp. 722–733. ISSN: 1534-5807.

Di Micco, Raffaella et al. (Nov. 2006). "Oncogene-induced senescence is a DNA damage response triggered by DNA hyper-replication". en. In: *Nature* 444.7119. Number: 7119 Publisher: Nature Publishing Group, pp. 638–642. ISSN: 1476-4687.

Dimauro, Teresa and Gregory David (Dec. 2010). "Ras-induced senescence and its physiological relevance in cancer". eng. In: *Curr Cancer Drug Targets* 10.8, pp. 869–876. ISSN: 1873-5576.

Dimri, G P et al. (Sept. 1995). "A biomarker that identifies senescent human cells in culture and in aging skin in vivo." In: *Proceedings of the National Academy of Sciences* 92.20. Publisher: Proceedings of the National Academy of Sciences, pp. 9363–9367.

Dolgin, Elie (2017). "A tour through the most studied genes in biology reveals some surprises." en. In.

Dou, Zhixun et al. (Oct. 2017). "Cytoplasmic chromatin triggers inflammation in senescence and cancer". en. In: *Nature* 550.7676. Number: 7676 Publisher: Nature Publishing Group, pp. 402–406. ISSN: 1476-4687.

Dyson, H. Jane and Peter E. Wright (Mar. 2005). "Intrinsically unstructured proteins and their functions". en. In: *Nat Rev Mol Cell Biol* 6.3, pp. 197–208. ISSN: 1471-0072, 1471-0080.

Elf, Johan, Gene-Wei Li, and X. Sunney Xie (May 2007). "Probing Transcription Factor Dynamics at the Single-Molecule Level in a Living Cell". en. In: *Science* 316.5828, pp. 1191–1194. ISSN: 0036-8075, 1095-9203.

England, Christopher G., Haiming Luo, and Weibo Cai (June 2015). "HaloTag Technology: A Versatile Platform for Biomedical Applications". In: *Bioconjugate Chem.* 26.6. Publisher: American Chemical Society, pp. 975–986. ISSN: 1043-1802.

Ernst, Jason et al. (Nov. 2016). "Genome-scale high-resolution mapping of activating and repressive nucleotides in regulatory regions". en. In: *Nat Biotechnol* 34.11, pp. 1180–1190. ISSN: 1087-0156, 1546-1696.

Esadze, Alexandre and Junji Iwahara (Jan. 2014). "Stopped-Flow Fluorescence Kinetic Study of Protein Sliding and Intersegment Transfer in the Target DNA Search Process". en. In: *Journal of Molecular Biology* 426.1, pp. 230–244. ISSN: 00222836.

Esadze, Alexandre and James T. Stivers (Dec. 2018). "Facilitated Diffusion Mechanisms in DNA Base Excision Repair and Transcriptional Activation". en. In: *Chem. Rev.* 118.23, pp. 11298–11323. ISSN: 0009-2665, 1520-6890.

Espinosa, Joaquin M. and Beverly M. Emerson (July 2001). "Transcriptional Regulation by p53 through Intrinsic DNA/Chromatin Binding and Site-Directed Cofactor Recruitment". English. In: *Molecular Cell* 8.1. Publisher: Elsevier, pp. 57–69. ISSN: 1097-2765.

Ezer, Daphne, Nicolae Radu Zabet, and Boris Adryan (July 2014). "Homotypic clusters of transcription factor binding sites: A model system for understanding the physical mechanics of gene expression". en. In: *Computational and Structural Biotechnology Journal* 10.17, pp. 63–69. ISSN: 20010370.

Fagagna, Fabrizio d'Adda di et al. (Nov. 2003). "A DNA damage checkpoint response in telomere-initiated senescence". en. In: *Nature* 426.6963, pp. 194–198. ISSN: 0028-0836, 1476-4687.

Faget, Douglas V., Qihao Ren, and Sheila A. Stewart (Aug. 2019). "Unmasking senescence: context-dependent effects of SASP in cancer". en. In: *Nat Rev Cancer* 19.8. Number: 8 Publisher: Nature Publishing Group, pp. 439–453. ISSN: 1474-1768.

Farley, Emma K. et al. (Oct. 2015). "Suboptimization of developmental enhancers". en. In: *Science* 350.6258, pp. 325–328. ISSN: 0036-8075, 1095-9203.

Fierz, Beat and Micheal G Poirier (2019). "Biophysics of Chromatin Dynamics". In: *Annu Rev Biophys* 48, pp. 321–345.

Fischer, M (July 2017). "Census and evaluation of p53 target genes". en. In: *Oncogene* 36.28, pp. 3943–3956. ISSN: 0950-9232, 1476-5594.

Fischer, Martin (May 2019). "Conservation and divergence of the p53 gene regulatory network between mice and humans". en. In: *Oncogene* 38.21, pp. 4095–4109. ISSN: 0950-9232, 1476-5594.

Franceschi, Claudio and Judith Campisi (June 2014). "Chronic inflammation (inflammaging) and its potential contribution to age-associated diseases". eng. In: *J Gerontol A Biol Sci Med Sci* 69 Suppl 1, S4–9. ISSN: 1758-535X.

Friedler, Assaf et al. (Apr. 2005). "Modulation of Binding of DNA to the C-Terminal Domain of p53 by Acetylation". en. In: *Structure* 13.4, pp. 629–636. ISSN: 0969-2126.

Fritsch, Christian C. and Jörg Langowski (Jan. 2011). "Chromosome dynamics, molecular crowding, and diffusion in the interphase cell nucleus: a Monte Carlo lattice simulation study". en. In: *Chromosome Res* 19.1, pp. 63–81. ISSN: 0967-3849, 1573-6849.

Fu, Yutao, Manisha Sinha, Craig L. Peterson, and Zhiping Weng (July 2008). "The Insulator Binding Protein CTCF Positions 20 Nucleosomes around Its Binding Sites across the Human Genome". en. In: *PLoS Genet* 4.7. Ed. by Bas van Steensel, e1000138. ISSN: 1553-7404.

Fumagalli, Marzia, Francesca Rossiello, Chiara Mondello, and Fabrizio d'Adda di Fagagna (Oct. 2014). "Stable Cellular Senescence Is Associated with Persistent DDR Activation". en. In: *PLOS ONE* 9.10. Publisher: Public Library of Science, e110969. ISSN: 1932-6203.

Fussner, Eden et al. (Nov. 2012). "Open and closed domains in the mouse genome are configured as 10-nm chromatin fibres". en. In: *EMBO Rep* 13.11, pp. 992–996. ISSN: 1469-221X, 1469-3178.

Gebhardt, J. Christof M. et al. (May 2013). "Single-molecule imaging of transcription factor binding to DNA in live mammalian cells". en. In: *Nat Methods* 10.5. Number: 5 Publisher: Nature Publishing Group, pp. 421–426. ISSN: 1548-7105.

Gillespie, Daniel T. (Dec. 1977). "Exact stochastic simulation of coupled chemical reactions". en. In: *J. Phys. Chem.* 81.25, pp. 2340–2361. ISSN: 0022-3654, 1541-5740.

Giresi, Paul G. et al. (June 2007). "FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin". en. In: *Genome Res.* 17.6, pp. 877–885. ISSN: 1088-9051.

Gorgoulis, Vassilis et al. (Oct. 2019). "Cellular Senescence: Defining a Path Forward". en. In: *Cell* 179.4, pp. 813–827. ISSN: 00928674.

Gorgoulis, Vassilis G and Thanos D Halazonetis (Dec. 2010). "Oncogene-induced senescence: the bright and dark side of the response". en. In: *Current Opinion in Cell Biology*. Cell differentiation / Cell division, growth and death 22.6, pp. 816–827. ISSN: 0955-0674.

Gorman, Jason et al. (Nov. 2007). "Dynamic Basis for One-Dimensional DNA Scanning by the Mismatch Repair Complex Msh2-Msh6". In: *Molecular Cell* 28.3, pp. 359–370. ISSN: 1097-2765.

Graha Subekti, Dwiky Rendra and Kiyoto Kamagata (Jan. 2021). "The disordered DNA-binding domain of p53 is indispensable for forming an encounter complex to and jumping along DNA". en. In: *Biochemical and Biophysical Research Communications* 534, pp. 21–26. ISSN: 0006-291X.

Granéli, Annette, Caitlyn C. Yeykal, Ragan B. Robertson, and Eric C. Greene (Jan. 2006). "Long-distance lateral diffusion of human Rad51 on double-stranded DNA". In: *Proceedings of the National Academy of Sciences* 103.5. Publisher: Proceedings of the National Academy of Sciences, pp. 1221–1226.

Guigas, Gernot, Claudia Kalla, and Matthias Weiss (Oct. 2007). "The degree of macromolecular crowding in the cytoplasm and nucleoplasm of mammalian cells is conserved". en. In: *FEBS Letters* 581.26, pp. 5094–5098. ISSN: 00145793.

Gunawardena, Jeremy (Apr. 2014). "Models in biology: 'accurate descriptions of our pathetic thinking'". In: *BMC Biology* 12.1, p. 29. ISSN: 1741-7007.

Hafner, Antonina, Martha L. Bulyk, Ashwini Jambhekar, and Galit Lahav (Apr. 2019). "The multiple mechanisms that regulate p53 activity and cell fate". en. In: *Nat Rev Mol Cell Biol* 20.4, pp. 199–210. ISSN: 1471-0072, 1471-0080.

Hafner, Antonina, José Reyes, et al. (June 2020). "Quantifying the Central Dogma in the p53 Pathway in Live Single Cells". en. In: *Cell Systems* 10.6, 495–505.e4. ISSN: 24054712.

Hamard, Pierre-Jacques, Dana J. Lukin, and James J. Manfredi (June 2012). "p53 Basic C Terminus Regulates p53 Functions through DNA Binding Modulation of Subset of Target Genes". en. In: *Journal of Biological Chemistry* 287.26, pp. 22397–22407. ISSN: 00219258.

Hammar, Petter et al. (June 2012). "The *lac* Repressor Displays Facilitated Diffusion in Living Cells". en. In: *Science* 336.6088, pp. 1595–1598. ISSN: 0036-8075, 1095-9203.

Hansen, Anders S. et al. (Mar. 2020). "Guided nuclear exploration increases CTCF target search efficiency". en. In: *Nat Chem Biol* 16.3. Number: 3 Publisher: Nature Publishing Group, pp. 257–266. ISSN: 1552-4469.

Henninger, Jonathan E. et al. (Jan. 2021). "RNA-Mediated Feedback Control of Transcriptional Condensates". en. In: *Cell* 184.1, 207–225.e24. ISSN: 00928674.

Hernandez-Segura, Alejandra, Tristan V. de Jong, et al. (Sept. 2017). "Unmasking Transcriptional Heterogeneity in Senescent Cells". en. In: *Current Biology* 27.17, 2652–2660.e4. ISSN: 09609822.

Hernandez-Segura, Alejandra, Jamil Nehme, and Marco Demaria (June 2018). "Hallmarks of Cellular Senescence". en. In: *Trends in Cell Biology* 28.6, pp. 436–453. ISSN: 09628924.

Hewitt, Graeme et al. (Feb. 2012). "Telomeres are favoured targets of a persistent DNA damage response in ageing and stress-induced senescence". In: *Nat Commun* 3, p. 708. ISSN: 2041-1723.

Hnisz, Denes et al. (Nov. 2013). "Super-Enhancers in the Control of Cell Identity and Disease". In: *Cell* 155.4, pp. 934–947. ISSN: 0092-8674.

Hoare, Matthew et al. (Sept. 2016). "NOTCH1 mediates a switch between two distinct secretomes during senescence". en. In: *Nat Cell Biol* 18.9. Number: 9 Publisher: Nature Publishing Group, pp. 979–992. ISSN: 1476-4679.

Hong, Liu and Jinzhi Lei (2009). "Scaling law for the radius of gyration of proteins and its dependence on hydrophobicity". en. In: *Journal of Polymer Science Part B: Polymer Physics* 47.2. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/polb.21634, pp. 207–214. ISSN: 1099-0488.

Huang, Fang et al. (Dec. 2009). "Multiple conformations of full-length p53 detected with single-molecule fluorescence resonance energy transfer". In: *Proceedings of the National Academy of Sciences* 106.49. Publisher: Proceedings of the National Academy of Sciences, pp. 20758–20763.

Huang, Weijun et al. (Oct. 2022). "Cellular senescence: the good, the bad and the unknown". en. In: *Nat Rev Nephrol* 18.10. Number: 10 Publisher: Nature Publishing Group, pp. 611–627. ISSN: 1759-507X.

Hug, Clemens B., Alexis G. Grimaldi, Kai Kruse, and Juan M. Vaquerizas (Apr. 2017). "Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription". en. In: *Cell* 169.2, 216–228.e19. ISSN: 00928674.

Hyle, Judith et al. (July 2019). "Acute depletion of CTCF directly affects MYC regulation through loss of enhancer–promoter looping". en. In: *Nucleic Acids Research* 47.13, pp. 6699–6713. ISSN: 0305-1048, 1362-4962.

Imbert, Arthur et al. (June 2022). "FISH-quant v2: a scalable and modular tool for smFISH image analysis". In: *RNA* 28.6, pp. 786–795. ISSN: 1355-8382.

Inga, Alberto, Paola Monti, et al. (Jan. 2001). "p53 mutants exhibiting enhanced transcriptional activation and altered promoter selectivity are revealed using a sensitive, yeast-based functional assay". en. In: *Oncogene* 20.4, pp. 501–513. ISSN: 0950-9232, 1476-5594.

Inga, Alberto, Francesca Storici, Thomas A. Darden, and Michael A. Resnick (Dec. 2002). "Differential Transactivation by the p53 Transcription Factor Is Highly Dependent on p53 Level and Promoter Target Sequence". en. In: *Molecular and Cellular Biology* 22.24, pp. 8612–8625. ISSN: 1098-5549.

Innes, Andrew J. and Jesús Gil (2019). "IMR90 ER:RAS: A Cell Model of Oncogene-Induced Senescence". eng. In: *Methods Mol Biol* 1896, pp. 83–92. ISSN: 1940-6029.

Isaacson, S. A., D. M. McQueen, and Charles S. Peskin (Mar. 2011). "The influence of volume exclusion by chromatin on the time required to find specific DNA binding sites by diffusion". en. In: *Proc. Natl. Acad. Sci. U.S.A.* 108.9, pp. 3815–3820. ISSN: 0027-8424, 1091-6490.

Isaacson, Samuel A. et al. (Nov. 2013). "The Influence of Spatial Variation in Chromatin Density Determined by X-Ray Tomograms on the Time to Find DNA Binding Sites". en. In: *Bull Math Biol* 75.11, pp. 2093–2117. ISSN: 0092-8240, 1522-9602.

Itoh, Yuji, Agato Murata, Satoshi Takahashi, and Kiyoto Kamagata (Aug. 2018). "Intrinsically disordered domain of tumor suppressor p53 facilitates target search by ultrafast transfer between different DNA strands". en. In: *Nucleic Acids Research* 46.14, pp. 7261–7269. ISSN: 0305-1048, 1362-4962.

Ivanov, Andre et al. (July 2013). "Lysosome-mediated processing of chromatin in senescence". eng. In: *J Cell Biol* 202.1, pp. 129–143. ISSN: 1540-8140.

Izeddin, Ignacio et al. (June 2014). "Single-molecule tracking in live cells reveals distinct target-search strategies of transcription factors in the nucleus". en. In: *eLife* 3, e02230. ISSN: 2050-084X.

Jacob, François and Jacques Monod (1961). "Genetic regulatory mechanisms in the synthesis of proteins". en. In: *J. Mol. Bio.*

Johmura, Yoshikazu et al. (July 2014). "Necessary and Sufficient Role for a Mitosis Skip in Senescence Induction". en. In: *Molecular Cell* 55.1, pp. 73–84. ISSN: 1097-2765.

Jun, Joon-Il and Lester F. Lau (July 2010). "The matricellular protein CCN1 induces fibroblast senescence and restricts fibrosis in cutaneous wound healing". en. In: *Nat Cell Biol* 12.7. Number: 7 Publisher: Nature Publishing Group, pp. 676–685. ISSN: 1476-4679.

Kaeser, M. D. and R. D. Iggo (Jan. 2002). "Chromatin immunoprecipitation analysis fails to support the latency model for regulation of p53 DNA binding activity *in vivo*". en. In: *Proc. Natl. Acad. Sci. U.S.A.* 99.1, pp. 95–100. ISSN: 0027-8424, 1091-6490.

Kamagata, Kiyoto et al. (Jan. 2020). "Liquid-like droplet formation by tumor suppressor p53 induced by multivalent electrostatic interactions between two disordered domains". en. In: *Sci Rep* 10.1, p. 580. ISSN: 2045-2322.

Kang, Chanhee et al. (Sept. 2015). "The DNA damage response induces inflammation and senescence by inhibiting autophagy of GATA4". In: *Science* 349.6255, aaa5612. ISSN: 0036-8075.

Kastenhuber, Edward R. and Scott W. Lowe (Sept. 2017). "Putting p53 in Context". en. In: *Cell* 170.6, pp. 1062–1078. ISSN: 00928674.

Kent, Samantha et al. (Oct. 2020). "Phase-Separated Transcriptional Condensates Accelerate Target-Search Process Revealed by Live-Cell Single-Molecule Imaging". en. In: *Cell Reports* 33.2, p. 108248. ISSN: 22111247.

Kilic, Sinan et al. (Aug. 2019). "Phase separation of 53BP1 determines liquid-like behavior of DNA repair compartments". In: *The EMBO Journal* 38.16. Publisher: John Wiley & Sons, Ltd, e101379. ISSN: 0261-4189.

Kim, Jiah et al. (Jan. 2019). "Nuclear speckle fusion via long-range directional motion regulates speckle morphology after transcriptional inhibition". en. In: *Journal of Cell Science*, jcs.226563. ISSN: 1477-9137, 0021-9533.

Kirschner, Kristina et al. (Mar. 2015). "Phenotype Specific Analyses Reveal Distinct Regulatory Mechanism for Chronically Activated p53". en. In: *PLOS Genetics* 11.3. Publisher: Public Library of Science, e1005053. ISSN: 1553-7404.

Kosar, Martin et al. (Feb. 2011). "Senescence-associated heterochromatin foci are dispensable for cellular senescence, occur in a cell type- and insult-dependent manner and follow expression of p16ink4a". In: *Cell Cycle* 10.3. Publisher: Taylor & Francis _eprint: https://doi.org/10.4161/cc.10.3.14707, pp. 457–468. ISSN: 1538-4101.

Kracikova, M et al. (Apr. 2013). "A threshold mechanism mediates p53 cell fate decision between growth arrest and apoptosis". en. In: *Cell Death Differ* 20.4, pp. 576–588. ISSN: 1350-9047, 1476-5403.

Kribelbauer, Judith F., Chaitanya Rastogi, Harmen J. Bussemaker, and Richard S. Mann (Oct. 2019). "Low-Affinity Binding Sites and the Transcription Factor Specificity Paradox in Eukaryotes". en. In: *Annu. Rev. Cell Dev. Biol.* 35.1, pp. 357–379. ISSN: 1081-0706, 1530-8995.

Kussie, Paul H. et al. (Nov. 1996). "Structure of the MDM2 Oncoprotein Bound to the p53 Tumor Suppressor Transactivation Domain". In: *Science* 274.5289. Publisher: American Association for the Advancement of Science, pp. 948–953.

Laberge, Remi-Martin et al. (Aug. 2015). "MTOR regulates the pro-tumorigenic senescence-associated secretory phenotype by promoting IL1A translation". en. In: *Nature cell biology* 17.8. Publisher: NIH Public Access, p. 1049.

Lambert, Samuel A. et al. (Feb. 2018). "The Human Transcription Factors". en. In: *Cell* 172.4, pp. 650–665. ISSN: 00928674.

Lämmerhirt, Lisa et al. (Jan. 2022). "Knockdown of Lamin B1 and the Corresponding Lamin B Receptor Leads to Changes in Heterochromatin State and Senescence Induction in Malignant Melanoma". en. In: *Cells* 11.14. Number: 14 Publisher: Multidisciplinary Digital Publishing Institute, p. 2154. ISSN: 2073-4409.

Lammers, Nicholas C., Yang Joon Kim, Jiaxi Zhao, and Hernan G. Garcia (Dec. 2020). "A matter of time: Using dynamics and theory to uncover mechanisms of transcriptional bursting". en. In: *Current Opinion in Cell Biology* 67, pp. 147–157. ISSN: 09550674.

Lane, David P. et al. (Dec. 2011). "Conservation of all three p53 family members and Mdm2 and Mdm4 in the cartilaginous fish". en. In: *Cell Cycle* 10.24, pp. 4272–4279. ISSN: 1538-4101, 1551-4005.

Langdon, Erin M. et al. (May 2018). “mRNA structure determines specificity of a polyQ-driven phase separation”. en. In: *Science* 360.6391, pp. 922–927. ISSN: 0036-8075, 1095-9203.

Laptenko, Oleg et al. (Mar. 2015). “The p53 C Terminus Controls Site-Specific DNA Binding and Promotes Structural Changes within the Central DNA Binding Domain”. en. In: *Molecular Cell* 57.6, pp. 1034–1046. ISSN: 10972765.

Lee, Kyung-Ha, Do-Yeon Kim, and Wanil Kim (Nov. 2021). “Regulation of Gene Expression by Telomere Position Effect”. In: *Int J Mol Sci* 22.23, p. 12807. ISSN: 1422-0067.

Lee, Tong Ihn and Richard A. Young (Mar. 2013). “Transcriptional Regulation and Its Misregulation in Disease”. en. In: *Cell* 152.6, pp. 1237–1251. ISSN: 00928674.

Lesne, Annick et al. (Nov. 2014). “3D genome reconstruction from chromosomal contacts”. en. In: *Nat Methods* 11.11, pp. 1141–1143. ISSN: 1548-7091, 1548-7105.

Li, H. et al. (Sept. 2014). “Integrated high-throughput analysis identifies Sp1 as a crucial determinant of p53-mediated apoptosis”. en. In: *Cell Death Differ* 21.9. Number: 9 Publisher: Nature Publishing Group, pp. 1493–1502. ISSN: 1476-5403.

Li, Jingyi Jessica, Peter J Bickel, and Mark D Biggin (Feb. 2014). “System wide analyses have underestimated protein abundances and the importance of transcription in mammals”. en. In: *PeerJ* 2, e270. ISSN: 2167-8359.

Li, Muyang, Jianyuan Luo, Christopher L. Brooks, and Wei Gu (Dec. 2002). “Acetylation of p53 Inhibits Its Ubiquitination by Mdm2 *”. English. In: *Journal of Biological Chemistry* 277.52. Publisher: Elsevier, pp. 50607–50611. ISSN: 0021-9258, 1083-351X.

Liou, Shu-Hao et al. (Mar. 2021). “Structure of the p53/RNA polymerase II assembly”. en. In: *Commun Biol* 4.1, p. 397. ISSN: 2399-3642.

Liu, Jie-Yu et al. (Feb. 2019). “Cells exhibiting strong p16INK4a promoter activation in vivo display features of senescence”. In: *Proceedings of the National Academy of Sciences* 116.7. Publisher: Proceedings of the National Academy of Sciences, pp. 2603–2611.

Loffreda, Alessia et al. (Aug. 2017). “Live-cell p53 single-molecule binding is modulated by C-terminal acetylation and correlates with transcriptional activity”. en. In: *Nat Commun* 8.1, p. 313. ISSN: 2041-1723.

Lujambio, Amaia et al. (Apr. 2013). “Non-cell-autonomous tumor suppression by p53”. In: *Cell* 153.2, pp. 449–460. ISSN: 0092-8674.

Luke-Glaser, Sarah, Heiko Poschke, and Brian Luke (2012). “Getting in (and out of) the loop: regulating higher order telomere structures”. In: *Frontiers in Oncology* 2. ISSN: 2234-943X.

Lupiáñez, Darío G. et al. (May 2015). “Disruptions of Topological Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer Interactions”. en. In: *Cell* 161.5, pp. 1012–1025. ISSN: 00928674.

Lyons, Heankel et al. (Jan. 2023). “Functional partitioning of transcriptional regulators by patterned charge blocks”. en. In: *Cell* 186.2, 327–345.e28. ISSN: 00928674.

Madsen, Jesper G. S. et al. (Nov. 2020). “Highly interconnected enhancer communities control lineage-determining genes in human mesenchymal stem cells”. en. In: *Nat Genet* 52.11, pp. 1227–1238. ISSN: 1061-4036, 1546-1718.

Maeshima, Kazuhiro, Kazunari Kaizu, et al. (Feb. 2015). “The physical size of transcription factors is key to transcriptional regulation in chromatin domains”. en. In: *J. Phys.: Condens. Matter* 27.6, p. 064116. ISSN: 0953-8984, 1361-648X.

Maeshima, Kazuhiro, Ryan Rogge, et al. (May 2016). “Nucleosomal arrays self-assemble into supramolecular globular structures lacking 30-nm fibers”. en. In: *EMBO J* 35.10, pp. 1115–1132. ISSN: 0261-4189, 1460-2075.

Malkin, David et al. (Nov. 1990). "Germ Line p53 Mutations in a Familial Syndrome of Breast Cancer, Sarcomas, and Other Neoplasms". en. In: *Science* 250.4985, pp. 1233–1238. ISSN: 0036-8075, 1095-9203.

Marklund, Emil et al. (Jan. 2022). "Sequence specificity in DNA binding is mainly governed by association". en. In: *Science* 375.6579, pp. 442–445. ISSN: 0036-8075, 1095-9203.

Marney, Christina B. et al. (June 2021). "p53-intact cancers escape tumor suppression through loss of long noncoding RNA Dino". en. In: *Cell Reports* 35.13, p. 109329. ISSN: 22111247.

Martin, Erik W. et al. (Feb. 2020). "Valence and patterning of aromatic residues determine the phase behavior of prion-like domains". en. In: *Science* 367.6478, pp. 694–699. ISSN: 0036-8075, 1095-9203.

Mazza, Davide et al. (Aug. 2012). "A benchmark for chromatin binding measurements in live cells". en. In: *Nucleic Acids Research* 40.15, e119–e119. ISSN: 1362-4962, 0305-1048.

Mazzocca, Matteo, Tom Fillot, et al. (June 2021). "The needle and the haystack: single molecule tracking to probe the transcription factor search in eukaryotes". en. In: *Biochemical Society Transactions* 49.3, pp. 1121–1132. ISSN: 0300-5127, 1470-8752.

Mazzocca, Matteo, Alessia Loffreda, et al. (Dec. 2022). *Chromatin organization drives the exploration strategy of nuclear factors*. en. preprint. Biophysics.

McSwiggen, David Trombley et al. (May 2019). "Evidence for DNA-mediated nuclear compartmentalization distinct from phase separation". en. In: *eLife* 8, e47098. ISSN: 2050-084X.

Mehta, Gunjan D. et al. (Dec. 2018). "Single-Molecule Analysis Reveals Linked Cycles of RSC Chromatin Remodeling and Ace1p Transcription Factor Binding in Yeast". en. In: *Molecular Cell* 72.5, 875–887.e9. ISSN: 10972765.

Melo, Carlos A. et al. (Feb. 2013). "eRNAs Are Required for p53-Dependent Enhancer Activity and Gene Transcription". English. In: *Molecular Cell* 49.3. Publisher: Elsevier, pp. 524–535. ISSN: 1097-2765.

Mendoza, Alex de et al. (Dec. 2013). "Transcription factor evolution in eukaryotes and the assembly of the regulatory toolkit in multicellular lineages". en. In: *Proc. Natl. Acad. Sci. U.S.A.* 110.50. ISSN: 0027-8424, 1091-6490.

Michael, Alicia K. and Nicolas H. Thomä (July 2021). "Reading the chromatinized genome". en. In: *Cell* 184.14, pp. 3599–3611. ISSN: 00928674.

Michaloglou, Chrysiis et al. (Aug. 2005). "BRAFE600-associated senescence-like cell cycle arrest of human naevi". eng. In: *Nature* 436.7051, pp. 720–724. ISSN: 1476-4687.

Miron, Ezequiel et al. (2020). "Chromatin arranges in chains of mesoscale domains with nanoscale functional topography independent of cohesin". In: *Sci. Adv.*

Misteli, Tom (Oct. 2020). "The Self-Organizing Genome: Principles of Genome Architecture and Function". en. In: *Cell* 183.1, pp. 28–45. ISSN: 00928674.

Mitchell, Pamela J. and Robert Tjian (July 1989). "Transcriptional Regulation in Mammalian Cells by Sequence-Specific DNA Binding Proteins". en. In: *Science* 245.4916, pp. 371–378. ISSN: 0036-8075, 1095-9203.

Muñoz, Denise P. et al. (2019). "Targetable mechanisms driving immunoevasion of persistent senescent cells link chemotherapy-resistant cancer to aging". In: *JCI Insight* 4.14, e124716. ISSN: 2379-3708.

Muñoz-Espín, Daniel, Marta Cañamero, et al. (Nov. 2013). "Programmed Cell Senescence during Mammalian Embryonic Development". en. In: *Cell* 155.5, pp. 1104–1118. ISSN: 0092-8674.

Muñoz-Espín, Daniel and Manuel Serrano (July 2014). "Cellular senescence: from physiology to pathology". en. In: *Nat Rev Mol Cell Biol* 15.7. Number: 7 Publisher: Nature Publishing Group, pp. 482–496. ISSN: 1471-0080.

Murakami, Yu et al. (2020). "CRISPR/Cas9 nickase-mediated efficient and seamless knock-in of lethal genes in the medaka fish Oryzias latipes". en. In: *Development, Growth & Differentiation* 62.9. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/dgd.12700, pp. 554–567. ISSN: 1440-169X.

Narita, Masashi et al. (June 2003). "Rb-Mediated Heterochromatin Formation and Silencing of E2F Target Genes during Cellular Senescence". en. In: *Cell* 113.6, pp. 703–716. ISSN: 0092-8674.

Nguyen, Thuy-Ai T et al. (Sept. 2018). "Revealing a human p53 universe". en. In: *Nucleic Acids Research* 46.16, pp. 8153–8167. ISSN: 0305-1048, 1362-4962.

Nguyen, Vu Q. et al. (Sept. 2021). "Spatiotemporal coordination of transcription preinitiation complex assembly in live cells". en. In: *Molecular Cell* 81.17, 3560–3575.e6. ISSN: 10972765.

Nitta, Kazuhiro R et al. (Mar. 2015). "Conservation of transcription factor binding specificities across 600 million years of bilateria evolution". en. In: *eLife* 4, e04837. ISSN: 2050-084X.

Nora, Elphège P. et al. (May 2012). "Spatial partitioning of the regulatory landscape of the X-inactivation centre". en. In: *Nature* 485.7398, pp. 381–385. ISSN: 0028-0836, 1476-4687.

Nott, Timothy J. et al. (Mar. 2015). "Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles". en. In: *Molecular Cell* 57.5, pp. 936–947. ISSN: 10972765.

Nozaki, Tadasu et al. (July 2017). "Dynamic Organization of Chromatin Domains Revealed by Super-Resolution Live-Cell Imaging". en. In: *Molecular Cell* 67.2, 282–293.e7. ISSN: 10972765.

Oh, Jaehak et al. (May 2008). "BAF60a Interacts with p53 to Recruit the SWI/SNF Complex". en. In: *Journal of Biological Chemistry* 283.18, pp. 11924–11934. ISSN: 00219258.

Ohanna, Mickaël et al. (June 2011). "Senescent cells develop a PARP-1 and nuclear factor-κB-associated secretome (PNAS)". In: *Genes Dev* 25.12, pp. 1245–1261. ISSN: 0890-9369.

Ou, Horng D. et al. (July 2017). "ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells". en. In: *Science* 357.6349, eaag0025. ISSN: 0036-8075, 1095-9203.

Pachitariu, Marius and Carsen Stringer (Dec. 2022). "Cellpose 2.0: how to train your own model". en. In: *Nat Methods* 19.12. Number: 12 Publisher: Nature Publishing Group, pp. 1634–1641. ISSN: 1548-7105.

Panne, Daniel (Apr. 2008). "The enhanceosome". en. In: *Current Opinion in Structural Biology* 18.2, pp. 236–242. ISSN: 0959440X.

Passos, João F et al. (Feb. 2010). "Feedback between p21 and reactive oxygen production is necessary for cell senescence". In: *Mol Syst Biol* 6, p. 347. ISSN: 1744-4292.

Pederson, Thoru (May 2000). "Diffusional protein transport within the nucleus: a message in the medium". en. In: *Nat Cell Biol* 2.5, E73–E74. ISSN: 1465-7392, 1476-4679.

Persson, Fredrik, Martin Lindén, Cecilia Unoson, and Johan Elf (Mar. 2013). "Extracting intracellular diffusive states and transition rates from single-molecule tracking data". en. In: *Nat Methods* 10.3, pp. 265–269. ISSN: 1548-7091, 1548-7105.

Pott, Sebastian and Jason D. Lieb (Jan. 2015). "What are super-enhancers?" en. In: *Nat Genet* 47.1. Number: 1 Publisher: Nature Publishing Group, pp. 8–12. ISSN: 1546-1718.

Raccaud, Mahé and David M. Suter (2018). "Transcription factor retention on mitotic chromosomes: regulatory mechanisms and impact on cell fate decisions". en. In: *FEBS Letters* 592.6. _eprint: https://febs.onlinelibrary.wiley.com/doi/pdf/10.1002/1873-3468.12828, pp. 878–887. ISSN: 1873-3468.

Ran, F. Ann et al. (Sept. 2013). "Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity". en. In: *Cell* 154.6. Publisher: NIH Public Access, p. 1380.

Reiter, Franziska, Sebastian Wienerroither, and Alexander Stark (Apr. 2017). "Combinatorial function of transcription factors and cofactors". en. In: *Current Opinion in Genetics & Development* 43, pp. 73–81. ISSN: 0959437X.

Ricci, Maria Aurelia et al. (Mar. 2015). "Chromatin Fibers Are Formed by Heterogeneous Groups of Nucleosomes In Vivo". en. In: *Cell* 160.6, pp. 1145–1158. ISSN: 00928674.

Ries, Ryan J. et al. (July 2019). "m6A enhances the phase separation potential of mRNA". en. In: *Nature* 571.7765, pp. 424–428. ISSN: 0028-0836, 1476-4687.

Roden, Christine and Amy S. Gladfelter (Mar. 2021). "RNA contributions to the form and function of biomolecular condensates". en. In: *Nat Rev Mol Cell Biol* 22.3, pp. 183–195. ISSN: 1471-0072, 1471-0080.

Rodier, Francis et al. (Jan. 2011). "DNA-SCARS: distinct nuclear structures that sustain damage-induced senescence growth arrest and inflammatory cytokine secretion". In: *J Cell Sci* 124.1, pp. 68–81. ISSN: 0021-9533.

Sabari, Benjamin R. et al. (July 2018). "Coactivator condensation at super-enhancers links phase separation and gene control". en. In: *Science* 361.6400, eaar3958. ISSN: 0036-8075, 1095-9203.

Sakaguchi, Kazuyasu et al. (Sept. 1998). "DNA damage activates p53 through a phosphorylation–acetylation cascade". In: *Genes Dev* 12.18, pp. 2831–2841. ISSN: 0890-9369.

Sako, Yasushi, Shigeru Minoghchi, and Toshio Yanagida (Mar. 2000). "Single-molecule imaging of EGFR signalling on the surface of living cells". en. In: *Nat Cell Biol* 2.3. Number: 3 Publisher: Nature Publishing Group, pp. 168–172. ISSN: 1476-4679.

Salama, Rafik, Mahito Sadaie, Matthew Hoare, and Masashi Narita (Jan. 2014). "Cellular senescence and its effector programs". en. In: *Genes Dev.* 28.2. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab, pp. 99–114. ISSN: 0890-9369, 1549-5477.

Saleh, Tareq and David A. Gewirtz (June 2022). "Considering therapy-induced senescence as a mechanism of tumour dormancy contributing to disease recurrence". en. In: *Br J Cancer* 126.10. Number: 10 Publisher: Nature Publishing Group, pp. 1363–1365. ISSN: 1532-1827.

Sammons, Morgan A, Thuy-Ai T Nguyen, Simon S McDade, and Martin Fischer (Sept. 2020). "Tumor suppressor p53: from engaging DNA to target gene regulation". en. In: *Nucleic Acids Research* 48.16, pp. 8848–8869. ISSN: 0305-1048, 1362-4962.

Sammons, Morgan A., Jiajun Zhu, Adam M. Drake, and Shelley L. Berger (Feb. 2015). "TP53 engagement with the genome occurs in distinct local chromatin environments via pioneer factor activity". en. In: *Genome Res.* 25.2, pp. 179–188. ISSN: 1088-9051, 1549-5469.

Sati, Satish et al. (May 2020). "4D Genome Rewiring during Oncogene-Induced and Replicative Senescence". en. In: *Molecular Cell* 78.3, 522–538.e9. ISSN: 10972765.

Schmitt, Clemens A., Boshi Wang, and Marco Demaria (Oct. 2022). "Senescence and cancer — role and therapeutic opportunities". en. In: *Nat Rev Clin Oncol* 19.10. Number: 10 Publisher: Nature Publishing Group, pp. 619–636. ISSN: 1759-4782.

Schosserer, Markus, Johannes Grillari, and Michael Breitenbach (2017). "The Dual Role of Cellular Senescence in Developing Tumors and Their Response to Cancer Therapy". In: *Frontiers in Oncology* 7. ISSN: 2234-943X.

Serrano, M. et al. (Mar. 1997). "Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a". eng. In: *Cell* 88.5, pp. 593–602. ISSN: 0092-8674.

Sexton, Tom et al. (Feb. 2012). "Three-Dimensional Folding and Functional Organization Principles of the Drosophila Genome". en. In: *Cell* 148.3, pp. 458–472. ISSN: 00928674.

Shah, Parisha P. et al. (Aug. 2013). "Lamin B1 depletion in senescent cells triggers large-scale changes in gene expression and the chromatin landscape". eng. In: *Genes Dev* 27.16, pp. 1787–1799. ISSN: 1549-5477.

Sheekey, Eleanor and Masashi Narita (Mar. 2023). "p53 in senescence – it's a marathon, not a sprint". en. In: *The FEBS Journal* 290.5, pp. 1212–1220. ISSN: 1742-464X, 1742-4658.

Shi, Xiaobing et al. (Aug. 2007). "Modulation of p53 Function by SET8-Mediated Methylation at Lysine 382". English. In: *Molecular Cell* 27.4. Publisher: Elsevier, pp. 636–646. ISSN: 1097-2765.

Silverstein, Timothy D., Bryan Gibb, and Eric C. Greene (Aug. 2014). "Visualizing protein movement on DNA at the single-molecule level using DNA curtains". en. In: *DNA Repair* 20, pp. 94–109. ISSN: 15687864.

Smeenk, Leonie et al. (Mar. 2011). "Role of p53 Serine 46 in p53 Target Gene Regulation". en. In: *PLoS ONE* 6.3. Ed. by Andrei Gartel, e17574. ISSN: 1932-6203.

Spitz, François and Eileen E. M. Furlong (Sept. 2012). "Transcription factors: from enhancer binding to developmental control". en. In: *Nat Rev Genet* 13.9, pp. 613–626. ISSN: 1471-0056, 1471-0064.

Spolar, Ruth S. and M. Thomas Record (Feb. 1994). "Coupling of Local Folding to Site-Specific Binding of Proteins to DNA". en. In: *Science* 263.5148, pp. 777–784. ISSN: 0036-8075, 1095-9203.

Steensel, Bas van and Andrew S. Belmont (May 2017). "Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression". English. In: *Cell* 169.5. Publisher: Elsevier, pp. 780–791. ISSN: 0092-8674, 1097-4172.

Sternberg, Samuel H. et al. (Mar. 2014). "DNA interrogation by the CRISPR RNA-guided endonuclease Cas9". en. In: *Nature* 507.7490. Number: 7490 Publisher: Nature Publishing Group, pp. 62–67. ISSN: 1476-4687.

Sturmlechner, Ines, Chance C. Sine, et al. (June 2022). "Senescent cells limit p53 activity via multiple mechanisms to remain viable". en. In: *Nat Commun* 13.1. Number: 1 Publisher: Nature Publishing Group, p. 3722. ISSN: 2041-1723.

Sturmlechner, Ines, Cheng Zhang, et al. (Oct. 2021). "p21 produces a bioactive secretome that places stressed cells under immunosurveillance". en. In: *Science* 374.6567, eabb3420. ISSN: 0036-8075, 1095-9203.

Su, Dan et al. (Jan. 2015). "Interactions of Chromatin Context, Binding Site Sequence Content, and Sequence Evolution in Stress-Induced p53 Occupancy and Transactivation". en. In: *PLoS Genet* 11.1. Ed. by Vivian G. Cheung, e1004885. ISSN: 1553-7404.

Sung, Myong-Hee, Michael J. Guertin, Songjoon Baek, and Gordon L. Hager (Oct. 2014). "DNase Footprint Signatures Are Dictated by Factor Dynamics and DNA Sequence". en. In: *Molecular Cell* 56.2, pp. 275–285. ISSN: 10972765.

Szabo, Quentin et al. (Feb. 2018). "TADs are 3D structural units of higher-order chromosome organization in *Drosophila*". en. In: *Sci. Adv.* 4.2, eaar8082. ISSN: 2375-2548.

Tafvizi, Anahita et al. (Jan. 2011). "A single-molecule characterization of p53 search on DNA". en. In: *Proc. Natl. Acad. Sci. U.S.A.* 108.2, pp. 563–568. ISSN: 0027-8424, 1091-6490.

Takahashi, Kazutoshi and Shinya Yamanaka (Mar. 2016). "A decade of transcription factor-mediated reprogramming to pluripotency". en. In: *Nat Rev Mol Cell Biol* 17.3, pp. 183–193. ISSN: 1471-0072, 1471-0080.

Tang, Zhanyun et al. (July 2013). "SET1 and p300 Act Synergistically, through Coupled Histone Modifications, in Transcriptional Activation by p53". In: *Cell* 154.2, pp. 297–310. ISSN: 0092-8674.

Teves, Sheila S et al. (Nov. 2016). "A dynamic mode of mitotic bookmarking by transcription factors". In: *eLife* 5. Ed. by Karen Adelman. Publisher: eLife Sciences Publications, Ltd, e22280. ISSN: 2050-084X.

Thurman, Robert E. et al. (Sept. 2012). "The accessible chromatin landscape of the human genome". en. In: *Nature* 489.7414, pp. 75–82. ISSN: 0028-0836, 1476-4687.

Tinevez, Jean-Yves et al. (Feb. 2017). "TrackMate: An open and extensible platform for single-particle tracking". en. In: *Methods*. Image Processing for Biologists 115, pp. 80–90. ISSN: 1046-2023.

Tremethick, David J. (Feb. 2007). "Higher-Order Structures of Chromatin: The Elusive 30 nm Fiber". en. In: *Cell* 128.4, pp. 651–654. ISSN: 00928674.

Tu, Zhigang, Xinying Zhuang, Yong-Gang Yao, and Rugang Zhang (May 2013). "BRG1 Is Required for Formation of Senescence-Associated Heterochromatin Foci Induced by Oncogenic RAS or BRCA1 Loss". In: *Molecular and Cellular Biology* 33.9. Publisher: Taylor & Francis _eprint: https://doi.org/10.1128/MCB.01744-12, pp. 1819–1829. ISSN: null.

Uehara, Ikuno and Nobuyuki Tanaka (June 2018). "Role of p53 in the Regulation of the Inflammatory Tumor Microenvironment and Tumor Suppression". In: *Cancers (Basel)* 10.7, p. 219. ISSN: 2072-6694.

Van Deursen, Jan M. (May 2014). "The role of senescent cells in ageing". en. In: *Nature* 509.7501. Number: 7501 Publisher: Nature Publishing Group, pp. 439–446. ISSN: 1476-4687.

Veprintsev, Dmitry B. et al. (Feb. 2006). "Core domain interactions in full-length p53 in solution". en. In: *Proc. Natl. Acad. Sci. U.S.A.* 103.7, pp. 2115–2119. ISSN: 0027-8424, 1091-6490.

Verfaillie, Annelien et al. (July 2016). "Multiplex enhancer-reporter assays uncover unsophisticated TP53 enhancer logic". en. In: *Genome Res.* 26.7, pp. 882–895. ISSN: 1088-9051, 1549-5469.

Vernon, Robert M and Julie D Forman-Kay (Oct. 2019). "First-generation predictors of biological protein phase separation". en. In: *Current Opinion in Structural Biology* 58, pp. 88–96. ISSN: 0959440X.

Verschure, Pernette J et al. (Sept. 2003). "Condensed chromatin domains in the mammalian nucleus are accessible to large macromolecules". en. In: *EMBO Rep* 4.9, pp. 861–866. ISSN: 1469-221X, 1469-3178.

Vitre, Benjamin D and Don W Cleveland (Dec. 2012). "Centrosomes, chromosome instability (CIN) and aneuploidy". en. In: *Current Opinion in Cell Biology* 24.6, pp. 809–815. ISSN: 09550674.

Walker, Kristen K. and Arnold J. Levine (Dec. 1996). "Identification of a novel p53 functional domain that is necessary for efficient growth suppression". en. In: *Proc. Natl. Acad. Sci. U.S.A.* 93.26, pp. 15335–15340. ISSN: 0027-8424, 1091-6490.

Wasserman, Wyeth W. and Albin Sandelin (Apr. 2004). "Applied bioinformatics for the identification of regulatory elements". en. In: *Nat Rev Genet* 5.4, pp. 276–287. ISSN: 1471-0056, 1471-0064.

Weidemüller, Paula, Maksim Kholmatov, Evangelia Petsalaki, and Judith B. Zaugg (Dec. 2021). "Transcription factors: Bridge between cell signaling and gene regulation". en. In: *Proteomics* 21.23-24, p. 2000034. ISSN: 1615-9853, 1615-9861.

Weirauch, Matthew T. et al. (Sept. 2014). "Determination and Inference of Eukaryotic Transcription Factor Sequence Specificity". en. In: *Cell* 158.6, pp. 1431–1443. ISSN: 00928674.

West, Lisandra E. et al. (Nov. 2010). "The MBT Repeats of L3MBTL1 Link SET8-mediated p53 Methylation at Lysine 382 to Target Gene Repression". In: *J Biol Chem* 285.48, pp. 37725–37732. ISSN: 0021-9258.

Westfall, Matthew D., Deborah J. Mays, Joseph C. Sniezek, and Jennifer A. Pietenpol (Apr. 2003). "The ΔNp63α Phosphoprotein Binds the p21 and 14-3-3σ Promoters In Vivo and Has Transcriptional Repressor Activity That Is Reduced by Hay-Wells Syndrome-Derived Mutations". In: *Molecular and Cellular Biology* 23.7. Publisher: Taylor & Francis _eprint: https://doi.org/10.1128/MCB.23.7.2264-2276.2003, pp. 2264–2276. ISSN: null.

Wiley, Christopher D. et al. (2017). "Analysis of individual cells identifies cell-to-cell variability following induction of cellular senescence". en. In: *Aging Cell* 16.5. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/acel.12632, pp. 1043–1050. ISSN: 1474-9726.

Willeit, Peter et al. (July 2010). "Telomere Length and Risk of Incident Cancer and Cancer Mortality". In: *JAMA* 304.1, pp. 69–75. ISSN: 0098-7484.

Woodstock, Dana L., Morgan A. Sammons, and Martin Fischer (July 2021). "p63 and p53: Collaborative Partners or Dueling Rivals?" en. In: *Front. Cell Dev. Biol.* 9, p. 701986. ISSN: 2296-634X.

Woringer, Maxime and Xavier Darzacq (Aug. 2018). "Protein motion in the nucleus: from anomalous diffusion to weak interactions". en. In: *Biochemical Society Transactions* 46.4, pp. 945–956. ISSN: 0300-5127, 1470-8752.

Wunderlich, Zeba and Leonid A. Mirny (Oct. 2009). "Different gene regulation strategies revealed by analysis of binding motifs". en. In: *Trends in Genetics* 25.10, pp. 434–440. ISSN: 01689525.

Xu, Caiyue et al. (Oct. 2020). "SIRT1 is downregulated by autophagy in senescence and aging". In: *Nat Cell Biol* 22.10, pp. 1170–1179. ISSN: 1465-7392.

Xue, Bin, Celeste J. Brown, A. Keith Dunker, and Vladimir N. Uversky (Apr. 2013). "Intrinsically disordered regions of p53 family are highly diversified in evolution". en. In: *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 1834.4, pp. 725–738. ISSN: 15709639.

Xue, Wen et al. (Feb. 2007). "Senescence and tumour clearance is triggered by p53 restoration in murine liver carcinomas". en. In: *Nature* 445.7128. Number: 7128 Publisher: Nature Publishing Group, pp. 656–660. ISSN: 1476-4687.

Yosef, Reut et al. (Aug. 2017). "p21 maintains senescent cell viability under persistent DNA damage response by restraining JNK and caspase signaling". In: *The EMBO Journal* 36.15. Publisher: John Wiley & Sons, Ltd, pp. 2280–2295. ISSN: 0261-4189.

Younger, Scott T. and John L. Rinn (Sept. 2017). "p53 regulates enhancer accessibility and activity in response to DNA damage". en. In: *Nucleic Acids Research* 45.17, pp. 9889–9900. ISSN: 0305-1048, 1362-4962.

Yu, Xinyang and Michael J. Buck (Jan. 2019). "Defining TP53 pioneering capabilities with competitive nucleosome binding assays". In: *Genome Res* 29.1, pp. 107–115. ISSN: 1088-9051.

Zandarashvili, Levani et al. (June 2012). "Asymmetrical roles of zinc fingers in dynamic DNA-scanning process by the inducible transcription factor Egr-1". en. In: *Proc. Natl. Acad. Sci. U.S.A.* 109.26. ISSN: 0027-8424, 1091-6490.

Zhang, Rugang et al. (Jan. 2005). "Formation of MacroH2A-Containing Senescence-Associated Heterochromatin Foci and Senescence Driven by ASF1a and HIRA". en. In: *Developmental Cell* 8.1, pp. 19–30. ISSN: 1534-5807.

Zhao, Yuanxiang and S.Steven Potter (Apr. 2002). "Functional Comparison of the Hoxa 4, Hoxa 10, and Hoxa 11 Homeoboxes". en. In: *Developmental Biology* 244.1, pp. 21–36. ISSN: 00121606.

Zhu, Haoran et al. (Apr. 2020). "Oncogene-induced senescence: From biology to therapy". eng. In: *Mech Ageing Dev* 187, p. 111229. ISSN: 1872-6216.

Annex 1: Derivation of the diffusion model

Derivation of the extension of volume exclusion with non-specific binding.

Allowing the ability of binding and unbinding leads to a new master equation as follows:

$$\begin{cases} \frac{dP_i^u}{dt} &= \sum_j \left[\alpha_{i\leftarrow j} P_j^u - \alpha_{j\leftarrow i} P_i^u\right] - k_i^{on} P_i^u + k_i^{off} P_i^b \\ \frac{dP_i^b}{dt} &= k_i^{on} P_i^u - k_i^{off} P_i^b \end{cases} \tag{8.2}$$

with k_i^{on} the binding rate at pixel i, and k_i^{off}. Note that we let k_i^{on} vary with the local chromatin signal, accounting for the higher number of 'binding site' in those volumes, however k_i^{off} remains the same, thus the index may be dropped. We can start to derive the steady-state by finding whan the null derivative entails:

$$\frac{dP_i^b}{dt} = 0 \Rightarrow P_i^b = \frac{k_i^{on}}{k_i^{off}} P_i^u \tag{8.3}$$

This can be plugged in the equation for the unbound state and cancel out the last two terms relating to the binding and unbinding, leaving us with:

$$P_i^u \propto e^{-RI_i} \tag{8.4}$$

$$P_i^b \propto \frac{k_i^{on}}{k^{off}} e^{-RI_i} \tag{8.5}$$

For which we are missing the normalization constant. We call this constant N, such that we now have the equation :

$$P_i^u = \frac{1}{N} e^{-RI_i} \tag{8.6}$$

$$P_i^b = \frac{1}{N} \frac{k_i^{on}}{k^{off}} e^{-RI_i} \tag{8.7}$$

The normalization constant can be found by summing the probability

$$N = \sum_i P_i^u + \sum_i P_i^b = \sum_i e^{-RI_i} + \sum_i \frac{k_i^{on}}{k^{off}} e^{-RI_i} \tag{8.8}$$

$$N = \sum_i (1 + K_D I_i) e^{-RI_i} \tag{8.9}$$

With $K_D = frack^{on}k^{off}$ and $k^{on} = k_i^{on} * I_i$, separating a constant value of k^{on} from the local signal of chromatin intensity. With this, we can then rewrite the expression of P_i^u and P_i^b :

$$P_i^u = \frac{e^{-RI_i}}{\sum_i (1 + K_{on} I_i) e^{-RI_i}} \tag{8.10}$$

$$P_i^b = \frac{K_{on} I_i \, e^{-RI_i}}{\sum_i (1 + K_{on} I_i) e^{-RI_i}} \tag{8.11}$$

We wish to express our final steady-state probability in relationship to the fraction of molecules in the bound and unbound states. For this we can write the unbound fraction F_u (fraction of molecules freely diffusing) as :

$$F_u = \frac{\sum_i P_i^u}{\sum_i P_i^u + \sum_i P_i^b} = \frac{\sum_i e^{-RI_i}}{\sum_i (1 + K_D I_i) e^{-RI_i}} \tag{8.12}$$

We can rearrange this equation to express K_D in terms of F_u:

$$K_D = \frac{1 - F_u}{I_i \, F_u} \tag{8.13}$$

With this, we can replace K_D in the expression of P_i^u and P_i^b in 8.11, and we arrive at:

$$\begin{cases} P_i^u &= F_u \frac{e^{-RI_i}}{\sum_i e^{-RI_i}} \\ P_i^b &= (1 - F_u) \frac{I_i}{\langle I_i \rangle} \frac{e^{-RI_i}}{\sum_i e^{-RI_i}} \end{cases} \tag{8.14}$$

where $\langle I_i \rangle = \frac{\sum_i I_i e^{-RI_i}}{\sum_i e^{-RI_i}}$ the average signal of the domain. Finally, with the probabilities, we can trivially write the log-likelihood of observing N particles, with n_i particle located within pixel i, and we arrive at

$$logL = \sum_i n_i log \left[(F_u + (1 - F_u) \frac{e^{-RI_i}}{\sum_i e^{-RI_i}}) \frac{I_i}{\langle I_i \rangle} \right] \tag{8.15}$$

Depending on the objective, this can be rewritten with the (non-specifically) bound fraction in mind easily with $F_u = 1 - F_b$

Annex 2: Supplementary Material

Figure S1 relates to the histogram fitting of fig.4.6.

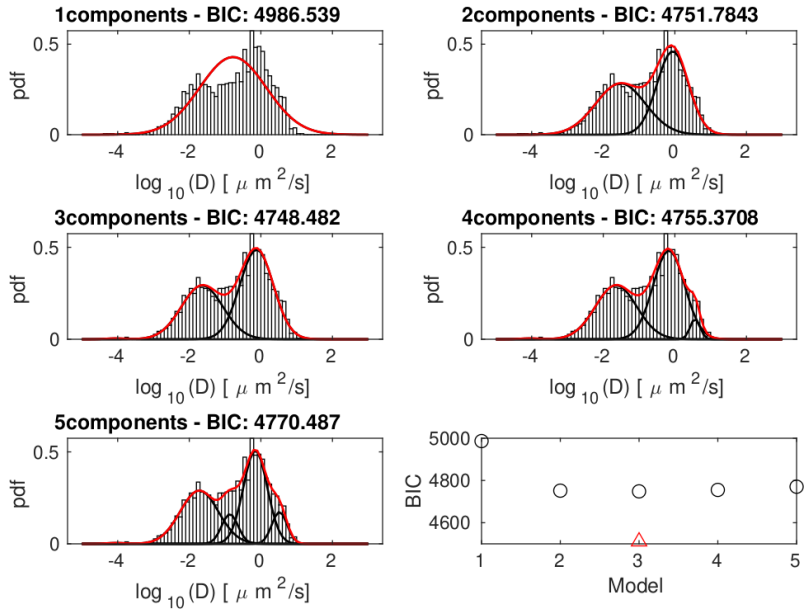

Fig. S1 **Model selection after tracking p53-HaloTag in BJ cells. Using the tracks obtained in the experiments, we compute the Mean Square Displacement, fit it to derive the coefficient of displacement, and populate an histogram with it. We used Matlab (Mathworks, Natick, MA, USA) to implement the models. To fit them to the p53 data, we employed custom routines based on the 'lsqnonlin' function, conducting a non-linear least squares minimization. The BIC is computed for each model of increasing size, and we choose the number of sub-population which minimizes it.**

The following tables report the experimental data used for the heatmaps presented in figures 4.14, 5.7 and 6.4.

chromatin class D	1	2	3	4	5	6	7
3.162278	1489	1614	1755	1690	1980	1926	2323
1	3513	4260	4363	4254	5014	4930	6018
0.316228	10170	13481	14028	14788	16851	17165	20184
0.1	15265	19512	21125	22962	25954	28938	35152
0.031623	11400	13399	14127	16579	18381	21926	27518
0.01	5392	5951	5992	7268	8160	9896	13195
0.003162	1974	2074	2256	2758	2917	3505	4747
0.001	935	1017	1051	1281	1446	1745	2372

Table 8.1 **Raw data for the dynamical enrichment of H2B as measured by SMT**

chromatin class D	1	2	3	4	5	6	7
3.162278	5421	5937	5622	5655	5259	4941	4045
1	5164	5538	5682	5610	5205	4961	4196
0.316228	4912	5639	5779	5600	5418	5153	4400
0.1	4616	5284	5327	4999	4869	5173	4550
0.031623	2863	3188	3109	2863	2913	3122	3036
0.01	1266	1309	1377	1215	1189	1398	1376
0.003162	507	494	434	416	380	490	495
0.001	234	211	244	189	191	225	215

Table 8.2 **Raw data for the dynamical enrichment of p53 as measured by SMT**

chromatin class D	1	2	3	4	5	6	7
3.162278	519	525	495	464	432	405	318
1	657	588	469	399	383	339	279
0.316228	754	774	514	474	462	355	335
0.1	691	729	500	490	499	493	364
0.031623	394	465	297	299	343	299	259
0.01	139	216	131	139	129	149	90
0.003162	57	73	75	42	41	89	42
0.001	22	43	24	19	19	28	18

Table 8.3 **Raw data for the dynamical enrichment of HaloTag as measured by SMT**

chromatin class D	1	2	3	4	5	6	7
3.162278	4671	4295	4271	3815	3624	3445	3240
1	4921	4441	4527	4082	3840	3506	3519
0.316228	6189	5818	5812	5383	5159	4576	4626
0.1	7653	7093	7165	7211	6918	5598	5950
0.031623	5843	5480	5410	5563	5415	4221	4594
0.01	2803	2665	2798	2723	2574	2063	2176
0.003162	1048	1036	1009	997	918	765	773
0.001	522	474	543	475	520	401	443

Table 8.4 **Raw data for the dynamical enrichment of RelA as measured by SMT**

chromatin class D	1	2	3	4	5	6	7
3.162278	1526	1485	1503	1535	1317	1278	1096
1	1977	2015	1905	1795	1738	1670	1419
0.316228	4142	4773	4883	4803	4348	4145	3803
0.1	6126	7299	7424	7560	7075	7657	7271
0.031623	4448	5129	5086	5259	5256	5915	5708
0.01	1878	2185	2176	2369	2443	2827	2617
0.003162	714	856	775	828	845	969	926
0.001	339	386	373	370	422	496	581

Table 8.5 **Raw data for the dynamical enrichment of CTCF as measured by SMT**

chromatin class D	1	2	3	4	5	6	7
3.162278	5421	5937	5622	5655	5259	4941	4045
1	5164	5538	5682	5610	5205	4961	4196
0.316228	4912	5639	5779	5600	5418	5153	4400
0.1	4616	5284	5327	4999	4869	5173	4550
0.031623	2863	3188	3109	2863	2913	3122	3036
0.01	1266	1309	1377	1215	1189	1398	1376
0.003162	507	494	434	416	380	490	495
0.001	234	211	244	189	191	225	215

Table 8.6 **Raw data for the dynamical enrichment of p53 in $BRAF^{V600E}$ as measured by SMT**

chromatin class D	1	2	3	4	5	6	7
3.162278	1150	1384	1354	1408	1415	1276	995
1	1245	1522	1564	1519	1433	1489	1161
0.316228	1309	1526	1549	1591	1559	1727	1252
0.1	1154	1333	1588	1542	1608	1745	1471
0.031623	717	897	1094	963	1085	1120	1085
0.01	318	433	510	482	489	489	534
0.003162	129	139	173	160	195	187	224
0.001	52	87	82	87	102	88	107

Table 8.7 **Raw data for the dynamical enrichment of p53 in BRAF^{WT} as measured by SMT**

chromatin class D	1	2	3	4	5	6	7
3.162278	470	556	672	612	582	517	379
1	626	690	818	803	775	686	432
0.316228	782	1041	1183	1066	931	805	748
0.1	901	1190	1294	1170	1101	874	1030
0.031623	595	682	781	660	632	529	821
0.01	247	267	341	256	268	228	371
0.003162	72	86	101	95	84	87	132
0.001	38	40	54	38	41	45	63

Table 8.8 **Raw data for the dynamical enrichment of p53 in nutlin treated BJ cells as measured by SMT**

chromatin class D	1	2	3	4	5	6	7
3.162278	483	523	505	539	529	419	398
1	580	595	628	577	574	543	416
0.316228	603	755	680	745	659	643	620
0.1	598	673	640	824	782	842	943
0.031623	410	405	366	483	534	604	670
0.01	173	179	160	202	225	294	284
0.003162	51	59	47	69	68	98	112
0.001	31	26	23	32	41	39	37

Table 8.9 **Raw data for the dynamical enrichment of p53 in control BJ cells as measured by SMT**