UNIVERSITA' VITA-SALUTE SAN RAFFAELE

CORSO DI DOTTORATO DI RICERCA INTERNAZIONALE

IN MEDICINA MOLECOLARE

CURRICULUM IN CELLULAR AND MOLECULAR PHYSIOPATHOLOGY

Role of the nuclear organization in p53 search mechanism and gene regulation

DoS: Prof. Davide Mazza 

Second Supervisor: Prof. Jane Mellor

Tesi di DOTTORATO di RICERCA di Matteo Mazzocca

matr. 014336

Ciclo di dottorato XXXIV

SSD BIO/11

Anno Accademico 2021/2002

CONSULTAZIONE TESI DI DOTTORATO DI RICERCA

Il/la sottoscritto/*I*	Matteo Mazzocca
Matricola / *registration number*	014336
nato a/ *born at*	Padova
il/*on*	07/10/1987

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

Role of the nuclear organization in p53 search mechanism and gene regulation

☒ AUTORIZZA la Consultazione della tesi / AUTHORIZES *the public release of the thesis*

☐ NON AUTORIZZA la Consultazione della tesi per …… mesi /DOES 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 part of it might be subject to patentability;*

□ 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 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 on its publication.*

Data /Date ..28/02/2022..... Firma /Signature ……

DECLARATION

This thesis has been composed by myself and has not been used in any previous application for a degree. Throughout the text I use both 'I' and 'We' interchangeably.

All the results presented here were obtained by myself, except for:

1) ***smRNA FISH experiments (Results, paragraphs 7.6 and 7.7, figures 7.21, 7.22, 7.25)*** *were performed by Alessia Loffreda from my group here in San Raffaele Scientific Institute, Milan, Italy.*

2) ***Custom MATLAB code for analysis of SMT/mSIM movies, diffusional anisotropy, DNA FISH and smRNA FISH (Methods, Paragraph 11)*** *was written by my DoS -Davide Mazza- and by Emanuele Colombo of the Imaging Gene Regulation Unit, Experimental Imaging Centre, San Raffaele Scientific Institute, Milan, Italy.*

3) ***Assembly and characterization of the mSIM microscope (Methods, Paragraph 10.2)*** *was performed by Emanuele Colombo and Zeno Lavagnino. Imaging Gene Regulation Unit, Experimental Imaging Centre, San Raffaele Scientific Institute, Milan, Italy.*

4) ***RNA-sequencing (discussed in Results, paragraph 7.5)*** *was performed by the Centre for Omics Sciences, Genomic facility, San Raffaele Scientific Institute, Milan, Italy.*

5) ***Differential Gene expression analysis (discussed in Results, paragraph 7.5)*** *was performed in collaboration with Emanuele Monteleone, San Raffaele Scientific Institute, Milan, Italy.*

All sources of information are acknowledged by means of references.

Acknowledgements

I would like to express my gratitude to my supervisor, Davide, for his tireless mentorship, for the constant support during my PhD and for all the fascinating discussions we had, that expanded my view on science beyond the canonical boundaries of biology. I would like to warmly thank Alessia, for the precious mentorship in the early phases of my PhD and for the unconditional support she gave me throughout time. And thanks a lot to Emanuele, for his important contribution on analysis pipelines which helped me much to speed up the data processing.

I want to say a special thanks to Cecilia, for her moral help in difficult times, but more generally for the emotional support she gives me every day.

Finally, I want to thank all the lab members, including members of the Experimental Imaging Center, for their continuous support to my PhD study and research, and in particular: Carlo, Tiziana, Marco, Daniela, Tom, Fulvio, Edda, Paola, Andrea, Amleto, Serenella.

ABSTRACT

A fundamental question in molecular biology is how the organization of the cell nucleus regulates gene expression. In its first steps, gene regulation requires transcription factors (TFs) to scan the genome and find their targets among a myriad of DNA random sequences, a process culminating with the association to specific promoters and enhancers. Bacteria have evolved special molecular mechanisms to elude the 'sequestering' effect of random DNA and speed up the TF search, making gene regulation extremely efficient. Whether eukaryotic TFs share similar mechanisms, how they find and selects their targets and whether the nucleus define local compartments that may facilitate or exclude their recruitment, is poorly understood. Recent evidence points out that TFs can engage weak and dynamic interactions with the surrounding environment, modulating the search process, but it is unclear what nuclear substructures are involved.

In this work we characterize how the human tumor suppressor p53 – an important TF regulating multiple pathways in response to DNA damage such as cell cycle arrest, senescence and apoptosis – searches for its target genes in the cell nucleus, and how nuclear organization controls this process. We apply a novel fluorescence live imaging technology to follow in real-time the search process of p53 through nuclear compartments. The approach combines single molecule tracking (SMT) -to study the dynamic interactions in living nuclei of individual p53 molecules- with multifocal structured-illumination (mSIM), to map distinct compartments in the nucleus at high resolution.

We find that p53 'perceives' the conformation of the nuclear environment through contacts mediated by its intrinsically disordered regions (IDRs). These interactions guide p53 through chromatin compartments of different permeability, to efficiently locate p53 targets and induce gene expression.

Table of contents

ACRONYMS AND ABBREVIATIONS

ChIP-seq	Chromatin immunoprecipitation
CTD	C-terminal domain
CDs	Chromatin domains
d_f	Dimension of diffusible space
DNAb domain	DNA binding domain
DNAb mutant	DNA binding mutant
d_w	Walker dimension
eRNAs	Enhancer RNAs
IC	Interchromatin compartment
IDRs	Intrinsically disordered regions
IR	Irradiation
xHOTs	Extremely occupied targets
KI	Knock-in
KO	Knock-out
LADs	Lamina associated domains
lnRNA	Long non-coding RNAs
MSD	Mean squared displacement
mSIM	Multifocal structured illumination microscopy
NBs	Nuclear bodies
paSMT	Photoactivated single molecule tracking
p53-ΔC	p53 C-terminal deletion
p53-ΔN	p53 N-terminal deletion
p53-HT	p53-HaloTag
PIC	Pre-initiation complex
PTMs	Post-translational modifications
RE	Regulatory element
shRNA	Short hairpin RNA
SMT	Single molecule tracking
τ_{search}	Search time
TAD	Transactivating domain
TADs	Topologically associating domains
TF	Transcription factor
WT	Wild-type
4OHT	4-Hydroxytamoxifen

LIST OF FIGURES

INTRODUCTION

1 The tumor suppressor p53

p53 is a transcription factor (TF) with a prominent role in cancer biology and very often mutated in tumors. The wild-type (WT) p53 acts as a tumor suppressor that orchestrates a complex transcriptional network in response to DNA damage, to neutralize dangerous consequences in terms of genome stability and cancer development.

In this section I illustrate how p53 regulates several and alternative pathways, recognizing hundreds of genes in the genome. This intense activity requires p53 to find the right genes at the right time, a complex task achieved through a molecular search in the cell nucleus. This process -poorly understood for all mammalian TFs- is referred to as the 'search mechanism', and as we will discuss later in the thesis, is strictly related to the nuclear organization.

We will conclude this first section describing the structure of p53. As it is always observed in biology, structure is strictly related to functions, and also in the context of the search mechanism, some structural features appear relevant.

1.1 p53 as a transcription factor involved in tumor suppression

The tumor suppressor p53 is a TF with strong anticancer functions, inhibiting formation and progression of malignancies. Indeed, in order to freely proliferate, most tumors inactivate p53 either directly by mutations of the p53 encoding gene (i.e. TP53 (Morris *et al*, 2019; Lawrence *et al*, 2014; Kandoth *et al*, 2013)), or indirectly by alterations of the p53 signaling pathways (Laurie *et al*, 2006; Tweddle *et al*, 2003). For instance, some cancers inhibit p53 WT by overexpressing its negative regulator MDM2 (Momand *et al*, 1998).

In first approximation, the antitumor functions of p53 can be ascribed to its fundamental role in directing the transcriptional response to DNA damage in the cell nucleus, articulated throughout all the several stages of tumorigenesis.

Normal cells are continuously exposed to a variety of genotoxic sources, both endogenous (e.g. oxidative stress (Kryston *et al*, 2011), replication fork collapse (Alexander & Orr-Weaver, 2016)) and exogenous (e.g. smoke carcinogens (Pfeifer *et al*, 2002), UV irradiation (Sinha & Häder, 2002)). The resulting DNA damage is a serious incident for the tissue homeostasis: if not correctly fixed, it can introduce genetic alterations (e.g. point mutations, chromosome aberrations) that may activate

oncogenes. DNA lesions are detected by a set of kinases (e.g. ATM, ATR) that phosphorylating p53 induce its rapid accumulation into the cell nucleus (Maréchal & Zou, 2013; Cheng & Chen, 2010). As a TF, p53 controls a complex transcriptional network, regulating genes involved in DNA repair, transient cell cycle arrest, senescence and apoptosis (Hafner *et al*, 2019; Aubrey *et al*, 2018; Sullivan *et al*, 2018). Globally, the combination of these multiple and sometimes alternative pathways (e.g. senescence vs apoptosis; Figure 1.1) effectively reduces the propagation of genetic errors in the tissue: either the DNA damage will be repaired, or the irreversibly damaged cells will be prevented to transmit their alteration to daughter cells, through the activation of senescence or apoptosis.

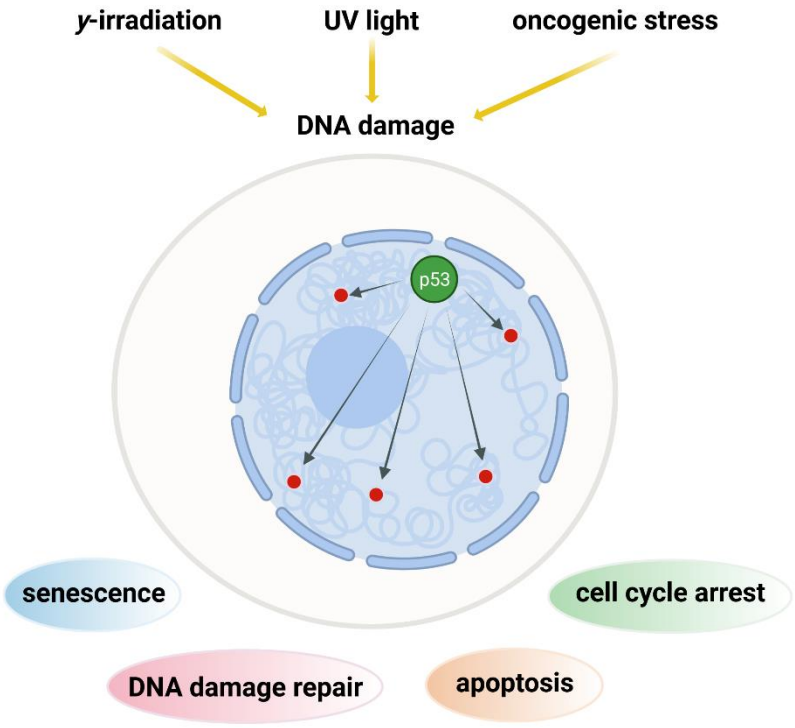


Figure 1.1 Cellular pathways activated by p53 in response to DNA damage. *Cells activate p53 in response to genotoxic agents that induce DNA damage. The activated p53 navigates the nucleus, searching for the right target genes within chromatin. Gene expression mediated by p53 avoids the transmission of dangerous lesions to cell daughters, by several cellular programs (e.g. cell cycle arrest, apoptosis). Image created using BioRender.com.*

Beyond its tumor suppressive functions in normal tissues which reduce the frequency of oncogene activation, p53 is fundamental in pre-cancerous cells (i.e. cells with induced oncogenes) where its activity represents a last barrier against tumorigenesis (Haigis & Sweet-Cordero, 2011; Bartek *et al*, 2007). How does p53 get activated in pre-cancer settings? Again, DNA damage plays a central role. Pre-cancerous cells are not only exposed to the conventional genotoxic sources as normal

cells, but they are also susceptible of an additional form of DNA damage (known as replication stress) due to the high cell proliferation rate. Even though it is still debated whether only the high level of replication-induced DNA damage or also other mechanisms activate p53 in these pre-oncogenic conditions, p53 strongly reduces the carcinogenic risk at this stage once again by inducing permanent cell cycle arrest or cell death (Haigis & Sweet-Cordero, 2011).

When tumors develop, eluding all the antineoplastic barriers, their metastatic potential and the prognosis are also influenced by p53 (Hainaut & Pfeifer, 2016). Typically, cancers retaining p53 WT are less vulnerable of the so-called genomic instability, a rapid accumulation of genetic alterations usually involving large DNA fragments (e.g. chromosomal rearrangements and aneuploidy; (Eischen, 2016)). Such increased instability drives cancer evolution, providing a rich source of genetic mutations that confers tumors higher capacity of immune evasion and metastasis (Vendramin *et al*, 2021). As a consequence, cancers with p53 WT -despite a reduced antitumor capability than physiological conditions- can still limit genomic instability better than tumors with mutated p53 (Eischen, 2016; Redman-Rivera *et al*, 2021).

How p53 coordinates these cell-fate decisions remains an open question. Although transcription independent non-canonical functions of p53 have been described (i.e. interference with autophagy and metabolic pathways (Hager & Gu, 2014)), most studies focus on p53 as a transcriptional regulator. In this context, a key question has been what kind of a transcriptional regulator p53 is, if an activator only or also a repressor (Fischer, 2017). A conspicuous number of works indicate that p53 can down-regulate some genes, through a repressive mechanism that would involve the DNA architecture of the regulatory elements bound by p53 (Wei *et al*, 2006; Schlereth *et al*, 2013; Smeenk *et al*, 2011; Cui *et al*, 2011; Nikulenkov *et al*, 2012). However other studies, including meta-analysis that collected data from independent works, highlight lack of strong evidence for p53 to be a transcriptional repressor (Tonelli *et al*, 2015; Fischer, 2017; Hafner *et al*, 2019; Su *et al*, 2015; Rashi-Elkeles *et al*, 2014, 53; Hafner *et al*, 2017; Fischer *et al*, 2014; Younger & Rinn, 2017; Fischer *et al*, 2016). An argument against the repressive function of p53 is also given by the observation that genes reported to be down-regulated usually do not display canonical p53 binding sites, raising the question of whether these sequences are really functional (Hafner *et al*, 2019; Hoffman *et al*, 2002; Lipski *et al*, 2012; Johnson *et al*, 2001). Finally, in a recent study, the functionality of more than 1500 potential p53 binding sites identified by ChIP-seq experiments was tested. In this study, any potential binding site was coupled to the same reporter gene. Among all the tested sequences only 40% were effectively bound by p53, and, interestingly, none of those

was significantly associated with down-regulation of the reporter gene (Verfaillie *et al*, 2016). Overall, there is a wide consensus in the p53 field to consider p53 as a direct activator of transcription and -only indirectly- of gene repression.

1.2 The importance of p53 binding activity

The activation of several pathways requires p53 to efficiently locate hundreds of genes in the genome, among a plethora of random DNA sites. The identification of these genes culminates into the binding of specific DNA response elements (REs) placed in their promoters or enhancers. A p53 RE is composed of two copies (half-sites) of a 10 bp sequence. The half-sites can be separated by a spacer of variable length (0-20 nucleotides), with the binding likelihood that is inversely proportional to the spacer length (Funk *et al*, 1992; Nguyen *et al*, 2018; Tonelli *et al*, 2015). The consensus of each half-site is given by RRRCWWGYYY (where R= A or G; W= A or T; Y= C or T). Initially discovered by in vitro studies (El-Deiry *et al*, 1992; Funk *et al*, 1992), the aforementioned consensus is also the preferential p53 binding motif in the cell nucleus, as confirmed by genomic profiles of p53 binding generated with different chromatin-immunoprecipitation approaches (e.g. ChIP-seq, ChIP-PET) from independent studies (Nguyen *et al*, 2018; Pavletich *et al*, 1993; Horvath *et al*, 2007; Botcheva *et al*, 2011; Fischer, 2017). In a recent meta-analysis, Menendez and colleagues estimated ~783,000 potential p53 REs throughout the human genome (Nguyen *et al*, 2018). By comparing 41 ChIP-seq datasets (some of which generated by them, some from other groups), the authors have investigated how p53 selects its target sites. As expected, the binding likelihood in vivo strongly depends on the composition and structure of the REs. In particular, two half-sites with exact consensus and without any spacer (indicated as perfect REs) have higher binding likelihood compared to other RE classes. However, most of the potential REs (98.4%) -including ~25% of the perfect REs- are found not to be bound in any of the examined dataset, indicating that additional factors regulate p53 selectivity. Moreover, in their natural position within chromatin, some well-known p53 target genes are left unbound and non-regulated depending on the cell type or stimulus, indicating a strong selectivity depending on the cellular context (Nguyen *et al*, 2018; Hafner *et al*, 2020).

What does guide the choice of p53 among all the potential REs? The selectivity process requires p53 to find the right genes in the cell nucleus, after a challenging search among the enormous amount of nonspecific DNA. Indeed, the selectivity process can be seen as a dual 'problem' where TFs have to both efficiently select

their specific REs over random DNA and - among all the specific genes- to regulate a subset of them, depending on the cellular conditions. As we will discuss in detail, the presence of nonspecific DNA may sequester TFs from their REs slowing down the search process and the efficiency of gene regulation. A successful target selectivity requires (a) TFs to efficiently explore chromatin in order to escape from random DNA and localize the right genes and (b) the target genes to be physically accessible to TFs diffusion and binding, with both conditions being strongly influenced by the nuclear architecture, as we will see in the next chapters.

1.3 Structure and regulation of p53

The human p53 is a protein composed by 393 amino acid residues, organized into several well-established domains with a precise function and a peculiar three-dimensional folding, in which a structured core in the middle is flanked by two N- and C-terminal intrinsically disordered regions (IDRs (Joerger & Fersht, 2008)). The modular structure with several IDRs is typical of proteins at the center of regulatory networks, because it confers a conformational adaptability that facilitates the interaction with a multitude of molecules different in nature (e.g. cofactors, RNAs) (Teilum *et al*, 2021).

The disordered N-terminus contains two contiguous transactivation domains (TADs) that stimulate the transcription of target genes. The TADs are followed by: a proline rich region, the structured core constituting the p53 DNA-binding (DNAb) domain, an oligomerization site and the disordered C-terminal domain (CTD; Figure 1.2).

The fully activated p53 is a tetramer (or, more precisely, a dimer of dimers; (Kitayner *et al*, 2010)). As seen previously, p53 REs are formed by two decamers (half-site) in tandem, in which each half-site can be further divided into two quarter sites (El-Deiry *et al*, 1992; Funk *et al*, 1992). Given the tetrameric structure of p53, it has been proposed that each monomer specifically binds to a single quarter site. Consistently, crystal structure of p53 in complex with DNA shows two p53 molecules that symmetrically bind a half-site (Kitayner *et al*, 2006, 2010).

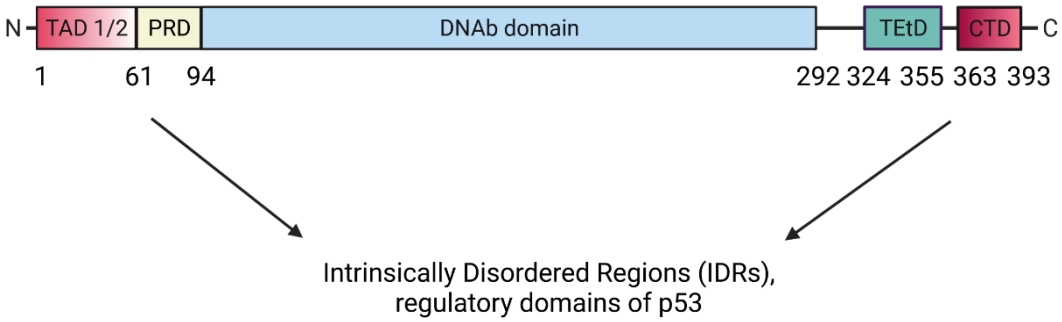


Figure 1.2 p53 protein domains. *The structure of p53 is composed of several domains with different structures and functions. The N-terminal transactivation domain and the C-terminal domain, regulate p53 activity through their intrinsically disorder. At the center, the structured DNA binding domain. Image created with BioRender.com.*

Like the activation domains of several TFs (e.g. glucocorticoid receptor), the disordered p53 N-terminus is enriched in acidic residues, a feature that first led to the hypothesis that p53 might be a transcriptional activator (Hollenberg & Evans, 1988; Raj & Attardi, 2017). Afterwards it was shown that this p53 region has a transcriptional potential similar to the strong herpes simplex virus protein VP16 (Fields & Jang, 1990). Commonly to VP16, p53 contains two autonomous TADs that activate transcription independently from each other (Candau *et al*, 1997; Chang *et al*, 1995). Recently, it has been observed that mice with mutations inactivating a TAD at a time, do not display higher risk of cancer than their WT counterpart, suggesting that a single TAD allows p53 for its tumor suppressor functions. Having two independent TADs may explain why tumors exert a selective pressure that fosters frequent mutations in the DBD of p53 but no in its N-terminal domain (Baugh *et al*, 2018; Sullivan *et al*, 2018).

The flexible conformation of the TADs allows p53 for promiscuous interactions with a myriad of different proteins, including its negative regulators MDM2 and MDMX (also called MDM4; (Joerger & Fersht, 2008)). MDM2 is an E3 ligase that binds and ubiquitinates p53 inducing its proteasomal degradation. MDMX does not have ubiquitin ligase properties by itself, but forms heterodimers with MDM2 that prevent MDM2 autoubiquitination (Biderman *et al*, 2012). Overall, in non-stressed conditions (i.e. absence of DNA damage), MDM2 and MDMX keep p53 at low nuclear level and, by binding p53 they physically obstruct the association between the TADs and transcriptional coactivators, inhibiting further the p53 transcriptional potential. Since MDM2 is a transcriptional target of p53, these two proteins form a negative feedback

loop that controls p53 levels in the absence of stress and during the return to homeostasis following stress.

When cells undergo DNA damage, p53 is phosphorylated at multiple sites within the TADs, reducing the affinity for MDM2. TADs are substrates for phosphorylation by several kinases involve in the DNA damage response, including ATM, ATR and Chk2 (Jenkins *et al*, 2012). In turn, p53 becomes less ubiquitinated and it rapidly accumulates in the nucleus. Lower MDM2-p53 binding renders the TADs more accessible to coactivators, chromatin modifiers, and general transcription factors (e.g. p300; (Raj & Attardi, 2017)). Indeed, gene regulation requires that TFs (e.g. p53, SOX2, MYC) coordinate a large variety of proteins to finally activate the RNA-polII (Lambert *et al*, 2018). Yet, it is not mechanistically clear at which level each cofactor is required. For instance, it is unknown whether cofactors modulate the search process of TFs and their binding selectivity. In principle, chromatin modifiers in complex with a TF could open the way to otherwise inaccessible genes, facilitating TF diffusion through chromatin and gene targeting. In line with this hypothesis, recently it has been observed that the mobility of the glucocorticoid receptor throughout the nucleus is modulate by its TAD (Garcia *et al*, 2021a). Similarly, the Hypoxia-Inducible Factor (HIF) 1-α and HIF 2-α that naturally display different nuclear mobilities, become interchangeable when swapping their TADs (Chen *et al*, 2021). It is unknown among TFs whether the search process is commonly regulated by TADs. If this phenomenon is mediated by cofactors and whether it determines what compartments a TF can visit in the nucleus and what it cannot, need to be elucidated further.

At the opposite side of p53, the disordered CTD is considered a fundamental domain regulating p53 activity. Differently from the acidic TADs, the CTD displays basic properties with positive charges provided by a stretch of lysine residues (Weinberg *et al*, 2004a). The positive electrostatic potential combined with its intrinsic disorder adapts the CTD to many different substrates, including proteins, but also DNA and RNA molecules (Laptenko & Prives, 2006). Importantly, the positive lysine charges mediate nonspecific contacts with DNA, relevant for the p53 search mechanism. In vitro, these transient interactions allow p53 to diffuse along the DNA, a mechanism known as 'DNA sliding' (Tafvizi *et al*, 2008, 2011; Itoh *et al*, 2016). As we will discuss later, bacteria TFs adopt a peculiar search process – named facilitated diffusion - in which they sometimes slide on DNA, sometimes freely move in the three-dimensional (3D) space (Berg *et al*, 1981; Riggs *et al*, 1970). Such an alternate diffusion, reduces the search time of TFs to their binding sites in vivo (Elf *et al*, 2007). In mammals, there has not been direct evidence of TFs sliding in vivo so far.

However, p53 is one among very few mammal TFs that have been reported to diffuse on DNA, at least in vitro. Whether p53 undergoes DNA sliding also in vivo, and what is the impact on p53 the search mechanism and selectivity is a question that still needs to be addressed.

Aside its potential role in sliding on DNA, the CTD is required for p53 binding selectivity (Hamard *et al*, 2012; Laptenko *et al*, 2015). For instance, in a ChIP analysis, the human tumor line H1299 has been engineered for the comparison of p53 mutants through inducible expression (Laptenko *et al*, 2015). p53 lacking the entire CTD recognized only a fraction (210/355) of the REs bound by its p53 WT counterpart. As a selectivity mechanism, the authors proposed that once p53 lands on its REs, the CTD induces a conformational change within the DBD, locking p53 on such sites, implying that the CTD mutant cannot stably bind some genes. However, it has not been investigated whether the mutant can still find the REs left unbound. In view of the reported CTD capability of sliding in vitro, the proposed 'locking' mechanism and a selectivity based on the search efficiency may be not mutually exclusive.

The complex CTD regulation is also given by multiple interactions with a variety of proteins. Among others, the CTD interacts with several cofactors, some of which in common with the TADs, including the acetyltransferase p300 (Kim *et al*, 2012). How does the ability of the CTD to bind cofactors reconciliate with its sliding properties on DNA? Whether protein-protein and protein-DNA contacts occur simultaneously within the CTD is not known. On the other hand, the recruitment of common cofactors may suggest that the N-terminal and C-terminal IDRs communicate to each other, exchanging molecular partners. While readily disposable to contact accessible DNA, the CTD would benefit from exchanging chromatin modifiers with the TADs when encountering sequences occluded by tight nucleosomes. Indeed, p53 belongs to a special category of TFs named '*pioneer*' factors that can bind their target genes also in condensed chromatin. For example, at basal conditions, several highly induced p53 target genes (e.g. *CDKN1A*, *PUMA*) are characterized by repressive histone modifications or CTCF binding, which is often a marker of repressed chromatin (Su *et al*, 2015; Gomes & Espinosa, 2010). Despite a close chromatin conformation, after stimulation p53 effectively reaches and binds its condensed targets.

To complicate the CTD regulation further, the abovementioned lysines are subject to extensive acetylation by p300 itself, yet their precise functions in vivo remain puzzling: while it is well established that acetylated p53 is associated to a stronger transactivation potential (Tang *et al*, 2008), controversies have arisen on whether p53 acetylation plays a role in modulating nonspecific binding on DNA, specific

binding at REs or the recruitment of cofactors (Laptenko *et al*, 2015; Loffreda *et al*, 2017a; Reed & Quelle, 2014, 5).

Although not much explored yet, several lines of evidence suggest that the p53 CTD can interact also with RNA molecules (Riley & Maher, 2007). Recently it has been shown that p53, after binding to the *CDKN1A* promoter, induces the transcription of a previously unknown long noncoding RNA (lnRNA), on the opposite strand of *CDKN1A* (Schmitt *et al*, 2016). Such lnRNA -named *DINO* (Damage Induced Noncoding)- physically interacts with the p53 CTD, and have been found to co-occupy with p53 other well-known p53 target genes (e.g. *DDB2*). The silencing of DINO reduces the binding of p53 to its targets and -in turn- their expression. Generally speaking, several studies show that lnRNAs can interact with specific genomic regions (Statello *et al*, 2021), but it is still unclear whether these RNAs guide the targeting of TFs to DNA.

Finally, the multivalency of protein IDRs (i.e. their propension to interact with multiple molecules, including RNAs) is particularly relevant for the formation of molecular condensates, in which multiple copies of TFs and coactivators converge forming nuclear clusters (Alberti *et al*, 2019; Chong *et al*, 2018; Sabari *et al*, 2018). As we will see later, molecular condensates typically display features of liquid droplets, like the capability to fuse together and generate bigger clusters. Despite their role in gene regulation is still fervently debated, some line of evidence suggest that biomolecular condensates may increase transcription (Leslie, 2021; Narlikar *et al*, 2021; Hnisz *et al*, 2017). Among several mechanistically models, condensates would enhance transcription because a locally higher TFs concentration should increase the binding rate to their target genes (Suter, 2020). Surprisingly, despite two large IDRs and a pronounced heterogeneity of interactions, there is little evidence that p53 can form self-condensates. In vitro studies have shown that p53 undergoes droplet formation only under certain physical-chemical circumstances, as high salt concentration or the addition of droplet formation agents (Kamagata *et al*, 2020; Boija *et al*, 2018). Moreover, p53 displays a lower propensity to form condensates compared to other TFs (e.g. MYC, SOX2; (Boija *et al*, 2018)).

Overall, the regulatory IDRs confer p53 with a large plasticity of interactions and functions that are only partially understood. Despite a large body of data, the contribution of p53 IDRs to the p53 search mechanism and binding selectivity have been not explored yet in living nuclei.

2 The role of search mechanism in TF activity and gene regulation

TFs regulate the expression of genes by the interaction with DNA regulatory elements (REs) at promoters or enhancers. The TF-RE interaction appears disfavored in vivo, because several molecular layers can reduce TF targeting efficiency.

In a hypothetical solution containing a TF and a DNA RE sequence, TF and DNA diffuse and meet with a constant frequency throughout time and, as a consequence, the DNA-RE interaction is governed only by their affinity, determining the stability of the molecular complex. However, in vivo, the DNA-RE molecular stability, taken alone, seems insufficient to ensure an effective interaction. Indeed, the massive presence of random DNA could sequester the TF reducing enormously the frequency by which a TF can find a given gene, with consequences in terms of gene expression up to case that the TF is so unlike to locate a gene to leave it unregulated. To overcome this problem, bacteria evolved a special search mechanism (already mentioned and thoroughly discuss later) called facilitated diffusion, where the nonspecific interactions with DNA, instead of having a sequestering effect, speed-up the search. Currently, however, it is unknown whether mammalian TFs adopt similar search strategies in vivo.

Another challenge that human TFs may face when 'searching' for genes, is given by molecular barriers in the crowded nuclear environment (e.g. nucleosomes), that could hamper the targeting process.

To highlight the search problem, in this section we first propose an analysis on p53 selectivity observed in vivo, and later we provide a quantitative reasoning on the impact that the large excess of nonspecific DNA in the human genome could have on TF activity, if we do not assume special search strategies.

2.1 The RE affinity fails to explain p53 binding selectivity

TFs locate their target genes by recognizing specific DNA elements in the genome, sequences of about 10-20 bp placed within promoters or enhancers. Yet, at any time, depending on the cellular context, TFs typically recognize a fraction of REs regulating only a subset of target genes (Biggin, 2011; Lambert *et al*, 2018). This can be relevant for the cellular output. For instance, even though p53 ubiquitously activates some targets (e.g. *CDKN1A*), other genes (e.g. *GADD45A*) are preferentially bound and regulated only in certain cell types or under specific stimuli (Nguyen *et al*, 2018). As elegantly demonstrated by Lahav's group, different p53 selectivity can lead cells

to opposite biological functions (Purvis *et al*, 2012). In their work, cells exposed to the same extent of DNA damage, but different abundance of p53 over time, experienced different p53 selectivity and cell fate. While cells with oscillatory p53 nuclear level reenter the cell cycle after a transient arrest, sustained p53 level induced senescence, with the two phenotypes that were linked to the p53 selection of different target genes. Under certain circumstances, differences in binding selectivity can be explained by the so called 'affinity model', where REs with low affinity are effectively bound only when p53 abundance is high enough (Chen *et al*, 1996; Schlereth *et al*, 2010). The affinity model is based on the stability of the p53-RE complex estimated in vitro (Weinberg *et al*, 2005). The binding strength of a TF-DNA complex depends on the number and type of non-covalent interactions (Hydrogen bonds, Van der Waals interactions) between the TF amino acids and RE nucleotides (Stormo & Fields, 1998; Coulocheri *et al*, 2007). As mentioned above, in their promoters, p53 target genes can have REs with different nucleotide composition and variations on the decameric motif RRRCWWGYYY (R= A or G; W= A or T; Y= C or T) that result in different p53-RE stability. Cell cycle genes often carry REs with higher affinity than those regulating apoptosis (Weinberg *et al*, 2004b; Qian *et al*, 2002; Noureddine *et al*, 2009; Inga *et al*, 2002; Schlereth *et al*, 2013). As such, the affinity model would imply that at low nuclear level, p53 occupies only high-affinity promoters leading to cycle arrest. At higher concentration (usually achieved by higher doses of DNA damage), p53 should accumulate also on apoptotic genes, inducing cell death. Despite some evidence consistent with this model (Chen *et al*, 1996; Schlereth *et al*, 2010), a not negligible number of genes with prominent apoptotic role (like *BBC3* (Kiraz *et al*, 2016; Elmore, 2007), carry high-affinity REs and are bound also at low p53 levels (Nguyen *et al*, 2018; Weinberg *et al*, 2004b; Kracikova *et al*, 2013; Kaeser & Iggo, 2002). Most importantly, such an affinity model based on the RE sequence, appears inadequate to predict not only the cell fate, but also whether p53 will recognize a certain gene in living cells.

To illustrate this concept, we propose a comparison between p53 binding frequency at specific genes in vivo with their affinities measured in vitro (Figure 2.1). We take advantage of affinity measurements provided by Alan Fersht and colleagues on a small subset of p53 targes and, for p53 binding profiles in vivo, we refer to 16 publicly available ChIP-seq datasets, including (a) different cell types and (b) different DNA damage treatments (Nguyen *et al*, 2018; Weinberg *et al*, 2005). For each gene we counted the number of datasets with observed p53 binding in vivo. From thermodynamics considerations, we would expect p53 to be more frequently bound to higher-affinity than lower-affinity REs (see **BOX 1**). This is not the case,

however, because high-affinity genes as *RRM2B* and *GADD45A* are recognized by p53 in less than 50% of datasets, while, for instance, the low-affinity *BAX* is bound in almost all cases, suggesting further mechanisms that may facilitate the recruitment of p53 at this locus. This latter case is paradigmatic, because *BAX* is a key apoptotic gene in the 'affinity model', proposed to have poor p53 binding in vivo.

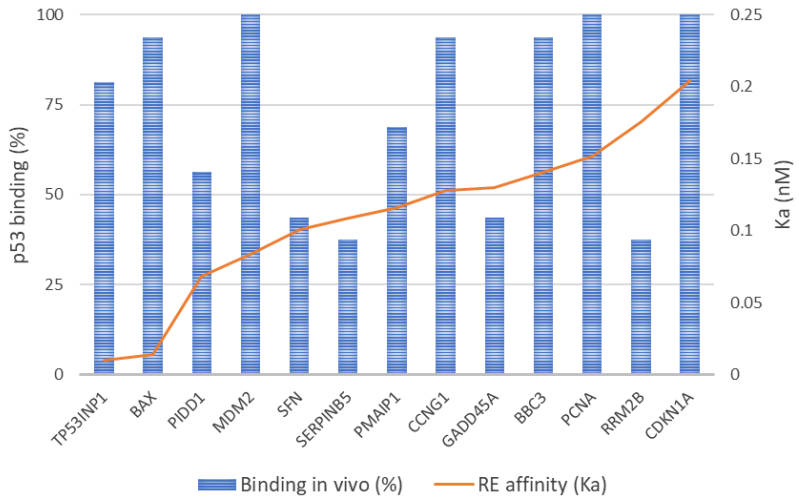


Figure 2.1 Relationship between p53 in vitro affinity and in vivo occupancy. *REs from well-established target genes, display different affinities in vitro (orange line), but a poor correlation with occupancy in vivo, as highlighted by the percentage (in blue) in which these REs are identified as effectively bound in cells (from available ChIP-seq datasets).*

Why these discrepancies between REs affinity and binding in vivo? As mentioned, such an affinity is derived from measurements in simple systems, where DNA fragments are incubated with multiple copies of p53. Under these conditions, the amount of TF-DNA interactions is dominated by non-covalent bonds that keep them associated (Coulocheri *et al*, 2007). In vivo, where the folded chromatin create sub-compartments and segregate parts of the genome in distinct physical and biochemical environments, additional factors beyond the pure DNA sequence contribute to TFs selectivity (Fraser & Bickmore, 2007). For instance, the organization of the nucleus often constrains the diffusion of proteins, and may transiently confine a TF in nuclear sub-compartments (Suter, 2020; Woringer & Darzacq, 2018a). A TF persistently exploring a defined region, has higher chance to meet possible genes, which may justify why some genes can be more frequently bound than expected by their affinity. On the other hand, genes placed at sites rather impermeable to diffusion (e.g. condensed chromatin), may be difficult to be reach for a TF, resulting in no binding.

BOX 1. *Expected in vitro occupancy: comparison of genes with different affinities.*

The simplest system to model TF-DNA binding is represented by a test tube containing the TF of interest and the consensus sequence to be bound. In this system, the TF-DNA binding is described by a first-order reversible reaction, represented by the chemical equation:

TF + DNA ⇌ TF-DNA (Eq. 1)

Here, the forward arrow represents association, and the speed at which these associations occur f_a depend by (a) the concentrations, [TF] and [DNA], of the two reactants and (b) by the association rate constant k_{on} with which an individual TF molecule -occupying a certain unit volume- will meet and interact with a DNA molecule in the same unit volume.

$$f_a = k_{on}[\mathrm{TF}][\mathrm{DNA}] \quad \text{(Eq. 2)}$$

For a diffusion-limited reaction (the TF binds its target as soon as they meet), k_{on} is determined by the Smoluchowski relationship, $k_{on} = 4\pi D_{TF} a$, where D_{TF} is the diffusion coefficient of the factor and a is the size of the target. Fast diffusing factors can find their targets more rapidly. Vice-versa, the complex dissociation speed f_d is provided by the product of the complex concentration by the dissociation rate constant k_{off}:

$$f_d = k_{off}[\mathrm{TF-DNA}] \quad \text{(Eq. 3)}$$

In the solution, the TF + DNA ⇌ TF-DNA reaction proceeds until the chemical equilibrium is reached, a condition in which formation and the disassembly of the TF-DNA complex occur at the same rate. At equilibrium:

$$k_{on}[\mathrm{TF}][\mathrm{DNA}] = k_{off}[\mathrm{TF-DNA}] \quad \text{(Eq. 4)}$$

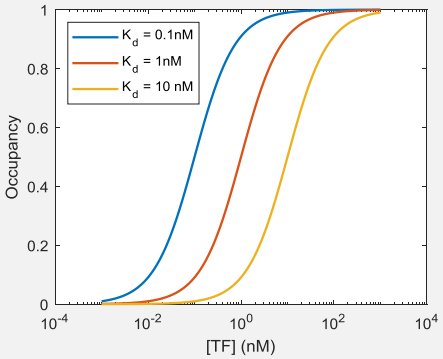
The same relationship can be written by introducing the dissociation constant $K_d = k_{off}/k_{on}$ so that at equilibrium:

$$K_d = \frac{[\mathrm{TF}][\mathrm{DNA}]}{[\mathrm{TF-DNA}]} \quad \text{(Eq. 5)}$$

From the definition of K_d, we can compute the probability for a DNA molecule to be bound by the TF as function of the concentration of free transcription factor. We will call this probability 'Occupancy', O, that is given by:

$$O([TF]) = \frac{[TF-DNA]}{[TF-DNA]+[DNA]} = \frac{1}{1+[DNA]/[TF-DNA]} = \frac{1}{1+\frac{K_d}{[TF]}} \sim \frac{1}{1+\frac{K_d}{[TF_{tot}]}} \quad \text{(Eq. 6)}$$

Where $[TF_{tot}]$ is the total concentration of transcription factor (bound and unbound) and the last approximate equality holds when the concentration of TF is much higher than KD. The condition with $[TF]= K_d$ corresponds to 50% of probability to have a DNA molecule occupied by the TF. The figure below illustrates the occupancy of TF, as function of its free concentration and of the dissociation constant of the reaction.



If the DNA described here was a DNA response element (RE) controlling the expression of a gene, Eq. 7 would in principle describe a direct link between thermodynamics and the TF biochemical function. Assuming a simple model where TF binding is directly coupled with transcription, one could predict target site occupancy and gene activity given only the TF concentration as an input.

2.2 A simple affinity model highlights that random DNA dramatically sequesters TFs

If we assume a TF selectivity based on the DNA affinity only, we undergo a relevant paradox -yet not always taken into account- which involves the enormous amount of nonspecific DNA in the genome. At a very low degree, TFs have affinity also for nonspecific sequences. Typically, TFs bind to random DNA sites with affinities 10^3 – 10^7 lower than specific REs (in the pM to nM range for specific sites and in the nM to μM range for nonspecific ones (Spolar & M. Thomas Record, 1994)). Despite their lower affinity, any random site on DNA could act as a nonspecific target. Nonspecific affinities are much smaller than specific ones, but the quantity of random DNA in the nucleus is amazingly higher than specific REs, leading -in human diploid cells- to $3x10^9$ nonspecific sites vs. about 10^5 specific ones. In **BOX 2**, we estimate the time

BOX 2. *Competition between specific and nonspecific binding.*

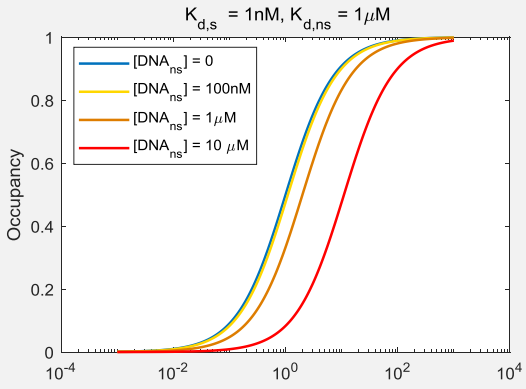
We can provide a simplified model of the competition between nonspecific and specific binding by considering a test tube containing a certain concentration of TF, a small amount of DNA molecules containing the consensus sequence for the TF and an excess of nonspecific DNA molecules. In these conditions the fraction of TF molecules not engaged in nonspecific binding will be provided by:

$$f_{free} = \frac{[TF]}{[TF-DNA_{ns}]+[TF]} = \frac{K_{d,ns}}{K_{d,ns}+[DNA_{ns}]} \quad \text{(Eq.7)}$$

Where $K_{d,ns}$ represents the dissociation constant from nonspecific sites. The occupancy at the specific sites can then be calculated by considering that only the non-sequestered population of TFs can be directed to the specific sites, by modifying Eq. 6 into:

$$O([TF_{tot}]) \sim \frac{1}{1+\frac{K_{d,s}}{f_{free}[TF_{tot}]}} \quad \text{(Eq.8)}$$

Where $K_{d,s}$, is the affinity of TFs for their specific sites. In other words, since a fraction of the TF is sequestered at nonspecific sequences, higher TF concentrations are needed to provide high occupancy at specific sites, as depicted in the figure below. We can use this toy model to schematize a TF searching for one specific binding site among $\sim 10^9$ potential nonspecific ones (in a volume comparable to the one of a eukaryotic cell nucleus, of approximately 200fL, leading to a nonspecific site concentration of approximately 10mM), with non specific dissociation constants in the μM range. Yet, this model would predict all specific sequences (with identical REs) to be occupied to the same extent, a condition that is not met in vivo. In the next section, we will therefore highlight how the complex nuclear eukaryotic environment might affect the TF search and its selectivity.



that a TF would spend on nonspecific DNA, assuming a simple model (i.e. a solution containing TF and DNA only), and using TF concentrations and specific-to-nonspecific DNA ratio encountered in a cell nucleus. The model predicts that 99.9% of the TF molecules would be sequestered at nonspecific sites, and the occupancy of the specific sites would be reduced of $> 1000 \times$ compared to the condition in which nonspecific interactions are not considered. Nonspecific sites can therefore have a non-negligible effect on the TF binding distribution throughout the genome, sequestering the TFs from their specific REs and inhibiting gene regulation.

For these reasons, it is generally thought that living organisms may have evolved some special mechanisms to efficiently recruit TFs at specific sites, involving both the organization of genome and the TF search strategy. For instance, large portions of the genome may be not directly available for binding, because already engaged by other molecules (e.g. histones, RNAs), reducing the actual amount of TF nonspecific interactions. As we will detail below, 1D sliding on DNA offers the possibility to transform nonspecific binding into a useful resource to speed up the search process.

It is assumed very often that also eukaryotic TFs can slide on DNA and efficiently select their targets in the genome. However, in living eukaryotic cells, TFs sliding has never directly observed due to considerable technical limitations. In recent years we have started to accumulate evidence of alternative TFs search processes in mammals, but we are only at the beginning of understanding how TFs and the nuclear architecture cooperate for the TF search and gene selectivity.

More specifically, about p53 several questions arise on the efficiency of its search process. First of all, about the paradoxical 'trapping effect' of the nonspecific genomic DNA. We do not know whether p53 escape from such sequences by (1) an efficient search strategy, by (2) an effective compartmentalization of nonspecific DNA, or (3) a combination of the two. Similar questions, involve the way p53 selects its target genes. It is unknown whether p53 can more efficiently locate some genes than others, whether this depends on their position within chromatin and if this can contribute to the observed discrepancies among the 'affinity model' and the binding observed in vivo.

3 How TFs search for their specific genomic REs

TFs need to find their subset of genes in a reasonable amount of time. A search time (τ_{search}) excessively long would not provide a rapid transcriptional response. Indeed, transcriptional activation orchestrated by p53 can be observed in 30 minutes following DNA damage (Allen *et al*, 2014). As outlined, TFs require efficient search strategies to locate their genes and associate to their REs. The search efficiency may be no equal for different genes, with the TF finding more easily some genes than others. This could translate into different binding frequency and gene expression levels, up to the extreme case of no binding and absence of gene regulation.

The τ_{search} is strictly related to the search strategy adopted by TFs and basically depends on how a TF explores the nuclear environment. Here, we introduce some fundamental notions about the search strategies, expanding some concepts that we recently reviewed (Mazzocca *et al*, 2021b). We will see that different types of exploration retain different target efficiency, in a context-dependent manner. We will introduce the factors that influence the search efficiency and, finally, we will review the current knowledge about the search mechanisms adopted by eukaryotic TFs.

3.1 Principles of the TFs search mechanism

The search mechanism problem can be faced by analyzing the time τ_{search} that a TF employs to meet a specific target. In an oversimplified scenario where a TF freely floats (by 3D random motion) in a homogeneous medium containing a single DNA target sequence, the τ_{search} can be derived by the Smoluchowski equation (Woringer & Darzacq, 2018b; Esadze & Stivers, 2018a):

$$\tau_{search} = \frac{V}{4\pi D a}$$

where V is the volume of the medium, D the TF diffusion coefficient and a the size of the DNA target. Assuming volumes compatible with the cell nucleus, the resulting τ_{search} spans between hours to days (Figure 3.1A), timescales that seem inconsistent with the rapid TF response observed in vivo (Jana *et al*, 2021). This long τ_{search} may be mitigated by having many TF copies searching for their targets simultaneously, but whether the concentration of TFs alone guarantees a rapid binding at promoters is not clear, especially considering the huge amount of nonspecific DNA that TFs encounters in the cell nucleus which -as described above- increase the τ_{search} even further.

For these reasons, it is thought that living organisms have evolved special mechanisms to enhance the TF search efficiency (Jana *et al*, 2021; Mirny *et al*, 2009;

Suter, 2020). Historically, the first evidence on this assumption came from bacteria, where it was shown that the LacI repressor finds its target RE orders of magnitude faster than the Smoluchowski limit (Riggs *et al*, 1970). This implies an exceptionally fast search strategy even for models that already neglect sequestration at nonspecific sites. Such astonishing finding was followed by a growing accumulation of data that led to a detailed description of LacI search mechanism, known as 'facilitated diffusion' (Berg *et al*, 1981; Halford & Marko, 2004). Nowadays, LacI facilitated diffusion is supported also by live-cell studies in vivo using single-molecule approaches. Elegant experiments on fluorescently-tagged LacI where synthetic 'roadblocks' were inserted near the LacI binding site in vivo, demonstrated that facilitated diffusion reduces the LacI τ_{search} in living cells by four times (Elf *et al*, 2007; Hammar *et al*, 2012).

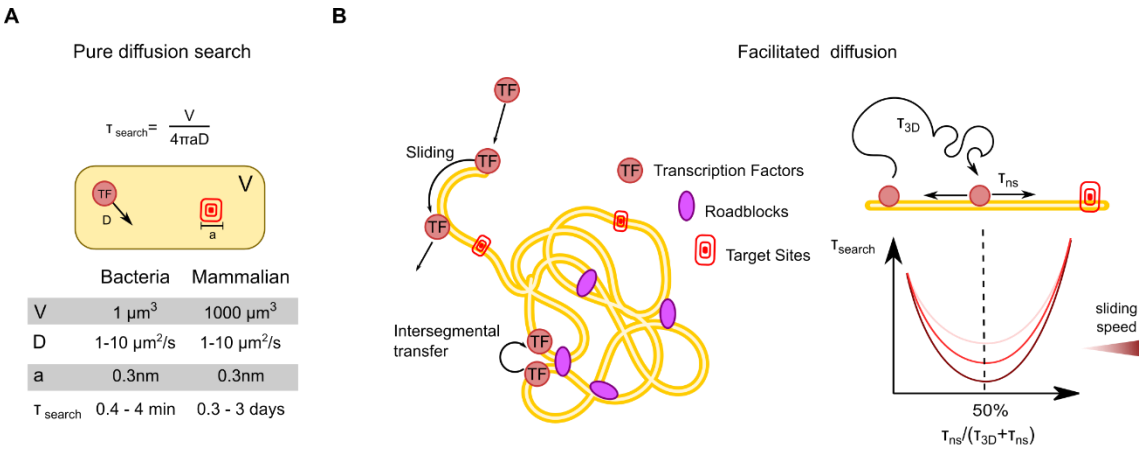


Figure 3.1 Diffusion-limited search and facilitated diffusion. *(A) Smoluchowski relationship for diffusion-limited reactions, where* τ_{search} *is the time taken by a TF to find a target site, V is the size of the explored volume, a is the target site (typically assumed as the size of a nucleotide, since the TF needs to align to its target with base-pair precision) and D is the diffusion coefficient of the molecule. Expected values for these parameters are provided in the table. (B) The mechanism of facilitated diffusion speeds-up the search process, with an efficiency that depends on both* τ_{3D} *-the time the TF spends diffusing in 3D between two nonspecific binding events- and* τ_{ns}*, the time spent sliding on DNA. Figure from 'The needle and the haystack: SMT to probe the transcription factor search in eukaryotes', open access article (Mazzocca* et al*, 2021b).*

Why facilitated diffusion is such an efficient strategy? Facilitated diffusion is an intermittent process where the TF alternates periods of 3D random diffusion with one-dimensional (1D) diffusion on DNA (i.e. sliding (Mirny *et al*, 2009)). DNA sliding (which involves nonspecific TF-DNA interactions) allows a TF to move away from unproductive sites while extensively sampling an entire DNA segment to locate specific REs (Figure 3.1B). When detached from DNA, the TF can freely move in the 3D space, covering long distances in short time. This alternating behavior reduces

the τ_{search} for several reasons. First, while moving by 3D diffusion, the TF is not compelled to find small REs with base-pair precision, but -instead- it can land on 'extended' DNA targets (i.e. the DNA piece that will be subsequently scanned). Second, by sliding, a TF instead of exploring all the 3D available space, is constrained on one dimension (which is actually the dimension containing its REs), speeding up the search. Finally, sliding allows the TF to efficiently escape from random DNA: otherwise, TFs would associate/dissociate multiple times with nonspecific targets until they find a RE, making the search very slow (Berg *et al*, 1981).

Many bacterial proteins have been found to slide on DNA (Esadze & Stivers, 2018b; Silverstein *et al*, 2014; Cravens *et al*, 2015), but whether also eukaryotic TFs can perform facilitated diffusion is unknown. Reports suggest that some TFs can perform DNA sliding (including p53) in vitro, but evidence in vivo are lacking and may be challenging for technical limitations (Elf & Barkefors, 2019). Importantly, however, strong evidence points out that nuclear proteins do not obey the assumptions of the Smoluchowski equation. For instance, when measuring the mean squared displacement (MSD) of proteins in living nuclei (a measure of how far molecules move over time from an initial position), they often appear not to diffuse as expected by random (i.e. Brownian) motion (Woringer & Darzacq, 2018b; Pederson, 2000; Fritsch & Langowski, 2010). While the MSD of random diffusing molecules increases linearly with time (MSD (τ)), the MSD of TFs typically increases less (and it is said subdiffusive), following a function of the type MSD (τ^{α}), with α<1 that is the anomalous diffusion exponent (Woringer & Darzacq, 2018b; Mardoukhi *et al*, 2015).

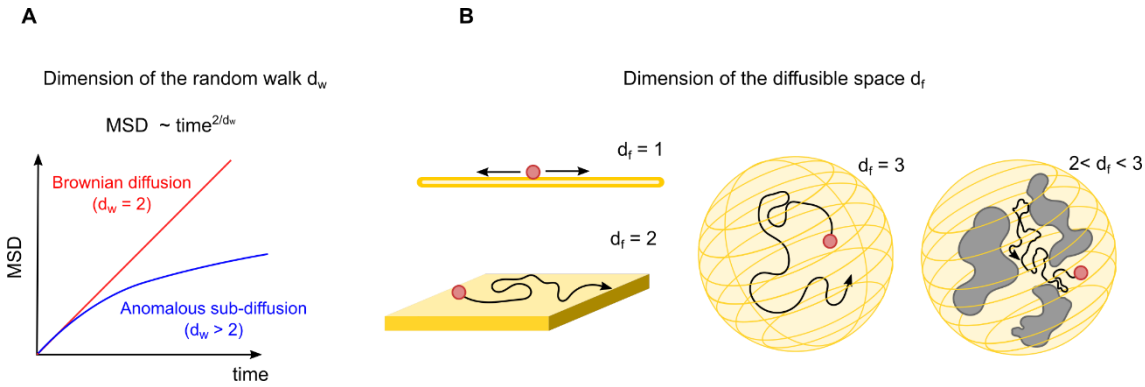


Figure 3.2 Exploration strategies in the eukaryotic cell nucleus. *The TF exploration strategy can be determined by two numbers, the walk dimension* $\boldsymbol{d_w}$ *and the diffusible space dimension* $\boldsymbol{d_f}$*. (A)* $\boldsymbol{d_w}$ *can be extracted by the analysis of the mean squared displacement of a TF and is equal to 2 for Brownian diffusion and larger than 2 for anomalous diffusion. (B) The dimensionality of the diffusible space* $\boldsymbol{d_f}$ *is equal to 1 for sliding on DNA, 2 for diffusion on a surface, 3 for unhindered diffusion and might have values between 2 and 3 for diffusion in a crowded environment or in fractal structures. Figure from 'The needle and the haystack: SMT to probe the transcription factor search in eukaryotes', open access article (Mazzocca* et al*, 2021b).*

This condition where TFs diffuses away from its initial position at slower rates than expected, highlights search strategies different from random motion and, more importantly, it directly impacts on the τ_{search}. Recent theoretical works have demonstrated that the search strategy adopted by a TF can be entirely determined by two parameters, the random walk dimension (d_w) and the diffusible space dimension (d_f; (Shlesinger, 2007; Bénichou *et al*, 2010)). d_w is associated with the anomalous diffusion exponent α, by the relationship $d_w = 2/\alpha$. In other words, when the TF is subdiffusive, $d_w > 2$. For example, in human nuclei, the elongation factor P-TEFb displays a subdiffusive behavior ($d_w \sim 3.3$), while the oncogenic transcription factor c-Myc largely diffuses by Brownian motion ($d_w = 2$; (Izeddin *et al*, 2014a)).

The parameter d_f -instead- is related to the diffusible space (i.e. the space where diffusion is allowed; (Woringer & Darzacq, 2018b)). For 3D random motion, $d_f = 3$, because the TF can move along all dimensions, whereas when the TF is constrained on less than 3D dimensions, it is said to move on a reduced dimensionality (e.g. for DNA sliding, with the TF linearly moving $d_f = 1$; Figure 3.2B (Woringer & Darzacq, 2018b; Woringer *et al*, 2014)).

In practical terms, independently on the molecular mechanisms underlying the search process, the combination of d_w and d_f defines two universal classes of diffusion, namely compact and non-compact exploration, which incorporate all possible diffusion mechanisms (e.g. facilitated diffusion), with different search efficiencies. The exploration is compact when $d_w > d_f$, and non-compact when $d_w < d_f$ (Figure 3.3; (Shlesinger, 2007; Bénichou *et al*, 2010; Meyer *et al*, 2012)). Exploration is non compact for example in case of 3D random diffusion (with $d_w = 2$ and $d_f = 3$, $d_w < d_f$), and compact in case of sliding ($d_w = 2$ and $d_f = 1$, $d_w > d_f$). Most importantly, only compact or non-compact exploration ultimately modulates the τ_{search} defining the search efficiency (Bénichou *et al*, 2010; Meyer *et al*, 2012).

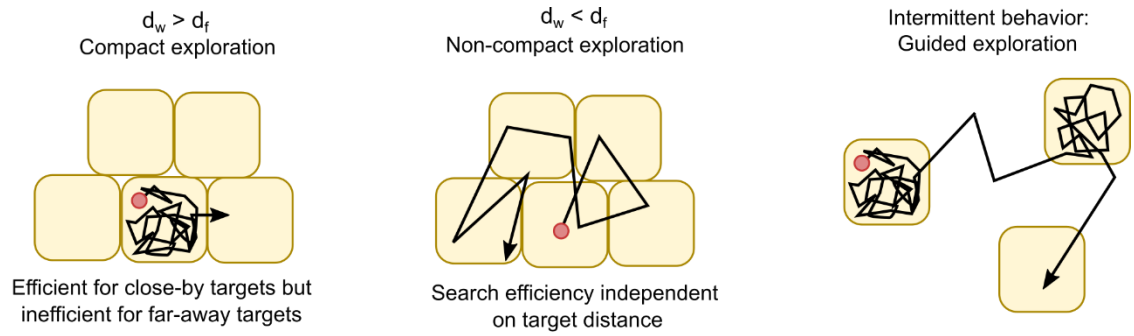


Figure 3.3 The dimension of the random walk d_w and the dimension of the diffusible space d_f define the search strategy. *Examples of how d_w and d_f influence the search process: compact exploration (left) leads TFs to entirely explore confined subregions, while these subregions are only rapidly visit by TFs undergoing non-compact exploration (middle).*

On the right an intermittent behavior, arising from values of d_w and d_f changing through time. Figure from 'The needle and the haystack: SMT to probe the transcription factor search in eukaryotes', open access article (Mazzocca et al*, 2021b).*

What is the effect of compact and non-compact exploration on τ_{search}? By compact exploration, TFs get confined in subregions (e.g. a DNA stretch in case of sliding) that they accurately sample. When this happens, TFs are likely to visit all the sites within a subregion before escaping. This has relevant consequences. If such region is enriched with genes, compact exploration increases the TF probability to find them. If instead there are no genes, the compact search is uselessly time consuming. This also implies that the τ_{search}, of compact exploration is distance-dependent (i.e. small in case the gene is nearby, bigger when target is far from the TF). On the contrary, by non-compact exploration TFs move without being constrained in subregions. On average, they rapidly diffuse from one region to another without completely explore them. The τ_{search} of non-compact diffusion is distance independent: the TF has equally low probability to find a nearby gene or one far apart, and the search time depends only on the diffusion coefficient of the TF and on the size of the target, as in the Smoluchowski relationship.

It is straightforward to note that the search efficiency is maximized when the TF undergoes compact exploration within gene-enriched regions, and non-compact in regions devoid of genes. A clear application of this concept is offered once again by facilitated diffusion, a search mechanism that is neither uniquely compact nor non-compact. Facilitated diffusion is an alternate search where TFs rapidly move through non-compact exploration in the 3D space devoid of genes landing on DNA stretches (where genes are of course more likely) and switching to compact exploration (by 1D sliding). Indeed, it can be theoretically demonstrated that facilitated diffusion maximizes the search efficiency when the factor spends half of its time in its 'non-compact' search mode (3D diffusion) and the other half in its compact search mode (1D sliding).

3.2 Factors influencing the search mechanism

Facilitated diffusion offers a formidable example of how proteins can efficiently locate their DNA targets, from both a molecular and a theoretical stand point. Yet, it may represent one among several mechanisms that TFs adopt to find their gene. This seems particularly plausible for eukaryotic TFs, considering the highly compartmentalized conformation of the cell nucleus, crowded with molecules (e.g. chromatin, RNAs) that occupy defined spaces (Misteli, 2007, 2020; Amiad-Pavlov *et*

al, 2021; Quinodoz *et al*, 2021). As we have seen, the search mechanism is modulated by the random walk (d_w) and the diffusible space (d_f) dimensions. In its most basic form (3D random motion), the search is non-compact ($d_w<d_f$). We thus consider all the factors that may push towards compact exploration ($d_w>d_f$), either by increasing d_w or reducing d_f. Many factors that we will illustrate here are hypothetical and speculative in nature, given that this field is still at its infancy. Yet, we think that referring to specific biological features may be useful to conceptualize the problem.

As mentioned d_w reflects the degree of sub-diffusion (the more a TF is subdiffusive, the higher d_w). How do TFs may become sub-diffusive? A typical example is provided by molecular crowding. The nucleus is a dense organelle (~100 mg/mL in euchromatin ~200–400 mg/mL in heterochromatin (Daban, 2000; Bohrmann *et al*, 1993)), filled with plenty of molecules and sub-compartments (e.g. DNA, nuclear bodies). Macromolecular crowding is a phenomenon due to the high concentration of macromolecules (e.g. chromatin) which -by limiting the space available for the smaller TFs- constrains their diffusion. Another important factor that may determine sub-diffusion is given by weak interactions that arise between TFs and the nuclear environment (Woringer & Darzacq, 2018b). A hypothetical example may be provided by molecular clusters of TFs and coactivators (e.g. Mediator, Brd4), as those suggested to form at enhancers (Boija *et al*, 2018; Sabari *et al*, 2018). These clusters could provide a favorable environment where a 'searching' TF transiently gets in touch with many different molecules, through weak interactions that deviate its free motion to subdiffusive behaviors.

How can be reduced d_f instead? As said, d_f is the dimension of the diffusive space. A TF explores a space of reduced dimensionality when it is obliged to move on less than 3D dimensions (e.g. 1D sliding). Here speculations may move from the observation that the nucleus is highly compartmentalized. For instance, nucleosomes often arrange in large domains (or clutches (Ricci *et al*, 2015; Strickfaden *et al*, 2020)). In principle, a TF forming chemical contacts with a clutch may reduce dimensionality of the search, exploring the clutch surface (2D). More generally, every time that a TF encounters an interacting compartment, two (not mutually exclusive) possibilities are plausible: the TF is allowed to diffuse through the compartment exploring its inside (3D motion), or the TF gets confined on a structure of reduced dimensionality (e.g. the compartment surface ~2D).

3.3 Search mechanism of eukaryotic TFs

We now focus on the evidence accumulated so far on the search mechanism of eukaryotic TFs. We start again from the 'benchmark' among the search strategies, facilitated diffusion. As mentioned, studies in vitro using naked DNA reported that some mammalian TFs (p53, Egr-1) are capable of 1D diffusion through DNA sliding (Tafvizi *et al*, 2008, 2011; Esadze & Iwahara, 2014; Zandarashvili *et al*, 2012). However, whether chromatinized DNA in living nuclei is still available for sliding or whether chromatin constitutes a barrier is unknown. Attempts to corroborate the sliding observed in vitro come from live imaging studies with single molecule tracking (SMT). SMT allows to visualize individual molecules fluorescently labelled and follow their motion in cell nuclei. As we will see, SMT is nowadays the elective approach to study TFs search mechanism, but the direct visualization of facilitated diffusion may be difficult because sliding is thought to occur over few hundreds of nucleotides, corresponding to ten nanometers at most (analogue to the resolution limit of SMT; (Elf & Barkefors, 2019)). Nevertheless, SMT provides kinetic parameters related to the TF search efficiency, that can be used to speculate on facilitated diffusion. A parameter is the amount of TF bound to chromatin. Since chromatin in living nuclei moves remarkably slowly (<0.1 $\mu m^2/s$ (Nagashima *et al*, 2019)), chromatin-bound molecules observed by SMT appear almost immobile. This allows to quantify the ratio of bound and diffusive molecules (usually referred to the bound and free fraction, respectively). Importantly, mutations on the DNA binding domain, drastically reduce the bound fraction of TFs, highlighting the relationship between mobility and TF-chromatin interactions (Mazzocca *et al*, 2021a). The TF bound fraction is a parameter also related to the search efficiency: the more molecules efficiently locate their targets on chromatin, the higher the expected bound fraction. For this reason, the bound fraction has been used to indirectly probe facilitated diffusion model in mammals (Chen *et al*, 2014). For example, previous studies on p53 showed that sliding in vitro is mediated by its disordered CTD (Tafvizi *et al*, 2011). Consistently, deletion of the CTD results in a lower p53 bound fraction in living cells, in line with a model in which -incapable of sliding- less p53 molecules reach their REs (Mazza *et al*, 2012a). Of course, these correlative observations need of further investigation, because it is still not demonstrated that the actual role of the p53 CTD on the search mechanism in vivo is to promote sliding.

Beyond facilitated diffusion, growing evidence point out other possible search strategies for eukaryotic TFs and, very often, a strong dependence on nuclear organization. Many recent findings derive from SMT studies, highlighting again the

importance of this approach (Garcia *et al*, 2021a; Izeddin *et al*, 2014b; Hansen *et al*, 2020). As discussed, the TF search process is defined by compact ($d_w > d_f$) or non-compact exploration ($d_w < d_f$). Even if only d_w can be readily measured by SMT (see later the paragraph **5.1 TF search investigated by Single molecule tracking (SMT)**), compact and non-compact exploration give rise to distinct diffusive behaviors. A TF involved in compact exploration, over time it visits repeatedly a certain subregion, appearing confined, whereas in the non-compact exploration TFs freely move with no constraints. Interestingly, several SMT studies have begun to investigate the nuclear mobility of diffusive TF molecules and the degree of their confinement (Garcia *et al*, 2021a; Nguyen *et al*, 2021; Hansen *et al*, 2020; Kent *et al*, 2020; Lerner *et al*, 2020; Izeddin *et al*, 2014b). It results that diffusing TFs very often get at least transiently confined, within subnuclear compartments not always well identified. For example, a remarkable study on ten components of the pre-initiation complex (PIC) -including Pol-II, general transcription factors (GTFs) and the coactivator Mediator- reported that several GTFs localize their target sites by confined exploration of nuclear subregions (Nguyen *et al*, 2021). Importantly, the confined behavior seems to be regulated by key members of the PIC (e.g. Mediator), indicating a cooperative mechanism among nuclear factors in the target process. Another study showed that CBX2 (a component of the chromatin repressive complex Polycomb) can form nuclear condensates providing 'baits' for further CBX2 diffusive molecules. Within condensates, CBX molecules appear to undergo confined exploration, speeding up the τ_{search} (Kent *et al*, 2020).

An important aspect to consider is that not all nuclear proteins display a signature of confined (compact) exploration (Izeddin *et al*, 2014b). This indicates that the search mechanism relies on intrinsic property of TFs. Moreover, among the proteins undergoing compact diffusion, not all behave equally. For instance, the elongation factor P-TEFb adopts compact exploration as a unique search strategy (Izeddin *et al*, 2014b), while the transcriptional regulator CTCF displays an intermittent strategy, alternating between compact and non-compact diffusion, a type of search that has been named 'guided exploration' (Hansen *et al*, 2020). Guided exploration shares with facilitated diffusion this alternate behavior (a TF sometimes freely navigate in the 3D space, sometimes it gets confined). The compact exploration of CTCF however, seems not to depend on sliding but on weak interactions with RNA compartments. This is relevant because it further confirms the variety of molecular mechanisms underlying compact exploration, and the strong relationship between diffusion of TFs and nuclear architecture.

We have seen the factors that can influence the targeting efficiency of TFs in living cells. Giving the strong influence that the environment plays on the TFs motion, we will describe next how the cell nucleus, with its many organizational layers, may regulate the search process of eukaryotic TFs.

4 Influence of the nuclear organization on TF search mechanism and the physical accessibility of genes

As discussed, the nuclear environment shapes the mobility of TFs and -in addition- the accessibility of genes. In this chapter we analyze how the cell nucleus may impact on the TF search and the targeting of genes. Far from being a homogeneous mixture of molecules, the mammalian nucleus is a complex organelle divided into specialized compartments with different structure and function (Misteli, 2007, 2020; Amiad-Pavlov *et al*, 2021; Quinodoz *et al*, 2021). This spatial division is thought to ensure a correct coordination in space and time of a large variety of nuclear processes (Misteli, 2007, 2020). For example, reactions like transcription and DNA replication do not take place everywhere in the nucleus but occur on discrete sites, in which the required factors (e.g. RNA-polII) accumulate to perform their functions (Misteli, 2007; Sutherland & Bickmore, 2009; Papantonis & Cook, 2013). DNA itself appears not randomly distributed but occupies regions depending on its functional state (Osborne *et al*, 2004; Bizhanova & Kaufman, 2021; Finn & Misteli, 2019; van Steensel & Furlong, 2019), with the transcriptional active euchromatin that physically segregates from the inactive heterochromatin, occupying distinct compartments (Bizhanova & Kaufman, 2021; van Steensel & Furlong, 2019).

How does such compartmentalized organization influence the search and selectivity of TFs? We will see that, despite the long notion that inactive genes often occupy heterochromatin regions, it is still poorly understood whether their localization within dense heterochromatin is necessary to prevent TF to find and regulate them.

As discussed earlier, TFs navigate in the nuclear space and encounter many different nuclear substructures that can establish chemical and physical interactions (e.g. DNA sliding and molecular crowding, respectively) tuning their mobility. Deviations from 3D random diffusion to compact exploration modulate the TF search mechanism and the τ_{search}.

With a focus on p53, we will review how the compartmentalized nucleus shapes TF activity. We will follow a zoom-out approach that starts from the nanoscale lengths -with nucleosomes- and gradually moves to higher structures up to the microscale level. In particular, we will see how such hierarchical organization impacts (i) on the TF mobility -with possible consequences on the type of exploration (i.e. compact vs non-compact), and (ii) on the physical accessibility of genes in terms of availability to be 'visited' and bound by TFs during their diffusion process.

4.1 Nucleosomes: the first obstacle for TFs

The diploid human genome with its 46 chromosomes is composed of ~6.5 billion base pairs (bp), corresponding to a physical size of about 2 meters of DNA (Piovesan *et al*, 2019). To fit inside a remarkably small nucleus, DNA is incorporated into histone octamers -complexes composed by two copies of each histone proteins H2A, H2B, H3, and H4 (Luger *et al*, 1997; Richmond & Davey, 2003)- forming nucleosomes. In a nucleosome, ~147 DNA nucleotides spontaneously wrap on the surface of a histone octamer through electrostatic attraction between DNA (negative charged for the phosphate groups of its backbone) and histones (enriched by positive charges for their basic amino acids (Parmar *et al*, 2019; Davey *et al*, 2002)).

This strong association (with a net free energy of ~40 k_BT for a single nucleosome (Parmar *et al*, 2019)) is the first potential obstacle to TFs because it may hamper TF-DNA interactions. Indeed, despite the genomic regions poorly occupied by nucleosomes are only the 2-3% of the human genome, they collect about the 90% of all TFs associated with chromatin (Thurman *et al*, 2012), indicating that nucleosomes tend to exclude TFs.

Nucleosomes potentially affect two sides of the search mechanism. On one hand they hide genomic REs preventing TFs to find their targets. On the other side, they may create a barrier to TF sliding. We start from considering how nucleosomes can obstruct the association of TFs to their REs, analyzing the problem in general terms and then focusing on the specific case of p53. Later, in a separated section, we will discuss the sliding 'problem' and more generally the effect of nucleosomes on the search mechanism.

4.1.1 Influence of nucleosomes on the accessibility of genomic REs

To better understand why nucleosomal DNA is poorly accessible to TFs, we need to consider the interactions that histones and DNA establish in a nucleosome. Overall, the histone octamer forms 14 contacts (e.g. by hydrogen bonds) with the DNA minor groove, and no more than ~4 consecutive DNA bp in a nucleosome remain exposed to solvent (Luger *et al*, 1997). Since most TFs bind DNA motifs of ~6–12 nucleotides (with the p53 RE being 20 bp (Bilu & Barkai, 2005)), this evidence alone might explain why nucleosomes often modulate the association of TFs to their genes. Nevertheless, the strength of the octamer-DNA interactions is not the same at each position, with stronger bonds close to the nucleosome symmetry axis (known as dyad), and weaker at enter and exit sites of DNA from a nucleosomes (Luger *et al*, 1997). For their weaker nature, the interactions at the entry/exit sites are perturbed by thermal

fluctuations, inducing DNA to temporarily dissociate from the histone octamer (Polach & Widom, 1995). These transient dissociations (or DNA 'breathing') have important consequences, since they expose the unwrapped nucleotides to the solvent and make them available for the association with TFs (Polach & Widom, 1995). In vitro measurements estimated that the entry/exit region are exposed to the environment for about the 10-30% of the time, allowing some TFs to bind them (Li *et al*, 2005; Tims *et al*, 2011; Wei *et al*, 2015). However, moving to closer positions to the dyad, the unwrapping becomes much more infrequent (with a constant rate of 1-10 minutes (Tims *et al*, 2011)), making the more internal sites poorly available to TFs, as confirmed also by the lower TF binding rates on these sites (Tims *et al*, 2011; Wei *et al*, 2015).

The remarkable inaccessibility of wrapped DNA severely limits the binding of TFs in vivo (Thurman *et al*, 2012; Michael & Thomä, 2021), except for a particular category of TFs called 'pioneers' that can bind their REs directly on nucleosomes (Zaret & Mango, 2016). Pioneer TFs are often associated with embryonic development (e.g. FOXA1 required for liver differentiation (Lee *et al*, 2005); SOX2 involved in the maintenance of pluripotency (Boyer *et al*, 2005)), and use their nucleosomal affinity to bind sequences inaccessible to other TFs, giving rise to a chromatin remodeling cascade (Zaret & Mango, 2016). Very interestingly to us, also p53 is a pioneer factor (Michael & Thomä, 2021), as documented by the extensive evidence on p53 binding nucleosomal DNA with high affinity, both in vitro and in vivo (Nili *et al*, 2010; Sammons *et al*, 2015; Yu & Buck, 2019). Initial studies using reconstituted chromatin and nucleosomal binding assays, showed that p53 binds to the *CDKN1A* promoter in presence of nucleosomes (Espinosa & Emerson, 2001; Laptenko *et al*, 2011). A genome-wide work based on chromatin immunopreciptation followed by microarray analysis (ChIP-chip), revealed that this behavior is not restricted to few genomic sites, and that p53 in cell nuclei preferentially associate with regions densely occupied by nucleosomes (binding to ~2000 of these regions (Nili *et al*, 2010)). This finding was further explored by ChIP-seq studies (Sammons *et al*, 2015; Younger & Rinn, 2017), confirming that at the genome-wide level, p53 binds especially to regions enriched by nucleosomes, highlighting that its pioneer capabilities are broadly employed in vivo.

What is the molecular mechanism allowing p53 to associate with nucleosomal DNA? Not all sites on a nucleosome are likely to engage interactions with p53. Indeed, several studies showed that p53 preferentially binds to the nucleosome edges, where DNA entry/exit sites reside (Yu & Buck, 2019; Laptenko *et al*, 2011). Whether this preference is given by the transient DNA unwrapping observed at these regions, is

unknown yet. Importantly, binding to nucleosomal DNA highlights once again the importance of the disordered p53 CTD, that plays a critical role also in this context. While the full-length p53 can bind to nucleosomal sequences even if lacking specific p53 REs (or, in other words, the full-length p53 can engage interactions with nonspecific nucleosomal DNA sequences), the deletion of the CTD strongly reduces the association with nucleosomes (Yu & Buck, 2019; Sahu *et al*, 2010), reaffirming the importance of the CTD in mediating p53 nonspecific interactions.

The findings on the 'pioneer' capabilities of p53 are important because suggest that nucleosomes do not negatively affect p53 binding activity. This implies that p53 reaches and associates with even apparently inaccessible genes (inaccessible from the nucleosome perspective, Figure 4.1). Thus, other factors may be responsible in determining p53 selectivity towards its target genes.

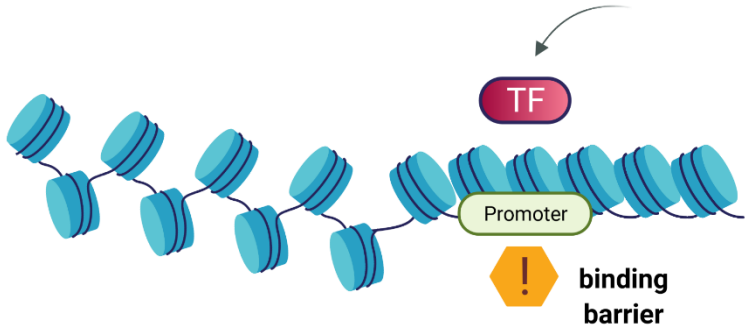


***Figure 4.1 Nucleosome impeding TF to associate with a promoter**. Nucleosomes may hamper the association of TFs to their DNA targets, except for a class of TFs (called pioneer factors) that can bind also nucleosomal DNA. Image created with BioRender.com.*

4.1.2 Influence of nucleosomes on the TF search mechanism

The impact of nucleosomes on the TF search mechanism, and in particular whether nucleosomes obstruct TFs that need to scan the genome to find their target genes, has been poorly elucidated so far.

A SMT study in yeast, indicated that nucleosomes slow down the search process of the TF Ace1p, reducing its target efficiency (Mehta *et al*, 2018). Consistently, by removing nucleosomes, the chromatin remodeler RSC creates a more accessible environment to Ace1p facilitating its exploration on DNA and speeding-up the targeting process.

Given the tight DNA wrapping on histones, a major question is whether TFs can perform sliding on nucleosomal DNA (Figure 4.2). Studies indicate that several proteins (including p53) slide by transient interactions with the DNA major groove (Hauser *et al*, 2016; Givaty & Levy, 2009; Khazanov & Levy, 2011). Since histones

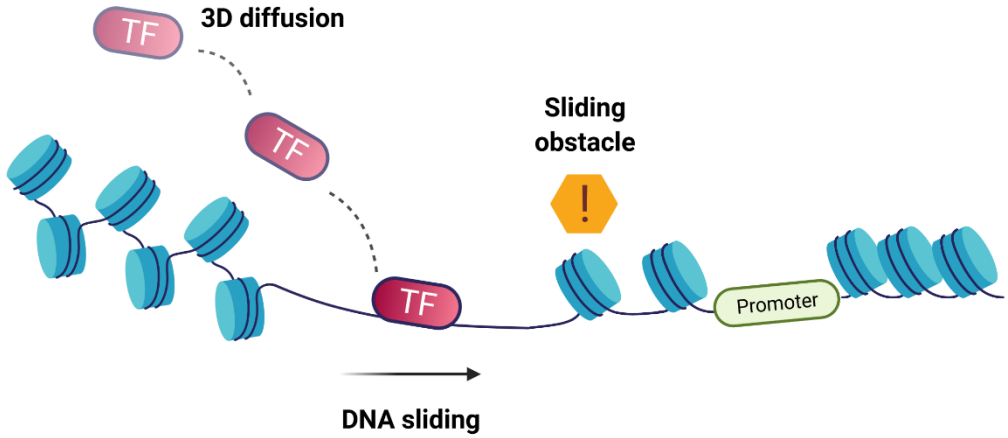


***Figure 4.2 Nucleosomes hampering TF sliding towards a promoter**. While studies shown that several TFs can slide on naked DNA, it is still unknown whether sliding may be obstructed by nucleosomes, preventing TFs to reach their promoters. Image created with BioRender.com.*

bind the minor groove (Luger *et al*, 1997), if for a moment we ignore steric hindrance factors, we may assume that TFs use the major groove as a 'rail' for sliding, disregarding nucleosomes on the other side. To our knowledge, however, evidence of TFs sliding on nucleosomes is lacking, and all the current data on p53 have been obtained in vitro on naked DNA (McKinney *et al*, 2004; Tafvizi *et al*, 2008, 2011; Leith *et al*, 2012; Murata *et al*, 2017; Subekti *et al*, 2020). This is somehow surprising given the documented affinity of p53 for nucleosomal DNA (Nili *et al*, 2010; Sammons *et al*, 2015). On the other hand, the lacking of data may imply that p53 cannot slide on nucleosomes in vitro, suggesting that other factors -perhaps available in the cell nucleus- may be necessary. For example, since histone PTMs (e.g. acetylation) can accelerate DNA unwrapping from nucleosomes and increase its accessibility (Fierz & Poirier, 2019; Michael & Thomä, 2021), the sliding process may be facilitated by cofactors (e.g. histone acetyltransferase) that act in concert with TFs opening dense nucleosomal DNA sequences. According to this hypothesis, we mentioned earlier that p53 IDRs (both the N-terminal and the C-terminal domains) recruit acetyltransferases (e.g. CBP): in principle, p53 could exploit these cofactors to make nucleosomal DNA more accessible for sliding. In addition, p53 recruits HMGB1 (Štros *et al*, 2004; Schröder *et al*, 2019; Rowell *et al*, 2012, 1) and PC4 (Batta & Kundu, 2007), architectural proteins that induce DNA bending (Štros, 2010; Garavís & Calvo, 2017), another factor promoting DNA to unwrap from nucleosomes (Fierz & Poirier, 2019; Michael & Thomä, 2021). Interestingly, even if the p53 affinity for HMGB1 and PC4 is relatively low (compared to other cofactors), HMGB1 and PC4 are so highly abundant in the cell nucleus that theoretical calculations suggest they may be

amongst the main p53 partners (Teilum *et al*, 2021). This is of note because p53-HMGB1 and p53-PC4 interactions have not been widely explored in the literature, and it is unknown whether these cofactors may cooperate for the p53 search mechanism.

Finally, we highlight that it is not easily predictable the general impact of nucleosomes on the TF search, because they could play not only a negative effect (limiting sliding) but also a positive role by reducing the accessibility of random DNA. Since nucleosomes bind the vast majority of the genome (~85-90%; (Michael & Thomä, 2021)), they also massively mask nonspecific DNA sequences, likely preventing TFs from being sequestered.

Overall, the data about how p53 by be influenced by this first level of chromatin organization, suggest that nucleosomes do not negatively affect p53 binding and gene selectivity. These pioneer properties are at least in part mediated by the disordered p53 CTD, but it is unknown whether also the disordered N-terminal TADs may cooperate. It is still uncertain whether p53 can slide on nucleosomal DNA in vivo, possibly accelerating the search process, and what factors may facilitate sliding on the histone octamer.

In the following paragraph, I will illustrate an additional layer of chromatin organization, provided by the hierarchical folding of nucleosomes in the 3D space.

4.2 Chromatin folding into higher order structures

4.2.1 Nucleosomes spontaneously folding in the 3D space

Nucleosomes can spontaneously arrange into 3D hierarchical structures depending on chemical and physical properties of the surrounding environment (e.g. ionic strength; (Strickfaden *et al*, 2020; Maeshima *et al*, 2016b, 2020)). On a histone octamer, only approximately half of the DNA negative charges are shielded by histones, determining a strong repulsion between adjacent nucleosomes (Hansen, 2002; Maeshima *et al*, 2020). In vitro, to minimize their repulsion, nucleosomes dispose along a stretched array, resembling 'beads on a string' (known as the 10-nm fiber for its diameter; (Baldi *et al*, 2020; Maeshima *et al*, 2014; Olins & Olins, 1974)). Increasing salt concentration (and mimicking so the nucleoplasm environment), available cations screen the DNA negative charges, inducing nucleosomes to move closer and to fold along the three dimensions (Maeshima *et al*, 2020; Hansen *et al*, 2017b). How thus do nucleosomes organize in the 3D space regulating the accessibility to TFs?

For a long time, textbooks have proposed the so called 30-nm fiber as the main chromatin folding, a structure where nucleosomes precisely rearrange into a compact zig-zag (or solenoid) composition of ~30 nm in diameter (Maeshima *et al*, 2020; Hansen *et al*, 2017b; Zheng *et al*, 2005; Kan *et al*, 2009). This model was supported by a large number of studies in vitro (Parmar *et al*, 2019; Dorigo *et al*, 2004; Song *et al*, 2014; Woodcock *et al*, 1984; Schalch *et al*, 2005; Krebs *et al*, 2017). Indeed, solutions at higher ionic strength -but still lower than physiological conditions (e.g., < 1 mM Mg^{2+})- induce nucleosomes to come into tighter proximity, but maintaining an ordered and repetitive structure.

More recently, however, many studies in vitro and on fixed nuclei questioned the existence of the 30-nm fiber (Eltsov *et al*, 2008; Nishino *et al*, 2012; Gan *et al*, 2013; Chen *et al*, 2016; Ou *et al*, 2017; Visvanathan *et al*, 2013; Boettiger *et al*, 2016; Fussner *et al*, 2012; Hsieh *et al*, 2015; Ohno *et al*, 2019). For example, electron microscopy (EM) and X-ray experiments performed on both interphase and mitotic chromosomes found no evidence of this structure (Eltsov *et al*, 2008; Nishino *et al*, 2012; Gan *et al*, 2013; Chen *et al*, 2016; Ou *et al*, 2017). Moreover, in vitro, at more physiological salt concentrations than those seen above, nucleosomes do no adopt any longer the canonical 30-nm configuration but self-assemble into larger globular domains, where an individual nucleosome can interact with any other without a pre-ordered hierarchy (Maeshima *et al*, 2020, 2016b; Strickfaden *et al*, 2020).

In a seminal study using super-resolution imaging, nucleosomes labelled by anti-histones antibodies were shown to form irregular (but discrete) domains of variable size (~30–50 nm), named nucleosome 'clutches' (Ricci *et al*, 2015). A single clutch has been estimated to contain few kilobases of genomic DNA (depending on the clutch size). Individual clutches appeared surrounded by a capillary network of nucleosome-depleted regions. Consistently, a large number of works using unrelated techniques (e.g. EM, super-resolution imaging, etc.), came to similar conclusion, showing that chromatin in cell nuclei is organized into heterogeneous domains, often indicated with different names (e.g. domains, clusters, condensates; (Eltsov *et al*, 2008; Nishino *et al*, 2012; Gan *et al*, 2013; Chen *et al*, 2016; Maeshima *et al*, 2016b; Ou *et al*, 2017; Visvanathan *et al*, 2013; Boettiger *et al*, 2016; Fussner *et al*, 2012; Hsieh *et al*, 2015; Ohno *et al*, 2019)). We take the organization depicted by clutches to analyze the possible influence on the TF search mechanism, but referring the same considerations to all similar 3D structures (i.e. domains, clusters, etc.).

When considering nucleosome clutches, a question arises: have TFs free access and can diffuse throughout? The question is relevant for the search process and -in turn- for gene regulation. For example, a clutch with limited accessibility would create

a barrier to the regulation of genes residing inside, since they cannot be easily reached by TFs. The very low scale lengths involved pose serious technical limitations for investigating whether clutches hamper the search process, however we can speculate on their accessibility by looking at the dynamics of nucleosomes. At the nanoscale level (~10 nm, the size of a single nucleosome), they display liquid properties as highlighted by the high dynamic of histones (Maeshima *et al*, 2016a). This suggests that TFs may freely diffuse inside clutches. However, at scales involving more than a nucleosome (>10 nm) the chromatin fiber acquires solid-like properties, resembling a hydrogel matrix, where individual clutches/domains can aggregate (forming bigger clusters), but cannot mix (Strickfaden *et al*, 2020). This second evidence argues against the free diffusion of TFs and suggest a poorly flexible chromatin scaffold that may limit TF to diffusion inside, with the diffusion probability decreasing as a function of the TF size (Axpe *et al*, 2019; Maeshima *et al*, 2015). In this regard, we need to consider that -when active - p53 is a tetrameric protein, with apparent molecular weight of ~212 kDa.

How thus may behave TFs encountering a clutch in the cell nucleus? The observation that p53, for example, binds with more affinity nucleosomal than naked DNA indicate that -in a first place- clutches may provide scaffold to recruit p53. It needs to be investigated whether, after its nucleosomal association, p53 undergoes diffusion on a reduced dimension either over the clutches surface (~2D) or on DNA (i.e. sliding), to densely sample the entire region searching for promoters (Figure 4.3).

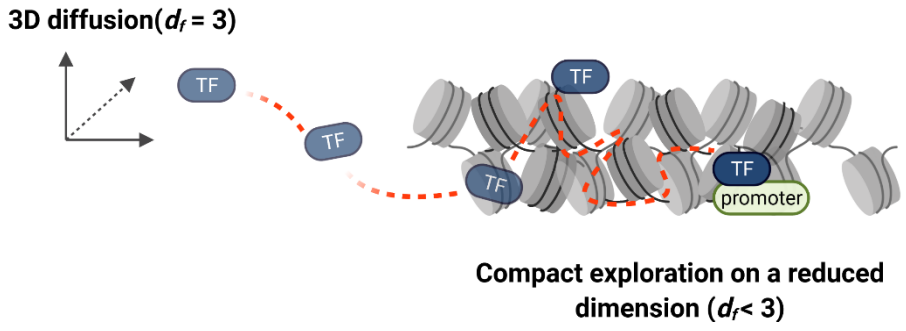


Figure 4.3 Search mechanism strategies to locate promoters in nucleosomal clutches*. A TF diffusing in the 3D promoter lands on nucleosomal clutch, scans the entire clutch surface (exploration in reduced dimensionality), and efficiently locates its promoter. Image created with Biorender.com.*

On the other hand, given the relative low number of nucleosomes involved (Pia Cosma and colleagues estimated a range of 4-8 nucleosomes for clutch (Ricci *et al*, 2015)), the effect played by clutches on the TF search may be strictly associated to

the sliding 'problem' that we discussed earlier. If we assume that pioneer TFs are assisted by coactivators to perform DNA sliding on nucleosomes (with acetyltransferase cofactors opening their way), then we should aspect a direct impact on the clutch structure itself. It seems reasonable that, by relaxing nucleosomes, coactivators may induce also a clutch rearrangement increasing their accessibility inside (Figure 4.4). This is purely speculative and needs a challenging validation.

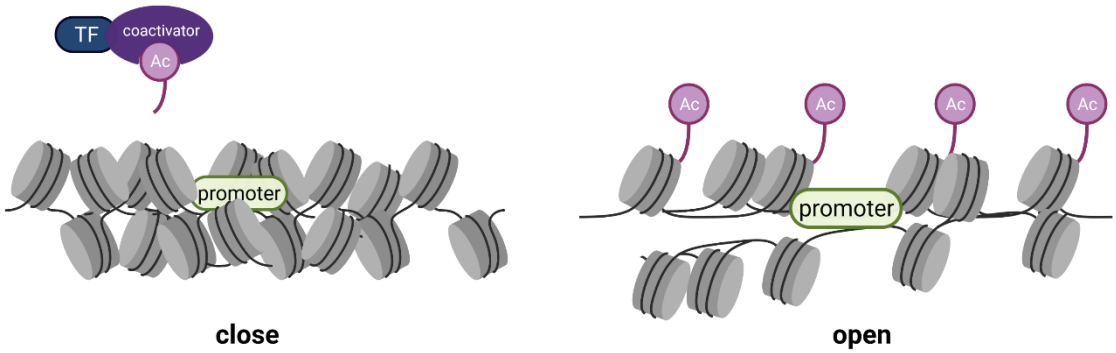


Figure 4.4 Search mechanism strategies to locate promoters in nucleosomal clutches*. TF recruits an acetyltransferase cofactor that opens a dense nucleosomal clutch, allowing the TF to reach its promoter. Image created with Biorender.com.*

Another study on chromatin organization at the nano-scale level was performed by labelling short DNA fragments (by high resolution DNA FISH (Boettiger *et al*, 2016)). Here, it was shown that genomic sequences with equal bp length can form clutches/domains of different size depending on the epigenetic state. In particular, the size was positive correlated to gene expression, with inactive regions occupying lesser space than active ones. However, it is currently unclear whether these larger, less condensed domains are required to facilitate the search and binding of TFs or if they are the result of gene expression. Given that, during transcription, the RNA-polII is supported by chromatin modifiers which displace nucleosomes and relax chromatin (Cramer, 2019), this could explain by itself why the active domains appear bigger. Whether their big size is also required before transcription, to facilitate the search/binding of TFs, is unknown, but theoretical studies suggest that TF diffusion could be slowed by high-density chromatin, affecting the search time (Maeshima *et al*, 2015; Kanada *et al*, 2019).

So far, we have considered the spontaneous 3D organization of nucleosomes, depending on the environmental properties (e.g. ionic strength) or the epigenetic state. However, nucleosomes are highly dynamic and can reposition on different clutches/domains, bringing in contact different DNA sequences. While in vitro this property is stochastic, in vivo it is highly regulated and lead to the formation of the

so-called topologically associated domains (TADs) where genetic loci specifically interact. We will analyze now the possible impact of TADs on the TF search.

4.2.2 Chromatin folding into Topologically Associated Domains (TADs)

At the sub-Mb level, the genome is further folded into non-random topologically associating domains (TADs). TADs are genomic regions -of hundreds kilobases up the Mb- in which distal genomic sequences are likely to interact together, appearing isolated from the rest of the genome (Dixon *et al*, 2012; Rao *et al*, 2014).

The organization of TADs is detected using cell population-based methods (e.g. genome-wide techniques as Hi-C, Micro-C (Szabo *et al*, 2019)). The resulting output represents sort of a snapshot of the genomic contacts occurring across the cell population and reflects the frequency by which a certain genomic interaction is observed. When these interactions are studied by imaging approaches -at the single cell level- to investigate their physical implications (e.g. the distance between two loci, A and B, belonging to the very same TAD), they display a high variability between cells. Nevertheless, TADs offer a picture of the nucleus where genomic sites, independently of whether they forms stable or transient contacts with other regions, are more likely to interact with sequences belonging to the very same TAD than with external sites (Szabo *et al*, 2019), resembling an isolated environment.

What is the mechanism providing specific and isolated interactions? An individual TAD is defined by precise boundaries (characterized by the tight association with the transcriptional regulator CTCF) and contains sequences engaged in long-range contacts (e.g. enhancer-promoter interactions). The formation of the TADs is an active process presumably mediated by loop extrusion, which involve the SMC protein complex and has an energetic cost. The DNA contacts arising within TADs -instead- are generated by loop extrusion, but also thermal fluctuations (Parmar *et al*, 2019; Davidson & Peters, 2021). In this respect, TADs represent confined regions where thermal fluctuations and looping can bring closer distant regions (Figure 4.5, next page).

Here we describe how this organization may impact on the activity of TFs, taking into account two main features of TADs. First, we analyze the effect of specific interactions between two genomic loci, with the paradigmatic case of enhancer-promoter (E-P) contacts. Second, we try to discuss the impact of the whole frequency of contacts that a given genomic site can engage within a TAD.

A longstanding question about E-P interactions is how they mechanistically regulate transcription. It is still poorly understood whether the interaction guides TFs

from enhancers to their promoters, and whether an actual physical contact is required.

Forcing E-P interactions at the *Hbb* locus has demonstrated to increase transcription of hemoglobin (Bartman *et al*, 2016), but there are also examples of enhancers that induce gene expression without any E-P contact (Schoenfelder & Fraser, 2019).

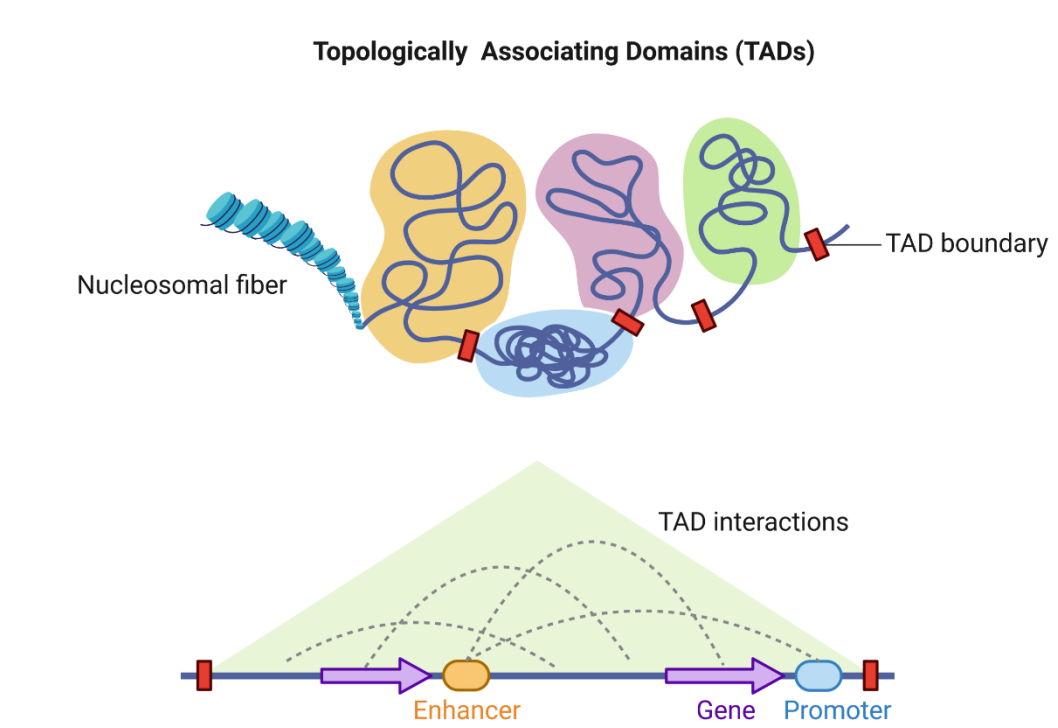


Figure 4.5 Chromatin organization in Topologically Associating Domains. *TADs are organizational units with specific boundaries, similar to isolated environments. Inside TADs, linear genomic sequences are brought in contact through 3D interactions (e.g. enhancer-promoter looping). Image created with BioRender.com.*

Recently, the literature focused on the interesting observation that at enhancers, TFs can stimulate transcription of long non-coding RNAs called enhancer RNAs (eRNAs). It is proposed that eRNAs favor chromatin looping, allowing the interaction with promoters (Li *et al*, 2016). Among many TFs, there is strong evidence that p53 binds at enhancers regions (Nguyen *et al*, 2018) and regulates the production of eRNAs (Melo *et al*, 2013; Link *et al*, 2013). Moreover, in a couple of works it has been shown that p53 can interact with a specific non-coding RNA, that in turn increases p53 binding at the *CDKN1A* promoter (Marney *et al*, 2021; Schmitt *et al*, 2016). Despite a molecular model is still to be defined, this suggests that RNAs may facilitate the targeting of TFs at specific genes.

We now analyze the effects generated by the whole contacts that a locus may experience within a TAD. As mentioned, the contact frequency reflects the amounts of regions that -depending on chromatin organization- a locus can encounter over time. This may have a strong impact on the TF search. Such issue has been mostly characterized by computational studies (Avcu & Molina, 2016; Cortini & Filion, 2018) showing that TFs could find more rapidly their REs on genomic regions at high connectivity (i.e. sites engaging chromatin loops, or making contacts by thermal fluctuations). This phenomenon is explained by models that assume TFs alternating between compact and non-compact exploration (specifically: sliding and 3D random motion). More in detail, TFs diffusing on the 3D space, land on DNA stretches that are exhaustively scanned by sliding. After this local exploration TFs dissociate, but the high connectivity (i.e. the dense presence of other DNA fibers) increases the TF likelihood to encounter other sequences to start a new local exploration. From the perspective of p53 this particularly intriguing, given its peculiar 'dimer of dimers' structure, where the two dimers can bind separate DNA sequences, independently from each other. We speculate that this may allow p53 to slide through one dimer only, while keeping the other 'free', readily available to make contacts with other DNA fibers on the surrounding space. This condition may ensure that as soon as the exploration of a DNA stretch is completed, p53 can hop through the other dimer on a new fiber, re-starting scanning.

All the levels of chromatin organization that we considered up to this section, allowed as to examined the 'search' problem with a specific focus on the local environments, providing potential structures for the TF compact exploration or barriers to the search mechanism. However, the TF search, with different combinations of non-compact and compact exploration, is a process involving the entire nucleus. In the next paragraph, we will try to provide a zoomed-out picture of the search process in relation to how chromatin is globally distributed in the nucleus.

4.3 Chromatin spatial distribution in the cell nucleus: the Cremer's model

Chromatin is the most abundant constituent of nuclei. Its hierarchical folding determines a compartmentalized spatial distribution that shapes the global nuclear architecture. Since the search process involves TF molecules navigating the entire nuclear space, this compartmentalization allows examining the search problem as a whole.

We start highlighting some elements of non-randomness in chromatin spatial distribution. Interphase chromosomes occupy preferential regions called chromosome territories (Misteli, 2020). In addition, active and inactive regions belonging to the very same chromosome, segregate from each other, occupying themselves distinct spaces. For example, silent heterochromatin typically associates with the nuclear periphery, while euchromatin points toward the nuclear inner (Misteli, 2020). Despite this non-random organization, interphase chromatin observed at the micron-scale level resembles an intricated tangle of fibers (Chen *et al*, 2016). Such heterogeneous distribution gives rise to chromatin 'domains' of different size and density, likely reflecting their nanoscale organization (e.g. different aggregations of nucleosomal clutches). Importantly, chromatin 'domains' are interspersed with low-density chromatin regions branching throughout the nucleus (Ricci *et al*, 2015; Ou *et al*, 2017). All these elements have been incorporated in a unified nuclear model theorized by Thomas and Christoph Cremer (Cremer *et al*, 2020). According to this model, the cell nucleus is described by two main structural and functional components, namely the chromatin domains (CDs) -enriched by DNA- and the 'interchromatin' compartment (IC), a reticulum of nuclear spaces devoid of chromatin. In this theorization, genes reside at the boundary of CDs, while IC provides a 3D network of channels connected with the nuclear pores, and allowing macromolecules (TFs, RNAs) to rapidly reach their site of action (Figure 4.6).

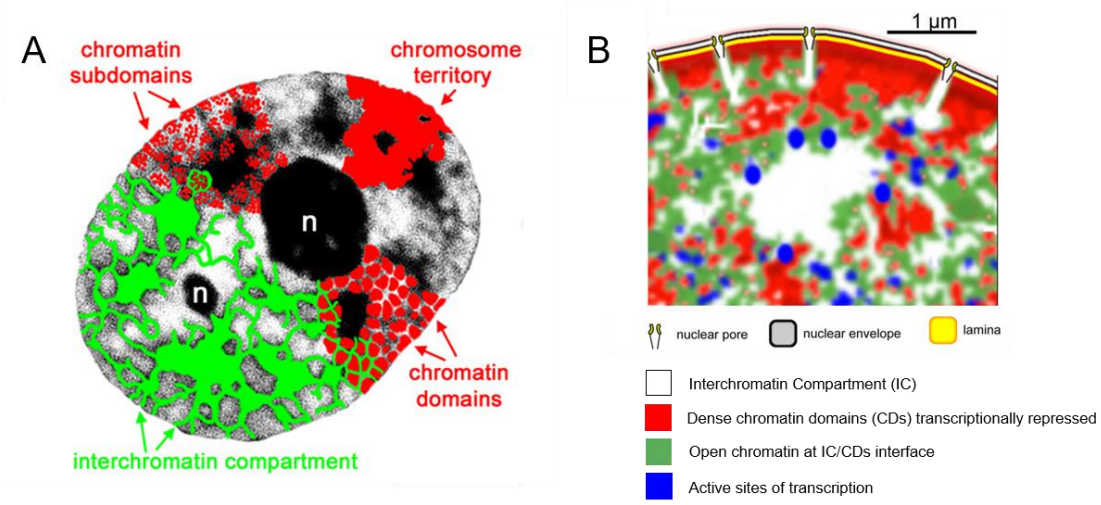


***Figure 4.6 Cremer's model of the cell nucleus**. (A) The global distribution of chromatin defines nuclear regions of different density. Poor-chromatin regions are connected to each other forming the interchromatin compartment (IC), allowing for the diffusion of molecules to their targets on denser chromatin domains (CDs). (B) Zoom-in of the nucleus, highlighting regions with different functions depending on chromatin permeability. The IC connected with nuclear pores regulate the molecular import/export of the nucleus. Dense CDs contain in their inside repressed genes, while transcription takes place in open chromatin regions at the boundaries of IC and dense CDs. Image adapted from 'The Interchromatin Compartment*

Participates in the Structural and Functional Organization of the Cell Nucleus', open access article (Cremer et al*, 2020).*

A main feature of the model involves the position of genes, strictly related to their functional state. Repressed genes occupy the deep and inaccessible interior of CDs, whereas active genes are positioned just outside, at the CD/IC interface directly accessible from the IC channels. This location should facilitate gene regulation. In particular, the IC would offer quick corridors to TFs for reaching their target genes, while inactive genes remain hidden to the transcriptional machinery inside CDs. The newly synthesized mRNAs, transcribed at the CD/IC interface, can exploit the IC corridors for exiting the nucleus.

Some propositions of the Cremer's model have been successfully tested. For instance, a recent work using super-resolution microscopy demonstrated that RNAs in the nucleus spatially segregates from DNA, and accumulate within the IC (Miron *et al*). However, it is still unknown if also TFs use this nuclear 'topography' in their search process, exploiting the IC as the main route to diffuse along distant sites and reach their genes. Yet, we can speculate on the potential consequences of the model. Here we analyze the problem with considerations on the search process and gene positioning.

4.3.1 How the Cremer's model could facilitate the TF search

Overall, the model is based on the assumption that TFs preferentially fill the IC channels, reflecting a non-uniform permeability of the nucleus that depends on chromatin density. Is it this a plausible view? Experimental evidence argues in favor of this assumption. For instance, imaging experiments showed that dense chromatin regions exclude inert tracers, depending on (i) chromatin density and (ii) the molecular size of the tracers (Görisch *et al*, 2005; Bancaud *et al*, 2009). On the same line, the current evidence that nascent mRNAs are enriched at boundaries of chromatin domains, is in agreement with the hypothesis that the big transcriptional machinery (assembling the pre-initiation complex, PIC) cannot penetrate the dense CDs. What are the implications on the TF search that precede PIC assembly (e.g. by p53)? If dense CDs are largely impermeable, they would exclude TFs from large part of the genome, saving them time from the 'sequestering' effect of random DNA. This could be relevant especially for 'pioneer' TFs that -differently from canonical TFs- do not find in nucleosomes an effective barrier against nonspecific sequences. On the other hand, excluded by dense chromatin, TFs can focus their search on more promising regions (enriched by potential target genes). In this perspective, while the

interior of chromatin domains according to Cremer's is poorly permeable, their surface is directly accessible from the IC channels. As a consequence, pioneer TFs may use the CD surfaces as platforms where landing (directly from the IC by 3D diffusion) and starting their local sampling searching for genes.

4.3.2 Positioning of genes at Cremer's chromatin domains

According to Cremer's model, repressed genes occupy the CD dense interior, whereas active genes are positioned at the CD periphery, directly in contact with the IC corridors.

In this regard, p53 inducible genes may represent a special category since they are inactive in absence of stress, and active following DNA damage. As a special category, there are two possibilities: they are permanently located in favorable positions at the CD/IC interface, or they change position from the CDs interior (when inactive) to outside (when active). Let's see the implications of both possibilities.

First hypothesis: inducible genes reside at CD/IC interface. In this situation, they may be immediately targetable from the IC by p53 -and this may facilitate the search process- but in this configuration they could be visited also by the transcriptional machinery. Since at basal condition they are not expressed, some mechanisms should prevent the RNA-polII from initiating transcription.

Second possibility. Inducible genes -when inactive- are inside the CDs, and moved outside after activation. We can discuss on the meaning of 'moved outside' (it could be either that they physically change position or that CDs are dismantled), but the important point here is whether they move out before or after p53 binding. It can be argued that they may move outside immediately after DNA damage, through mechanisms of chromatin reorganization, to favor p53 targeting at the CD/IC interface. Indeed, after damage, chromatin is reported to undergo global changes (Hauer & Gasser, 2017). However, treatments by Nutlin (a chemical agent blocking the p53 negative regulator MDM2), induce relevant p53-mediated gene expression without apparent DNA damage (Nguyen *et al*, 2018; Fischer, 2017), suggesting that the genes are already targetable by p53. Thus, we now consider the possibility that inactive p53 targets actually reside inside CDs. Is it a plausible hypothesis? How can p53 reach them in dense regions? We have mentioned that large inert tracers are actually excluded by chromatin depending on its density. However, this is not a binary property (no diffusion/all diffusion), and is -instead- more similar to a gradient, where also dense regions can be visited, although at a lesser extent. This implies that a fraction of TF molecules may still penetrate. Of course, this may be true also for the

transcriptional machinery, but with a significant difference in likelihood. The RNA-polII alone (even without all the remaining PIC components) is 550 kDa, while tetrameric p53 is less than half its size. We then need to consider aspects of the search mechanism itself to define this likelihood, since p53 for example can bind nucleosome and slide on DNA (at least in vitro). Combining these two features, p53 may enhance its capability to penetrate dense regions. Interestingly, a recent study on Cas9 and TALEN proteins showed that when they were instructed to target heterochromatin DNA, they can effectively reach dense regions, with a strong difference in efficiency (TALEN was 5 times faster), dictated by their search strategies (Jain *et al*, 2021). This reveals that (i) nuclear factors can reach dense sites, and -importantly- (ii) different proteins can adopt distinct search mechanism, with important consequences in the final output. Further evidence that TFs can reach and activate genes positioned in dense chromatin regions, come from an interesting work using the bacterial LacI in mammalian cells (Tumbar *et al*, 1999). The authors inserted an artificial gene containing both the LacI binding site and a DNA sequence regulating heterochromatin formation. Keeping LacI inactive, the artificial gene appeared positioned within a large and dense heterochromatin domain. Following LacI activation, the gene was activated, occupying positions in relaxed and accessible chromatin.

Overall, these data suggest that -in principle- repressed genes occupying the dense CDs inside could be still found and activated. We argue -though- that genes placed in remote positions in the inner CDs may be reached less efficiently, with a likelihood that correlated to the distance from the CD surface.

4.3.3 A short summary

The nuclear structure depicted by Cremer's has some intrinsic limitations, mainly given by the fact that the CD/IC structure, in reality, could be more complex. For instance, CDs instead of representing chromatin compartments defined by precise boundaries, could be formed by a continuum of fibers becoming progressively less dense until appearing as IC channels. However, Cremer's model, retains a strong concept: not all sites are equally available, with a likelihood given by chromatin density. This is important because it can reduce the search space of a TF.

The position of p53 target genes in Cremer's nuclear model is unknown, and can technically challenging to be measured. However, for the considerations illustrated before, we consider plausible two possibilities, with genes either readily available at the CDs/ICs interface, or positioned inside the CDs. We speculate that an

intermediate position could be a good balance for the necessity of an efficient targeting and the preventing an anticipated recruitment of the transcriptional machinery.

4.4 Nuclear bodies and transcriptional condensates

Beyond the chromatinized genome, living nuclei are further characterized by several membrane-less sub-compartments (usually formed by proteins or proteins and RNAs) generally called nuclear bodies (NBs (Misteli, 2007; Spector & Lamond, 2011)). Examples of NBs are provided by the nucleolus (dedicated to the production of ribosomal RNA), nuclear speckles (organelles involved in the synthesis and processing of mRNA) and PML bodies. An important observation is that some NBs recruit TFs to regulate fundamental aspects of their activity (Galganski *et al*, 2017). With this in mind, here we analyze the search process from the perspective that TFs need to find specific NBs to regulate their functions, focusing on PML bodies that have been shown to associate with p53.

Finally, a conclusive and important property of nuclear organization is the emerging evidence that TFs can form local hubs in the nucleus, generally named molecular condensates. TF condensates often display liquid properties (e.g. a high mobility of molecules in their inside), and are thought to have a relevant -but still debated- role in transcription. The current view states that the higher concentration of TFs inside condensates would increase the binding at specific genes. We will examine this possibility highlighting some weaknesses, and illustrating recent evidence that suggest alternative perspectives.

4.4.1 PML bodies

NBs confer the nucleus with spatially defined regions where specific molecular reactions take place, contributing to the regulation of TFs activity. In this context, the so-called PML bodies play a significant role. PML bodies are nuclear sub-compartments enriched with the promyelocytic leukemia (PML) protein (Bernardi & Pandolfi, 2007). Among many other functions, PML bodies are suggested to provide subnuclear factories that interact with TFs, mediating the accumulation of PTMs (Bernardi & Pandolfi, 2007; Lallemand-Breitenbach & Thé, 2010). In this regard, an important issue still poorly understood is how TFs meet their modifiers and co-regulator (e.g. acetyltransferases) in the cell nucleus, allowing for a spatiotemporal coordination of their functions. In principle, they may encounter randomly through diffusion in the nuclear space, although this appears not very efficient and could be

counterbalanced by encountering modifiers that remove PTMs (e.g. deacetylases). On the other hand, the evidence that TFs can be regulated at the PML bodies suggest that these sub-compartments may facilitate TFs and coactivators to find each other. Indeed, some studies reported that p53 can co-localizes with the PML bodies (although, to our knowledge, a strong co-localization was observed only by PML overexpression; (Pearson *et al*, 2000; Fogal, 2000)) and that p53 can interact with several PML isoforms (Fogal, 2000; Ivanschitz *et al*, 2015). Interestingly, important coactivators (e.g. p53 modifiers as p300) co-localize with PML bodies, and it has been shown that PML regulates p53 phosphorylation and acetylation (Lallemand-Breitenbach & Thé, 2010; Pearson *et al*, 2000; Rokudai *et al*, 2013).

Together, these data suggest a model where inactive p53 reaches a nuclear PML body and encounters coactivators that mediate specific PTMs. To be effective, this process requires that TFs and coactivators have higher probability to encounter within the PML nuclear bodies than outside. Some questions -however- remain to be addressed. What is the mechanism underlying the recruitment of p53 and coactivators? A possibility is that -after going inside a PML body- p53/coactivators get entrapped by molecular interactions (as for example those occurring between p53 and different PML isoforms) that retain them inside for a time sufficiently long to ensure their encounter. Alternatively, the time spent inside by p53/coactivators could be not that long by itself, but the frequency of molecules going inside at any moment may by very high, due to an extremely efficient search strategy allowing p53 and coactivators to rapidly locate PML bodies. On this possibility, we highlight that PML bodies have a peculiar surface with a shell-like structure (Lallemand-Breitenbach & Thé, 2010), that in principle may be used by molecules to reduce the dimensionality of their diffusion process ($df < 3$), exploring repeatedly the shell until they get inside the PML compartment. The two possibilities (efficient search process or long time spent inside) are not mutually exclusive, but need to be verified.

4.4.2 Transcriptional condensates and the liquid-liquid phase separation controversy

In recent years, it has become increasingly appreciated that some TFs can form local aggregates (or molecular condensates) in the nucleus. Condensates are thought to be generated by a process of liquid-liquid phase separation, a phenomenon observed for several proteins in vitro under circumstances that depends on their structure and on environmental parameters (e.g. pH, temperature; (Alberti *et al*, 2019)). A protein uniformly distributed in a solution, forms a unique chemical phase

with invariable concentration. Increasing the protein amount up to a critical limit, the uniform distribution can separate into two different phases, dilute and dense, with the dense one formed by liquid-like droplets (or condensates; Figure 4.7). The transition is said to be 'liquid-liquid' because the two phases display liquid properties, as the capability of dense droplets to fuse together when contacting each other, forming bigger condensates.

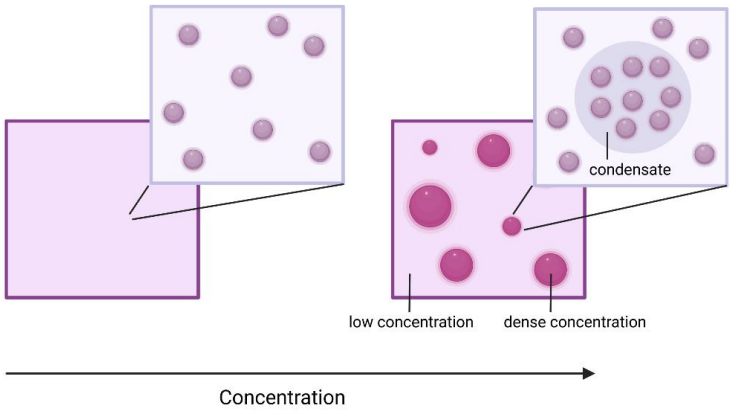


Figure 4.7 Formation of molecular condensates through phase-separation. *At low concentration (left), a protein solution appears uniform, forming a single chemical phase with fixed properties. Above a critical concentration (right), the solution separates into two different phases, forming dense molecular condensates disperse in the lower-concentrated phase. Image created with BioRender.com.*

Driving force of phase-separation resides on the intrinsically disorder regions (IDRs) of proteins (Boija *et al*, 2018; Alberti *et al*, 2019; Chong *et al*, 2018). As discussed in other sections, IDRs confer proteins with a strong adaptability to different substrates allowing them to create weak interactions with several molecules simultaneously (a property called multivalency).

Now, TFs typically contain IDRs fundamental for their activity but for which a clear mechanistically model of function was -and possibly is still- lacking. Several TFs have been demonstrated to undergo phase-separation in vitro, and this property has been correlated with the observation that they often also form local hubs in the nucleus. These hubs often co-localize cofactors. Some TFs however (and p53 in this is a paradigmatic example) despite strong IDRs, neither form liquid droplets in vitro (in physiological-like conditions) nor form visible condensates in the nucleus.

Several transcriptional condensates have been reported to form at enhancer regions during development, formed by co-activators (Sabari *et al*, 2018), RNA-PolII (Cho *et al*, 2018), or co-clusters between TFs and coactivators (Boija *et al*, 2018). Moreover, a study showed that several IDRs artificially fused to the bacterial LacI

binding domain, can form condensates at synthetic arrays containing repetitive LacI binding sites (artificially inserted in mammal cells), but also in random locations throughout the nucleus (Chong *et al*, 2018).

Some studies proposed that phase-separated nuclear condensates increase transcription (Sabari *et al*, 2018; Cho *et al*, 2018; Wei *et al*, 2020), with a model suggesting binding cooperativity between TFs and cofactors (Hnisz *et al*, 2017). If we restrict the focus on TFs only, a consequence of having many copies of the same TF in condensates positioned on genes, is that as soon as a TF molecule dissociate from a DNA RE (enhancer/promoter), can be replaced by another one, increasing the binding rate, and in turn transcription.

Most concerns on transcriptional condensates (Narlikar *et al*, 2021), can be summarized as (i) whether the nuclear condensates are actually formed by phase-separation, (ii) whether the condensates primarily form on correct genomic locations and (iii) what is their actual role in transcription.

According to some researches (e.g. the Robert Tjian and Xavier Darzacq group), generic TF condensates are not all necessarily formed by liquid-liquid phase separation but also other mechanisms can intervene (McSwiggen *et al*, 2019b). A clear example came from the same group, with a work on the Herpes Simplex Virus, capable to induce visible aggregates of RNA-PolII (McSwiggen *et al*, 2019a). Despite some features of these condensates are in line with phase separation (e.g. condensates could fuse forming bigger droplets), other fundamental aspects were not observed (as no difference in the diffusion coefficient inside and outside condensates), highlighting the possibility that other mechanisms may induce condensates formation. This paradigmatic case was used as a warning for more systematic and quantitative analysis of molecular condensates in cells.

The second layer of this controversy is on the actual role of condensates in transcription, highlighted by the conflicting evidence from the recent literature indicating that at least some condensates can be detrimental for gene expression. According to the 'canonical' model, the formation of condensates would favor the binding frequency of TFs to their target genes, increasing transcription. However, the binding frequency does not exclusively depend on TF local concentration, but -as we have largely discussed- also on the TF search efficiency. As for condensates, an efficient search leads TFs to frequently find and associate to their genes. In other words, condensates and efficient search mechanisms are separated phenomena (although not mutually exclusive (Kent *et al*, 2020)) that can enhance the association rate of TFs to their genes and, in turn, gene expression.

We stress this notion because, despite being distinct phenomena, they display common molecular bases. As mentioned, the formation of condensates strongly depends on protein IDRs, which allow TFs to engage weak interactions. Similarly, the TF search is strictly related to IDR-mediated interactions, inducing compact exploration and guiding TFs towards an accelerated gene targeting (Hansen *et al*, 2020; Kent *et al*, 2020). As described, IDRs were demonstrated to guide compact exploration both in vitro (see for example the p53 CTD (Tafvizi *et al*, 2011)) and in vivo (Garcia *et al*, 2021a; Hansen *et al*, 2020).

Since the TF IDRs can influence two different outputs (condensate formation and accelerated search), this situation can create potential issues in the interpretation of data. We propose some recent examples from the literature, to highlight the problem. In a publication was studied HOXA9, a TF that often undergoes genetic translocation with strong IDRs (e.g. NUP98), inducing leukemia (Ahn *et al*, 2021). Using cancer cell lines models, it was demonstrated that the fusion protein can form phase-separated nuclear condensates. A subsequent analysis by ChIP-seq revealed higher genomic peaks of the TF fusion protein compared to its WT counterpart (incapable of making condensates). This is in line with the canonical model of nuclear condensates increasing the TF binding frequency, possibly because condensates arise on correct genomic positions. However, the results do not disprove the alternative 'accelerated search' model either, since it is still possible that the binding frequency increases independently of condensates and through -instead- an efficient search mechanism *outside of condensates*. To distinguish among these two possible models, it would be important (i) to show whether actively transcribed genes are positioned within condensates; (ii) to analyze the relationship between TF abundancy, TF condensation and transcription at the single cell level. This is a relevant point, especially if considering two recent works from different groups, suggesting that transcription is independent of the actual formation of condensates (Trojanowski *et al*, 2021) and that condensates can even be detrimental, repressing gene expression (Chong *et al*, 2021). In both publications, imaging methods were used to correlate condensate formation and gene expression at the single cell level. Data on different TF IDRs fused to the same DNA binding domain, showed that the IDR multivalency increases transcription but does not correlate with the formation of condensates (Trojanowski *et al*, 2021). Moreover, the oncogenic fusion TF EWS/FLI1, when forming condensates represses the expression of its target genes (Chong *et al*, 2021).

5 Live-cell imaging methods to investigate the TF search mechanism in the nucleus

After having discussed how the cell nucleus -organized at multiple levels- can influence the TF search process and selectivity of genes, we now conclude the introduction with a last chapter, focusing on the techniques that can be applied to study the search mechanism in vivo, in the context of nuclear organization.

5.1 TF search investigated by Single molecule tracking (SMT)

As mentioned in previous paragraphs, SMT is fluorescence microscopy approach largely applied to investigate diffusion and search mechanism of TFs. A great advantage of SMT is the possibility to observe individual molecules in real-time, to extract information on the search process directly from their kinetic behavior. SMT is based on labelling the protein of interest with fluorescent ligands. The protein of interest is usually fused to engineered peptides (e.g. HaloTag) that can be labelled with cell permeable organic dyes (England *et al*, 2015). Typically, the concentration of TF-tagged molecules (either expressed endogenously by genetic knock-in or transient transfection) is high, and labelling the entire population generates a diffuse nuclear staining that does not allow to discriminate single molecules. To overcome this limitation, a low amount of fluorescent dye is used, labeling only a small fraction of molecules. As a result, individual molecules appear as diffraction-limited spots that are imaged by live-imaging movies (Figure 5.1 A, next page).

To study the TF search mechanism, it is necessary to focus on diffusing molecules. More precisely, since during a SMT movie a molecule can change kinetic state (e.g. a diffusing molecule finding a gene, would switch from a diffusing state to an immobile bound state), the segments of each single molecule track are analyzed to isolate the diffusing ones from those bound (Figure 5.1B, next page). This is usually performed using unbiased approaches, based on Hidden Markow models (Persson *et al*, 2013; Vega *et al*, 2018) or machine learning (Granik *et al*, 2019). Finally, the isolated diffusing tracks are analyzed to identify features associated to the type of diffusion, thus directly investigating the search process. As seen in chapter **3.1 Principles of the TFs search mechanism**, the parameters d_w and d_f govern the search efficiency determining whether a TF explores the nucleus via non-compact or compact diffusion. d_w (related to the TF diffusion coefficient) can be estimated by the mean squared displacement (MSD) of diffusing molecules, a coefficient that can be estimated from SMT data.

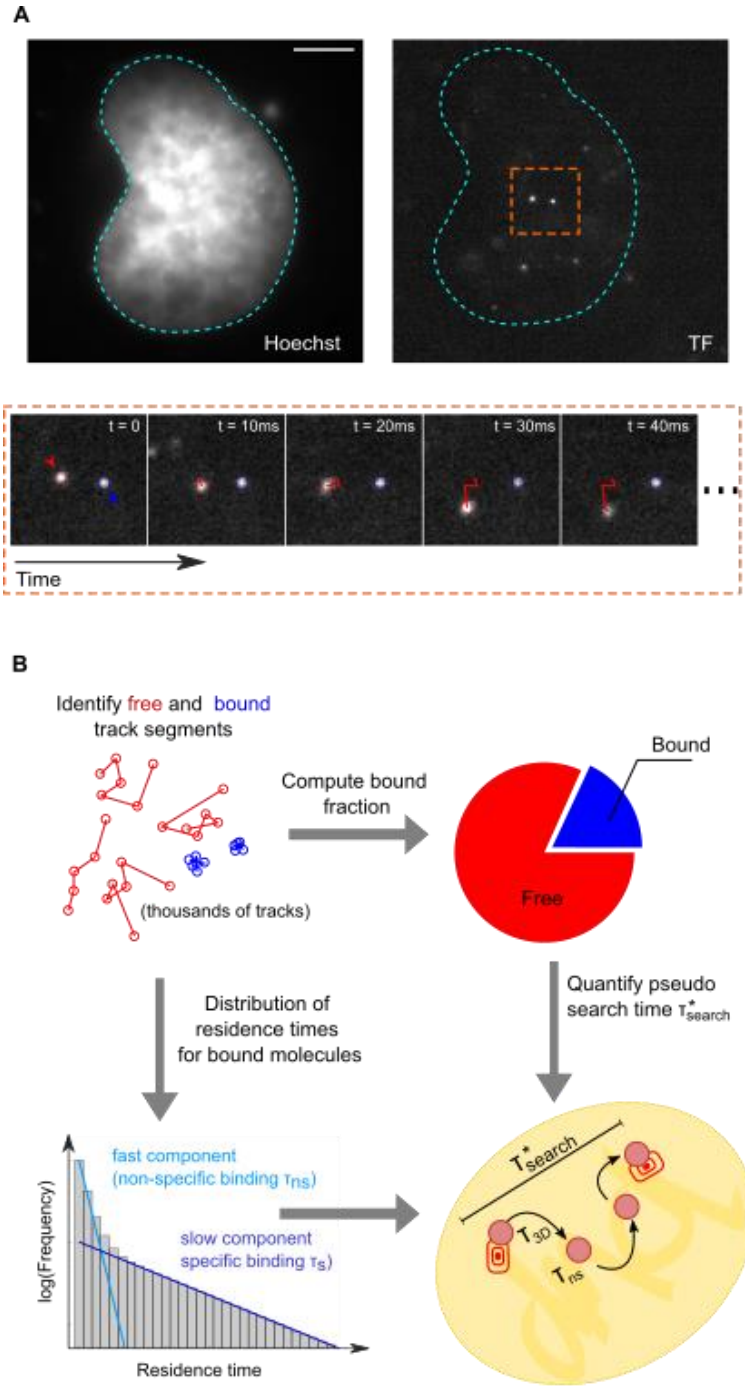


Figure 5.1 Single Molecule Tracking (SMT) to quantify the search mechanism. *(A) SMT is based on labelling the protein of interest with low concentrations of fluorescent dyes, so that the diffraction limited spots corresponding to individual TF molecules can be tracked over time. (B) The single-molecule tracks are segmented into 'bound' and 'free' segments. Free segments are used to characterize the TF search mechanism, by either combining kinetic parameters into a pseudo-search time (i.e. the time it takes to a single molecule to reach one of its many specific sites) or by analyzing the anisotropy of diffusion (see below). Figure from 'The needle and the haystack: SMT to probe the transcription factor search in eukaryotes', open access article (Mazzocca* et al*, 2021b).*

However, d_f (the diffusible space) cannot be quantified with SMT and it would require difficult measurements from 3D high resolution images of chromatin (Récamier *et al*, 2014).

A simple way of discriminating between compact and non-compact exploration consists in analyzing the diffusional anisotropy parameter: in compact exploration a TF visits repeatedly a region, appearing confined, whereas in the non-compact case TFs freely move with no constraints. Confinement (or compact exploration) therefore implies the TF walking back on its steps, a property that can be quantified by measuring the angles between three consecutive positions in single molecule trajectories. For confined molecules, the backward direction (180°) is more frequent than any other (and the diffusion is said anisotropic), while for freely diffusing molecules all angles are equally probable (i.e. a molecule can take any direction) and the diffusion is said isotropic. Using this approach, it has been observed that different nuclear proteins search for their targets adopting different strategies.

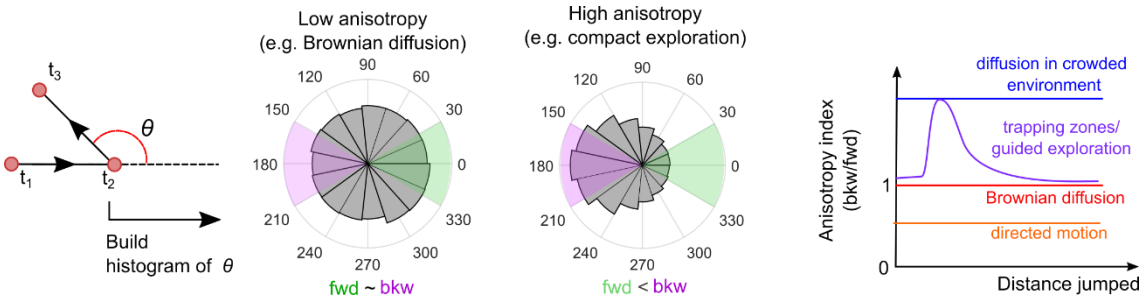


Figure 5.2 Analysis of diffusional anisotropy. *Anisotropy allows to discriminate among different TF search strategies (i.e. compact vs non-compact). Figure from 'The needle and the haystack: SMT to probe the transcription factor search in eukaryotes', open access article (Mazzocca* et al*, 2021b)*.

A useful metric, providing a more precise degree of the confinement, is the 'fold-anisotropy' $f_{180/0}$, that quantifies how many times a step backward (180°±30°) is more frequent than a step forward (0°±30°; Figure 5.2). Confined (compact) molecules display $f_{180/0}$ >1, while free (non-compact) molecules show $f_{180/0}$ =1. In addition, the $f_{180/0}$ metrics can be applied to investigate whether TFs explore the nucleus always by compact, by non-compact or by a combination of the two processes. For example, a TF alternating between non-compact and compact exploration, display backward anisotropy ($f_{180/0}$ >1) not at all scale lengths but only at the scale corresponding to the size of the trapping region (Hansen *et al*, 2017a). This behavior is expected for facilitated diffusion or any other search process with similar alternate diffusion. Interestingly, such search strategy has been recently

observed for the transcriptional regulator CTCF and has been named 'guided diffusion'.

On the contrary, TFs undergoing uniquely compact or non-compact explorations display backward anisotropy ($f_{180/0}$ >1) or isotropic diffusion ($f_{180/0}$ = 1) at all spatial scales.

5.2 Microscopy approaches to investigate the role of chromatin organization on TF search.

The possibility to study TF mobility while knowing the spatial organization of the nucleus, would allow to investigate how TFs perceive the surrounding environment and whether they modulate the search, accordingly. An example is provided by Cremer's model: to study whether TFs reach their targets by diffusing throughout the interchromatin channels, maps of chromatin (e.g. labelled by DNA binders like Hoechst) at high resolution would be required. However, SMT tracking is a technique based on widefield microscopy, which by definition does not allow to image the chromatin space at high resolution, mostly because of the limited optical sectioning associated to this illumination scheme.

Super-resolution microscopy based on imaging single molecules (e.g., Photoactivation Localization Microscopy (PALM) or direct Stochastic Optical Reconstruction Microscopy (dSTORM)) have been used to provide high-resolution maps of chromatin organization, down to ~20nm in lateral resolution (Ricci *et al*, 2015; Boettiger *et al*, 2016; Nozaki *et al*, 2017). However, these approaches are ill-suited to imaging live cells, due to the specific requirements on the chromatin labeling strategy and on the long acquisition time (in the order of minutes) required to collect all the localizations necessary to provide a suitable reconstruction of the image. An interesting alternative approach is represented by structured illumination microscopy (SIM), that while providing a lower improvement in resolution (about 100-150 nm laterally), allows for optical sectioning, flexibility in the choice of chromatin labelling strategy, and rapid acquisition times (in the order of seconds (Heintzmann & Huser, 2017)). Recently, SIM has been used to test the Cremer's model for chromatin organization in fixed cells (Miron *et al*).

The work described in this thesis – focusing on characterizing the role of nuclear organization on TF search – relies on a novel microscope developed in our lab (described in the Material and Methods section, **10.2 SMT/mSIM**) that combines SMT of TFs with structured illumination Imaging of the nuclear architecture, using either commonly used DNA labelling dyes (e.g. Hoechst) or GFP-tagged proteins.

AIM OF THE WORK

As an inducible TF activated by DNA damage, p53 needs to rapidly scan the human genome to locate hundreds of genes and regulate several cell functions (e.g. cell cycle arrest). This task seems extraordinary difficult given the enormous amount of nonspecific DNA, potentially segregating p53 (and all TFs indeed) far from their target genes.

Organisms are thought to have evolved molecular strategies (or 'search mechanisms') to rapidly direct TFs to their DNA sites to regulate transcription. In bacteria for example, facilitated diffusion accelerates the TF search process, with genes that are reached faster than expected by random diffusion. Importantly, facilitated diffusion accelerates the search not 'in spite of' but 'because of' random DNA, since TFs take advantage of DNA nonspecific contacts to rapidly scan the genome. This highlights the importance of weak and transient interactions in the search process.

It is still poorly understood how human TFs effectively locate their targets in the large genome, and evidence of facilitated come only from in vitro studies (e.g. on p53). However, recent data suggest that also eukaryotic TFs adopt their wide repertoire of weak interactions in the search process, contacting multiple nuclear substructures. Importantly, the highly compartmentalized nucleus creates local environments that can either exclude TFs, or trap them to facilitate the recruitment at target genes.

While there has been remarkable progress in our understanding of nuclear organization and TF dynamics, these two aspects are often treated separately, limiting a comprehensive analysis of the TF targeting process.

The general aim of this thesis is to characterize the p53 search mechanism in living cells, investigating its molecular bases and the influence exerted by the nuclear compartmentalization. We apply a cutting-edge live imaging approach (SMT/mSIM) developed by our group, that allows to follow the search process of individual molecules throughout the nuclear spatial organization, with a focus on compartments defined by chromatin, and -marginally- by PML nuclear bodies.

We take the Cremer's model of the cell nucleus as archetype to dissect how the crowded nuclear environment can guide p53 towards its targets. More precisely, we focus on (a) how p53 explores the genome through diffusion, (b) whether such exploration is tuned depending on the nuclear compartments (e.g. chromatin, nuclear bodies) and -finally- (c) what local interactions underlie the p53 search process, with implications on gene targeting process and gene expression.

RESULTS

6 Cellular model to investigate p53 search mechanism

In this first part of results, we describe the cell model adopted in this study and the model validation, focusing on three aspects: (i) the capability of p53 of activating transcription following DNA damage, (ii) the generation of both p53-null cells and cells with endogenous expression of p53-HaloTag (required for live imaging SMT), and (iii) the identification of an early time point following DNA damage, where the activity of p53 reaches its peak, that will be used for subsequent analysis throughout the thesis.

6.1 Choice of the cellular model

To investigate the relationship between p53 search and nuclear organization we decided to adopt a U2OS-derived cell line from Gaëlle Legube group, called DIvA (DSB Inducible via AsiSI (Aymard *et al*, 2014)).

We chose this cell line for two reasons. First, DIvA retains p53 WT, an important feature considering that most cancer cell lines express p53 DNA binding mutants, incapable of regulating p53 target genes. Second, this cell line appeared particularly appealing for the readily availability of Hi-C data (including a recent publication (Arnould *et al*, 2021)), that instructs about genome organization and genome-wide chromatin compaction, features that we would like to relate to the p53 search mechanism (see **4.2.2 Chromatin folding into Topologically Associated Domains (TADs)**).

The DIvA cell line is characterized by the stable expression of an inducible nuclease, AsiSI, that can cut DNA determining double strands breaks (DSBs). This is a type of DNA damage induced also by γ-irradiation, a method widely employed to activate p53. AsiSI is fused to the estrogen receptor (ER) ligand binding domain and, at basal conditions, it is retained in the cytoplasm. AsiSI-RE translocation in the nucleus can be induced by adding the hormone-analogous compound 4-Hydroxytamoxifen (4OHT) to the culture medium. In the nucleus, AsiSI recognizes about 150 restriction sites throughout the genome, inducing DSBs (Aymard *et al*, 2014).

We decided to induce p53 activity in this model exploiting the 4OHT treatment along with the canonical γ-irradiation (IR). Originally, we also planned to expand the already available Hi-C datasets with additional data generated by us, in conditions of DNA damage by irradiation (for which data are not currently available). Given a lack

of expertise in our lab, we scheduled my stay in the Gaëlle Legube lab for 6 months, to learn the technique and generate new data. Unfortunately, under the Sars-Cov2 pandemic we had to reconsider our program. However, preliminary Hi-C data kindly shared with us by Legube's group, showed no major difference in terms of TAD organization before and after 4OHT treatment, convincing us to focus on the available Hi-C dataset in basal condition.

We also anticipate that the characterization of the DIvA cell line in terms of p53 activity after DNA damage (next section), showed different yields depending on the treatment, with 4OHT (to induce AsiSI) activating p53 moderately relative to γ-IR. This led us to adopt γ-IR as preferential method for p53 activation in our cellular model.

6.2 Characterization of p53 activation in DIvA cells

After DNA damage, several regulatory proteins (e.g. ATM) modify p53 with multiple PTMs (Liu *et al*, 2019). Among these, phosphorylation of the p53 TAD inhibits the interaction between p53 and the ubiquitin ligase MDM2, reducing in turn p53 ubiquitination and degradation, and making the TAD available for recruiting cofactors (e.g. p300). As a result, p53 accumulates and induces the expression of its target genes (e.g. *CDKN1A*).

The DIvA cell line has been generated from the U2OS osteosarcoma. Despite U2OS cells have been largely employed as a model for studying p53 WT (Nguyen *et al*, 2018), to our knowledge p53 activity in DIvA has not been characterized.

To assess p53 functionality in DIvA cells, we have generated a population of DIvA p53-null cells, to be used as a control cell line in our experiments. We obtained this population by transfecting DIvA cells with knock-out (KO) CRISPR/Cas9 vectors (using a commercial kit), to induce the deletion of p53 gene. After transfection and negative selection (see details in Material and Methods, **9.2 Generation of p53-HaloTag and p53 KO cell lines**), cells were tested by western blot before and after p53 activation by DNA damage (Figure 6.1). The resulting population contains p53-null cells, with a fraction of residual p53 WT ones, as showed by the western blot bands. Hereafter, we will refer to this mixed population as the 'p53 KO pool'. As we will describe in next paragraphs, the KO pool has been employed to isolate individual p53 KO clones. At this stage however, the pool was used as a control for preliminary tests on our cellular model.

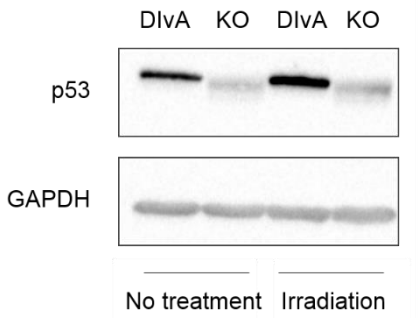


Figure 6.1 Western blot to measure p53 level on cells genetically modified to induce a p53 knock out. *The panel shows p53 level following activation with DNA damage (3 hours after irradiation, 10 Gy) in DIvA cells, and in a p53 KO pool with a residual p53 expression.*

To verify the activation and the transcriptional potential of p53 following DNA damage, we exposed cells to γ-irradiation (10Gy). γ-rays are ionizing radiations that generate DNA double strands breaks (DSBs) activating the p53 transcriptional activity (Hafner *et al*, 2020; Purvis *et al*, 2012). In a series of time-course experiments, we measured the protein levels of both p53 and a well-established p53 target (p21). To investigate whether p21 is specifically regulated by p53, we compared the DIvA (p53 WT) to the p53 KO pool. I collected cells at several time points after irradiation, and analyzed by western blot (Figure 6.2).

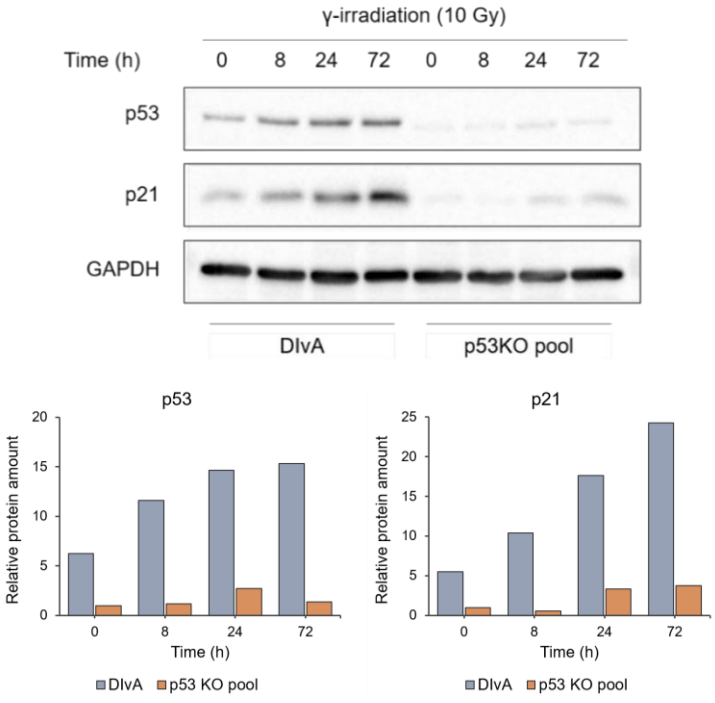


Figure 6.2 Time-course of p53 and p21 protein expression upon the induction of DNA damage. *DIvA cells and our p53 KO pool are compared to investigate p53 activation and the capability to induce p21.*

After DNA damage, p53 gradually increases, and so does p21. The accumulation of p21 appears largely dependent on p53, confirming that DNA damage effectively induces p53 activity in the DIvA model.

As mentioned above, DNA damage can be also generated in DIvA cells by inducing the translocation in the nucleus of the restriction enzyme AsiSI, using 4OHT.

We first tested that stimulation with 4OHT could promote the translocation of AsiSI in the cell nucleus and the induction of DNA damage. To this end, we run immunofluorescence experiments (Figure 6.3), evaluating AsiSI nuclear accumulation and the formation of DNA damage foci, using the DNA DSB marker γH2AX.

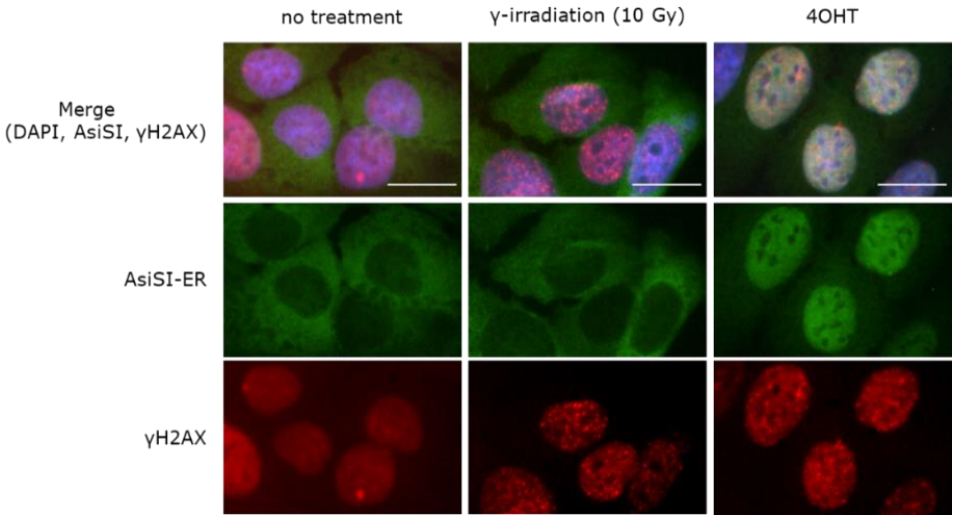


Figure 6.3 Immunofluorescence on the capability of 4OHT to induce AsiSI-ER translocation and DNA damage. *DIvA cells were treated to induce DNA damage by either γ-irradiation (10 Gy) or by 4OHT (for 4 hours). The AsiSI endonuclease (green) migrates in the nucleus after 4OHT treatment inducing DNA damage (labelled by γH2AX, in red).*

In non-treated and irradiated cells, AsiSI-ER correctly localizes in the cytoplasm. Differently, cells exposed to 4OHT display AsiSI-ER accumulation into the nucleus and induction of DSBs, as indicated by the increased level of nuclear γH2AX.

To examine to what extent AsiSI-mediated DSBs can induce p53 activity, we next treated our cells with 4OHT, comparing the expression levels of the p53 target p21 to those observed in the parental cell line U2OS, that is reported to have no detectable functional ER (Quaedackers *et al*, 2001) and should not respond to 4OHT stimulation.

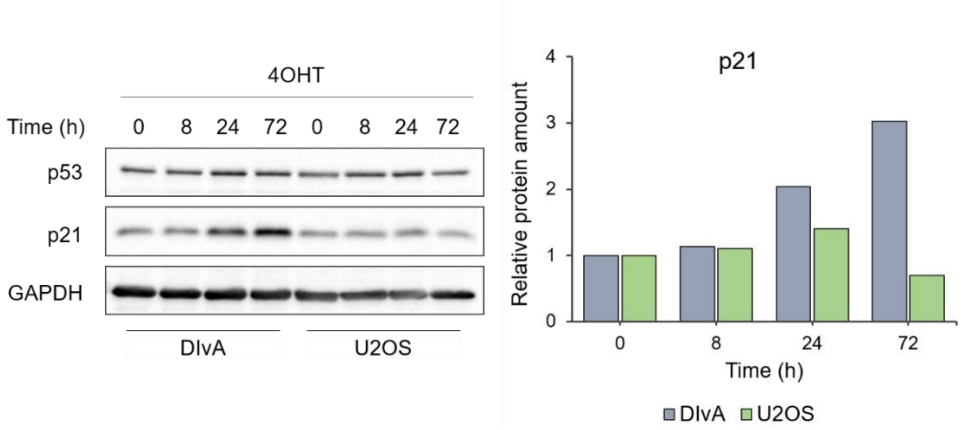


Figure 6.4 Western blot to evaluate the induction of p21 following 4OHT-induced activation of the AsiSI endonuclease. *We compared DIvA cells (expressing the AsiSI endonuclease) and the parental U2OS (not expressing AsiSI).*

After 4OHT treatment, p21 increases in the DIvA cell line, but not in U2OS, suggesting that p53 activity is specifically induced by AsiSI-mediated DNA damage. However, the extent of p21 accumulation over time appears slower and lower than previously observed by IR (comparison of Figure 6.2 with Figure 6.4). This may reflect a limited activation of p53, possibly because AsiSI confers a reduced level of DNA damage compared to IR.

Finally, we performed a time-course RT-qPCR experiment to evaluate the expression of two high-fidelity p53 target genes (*CDKN1A* and *BBC3*) at the mRNA levels (Figure 6.5), in response to either IR or 4OHT-mediated AsiSI activation. Both DNA damage agents induce the expression of the two genes. Again, however, at later time-points, their mRNA levels appear more pronounced following γ-irradiation than 4OHT treatment.

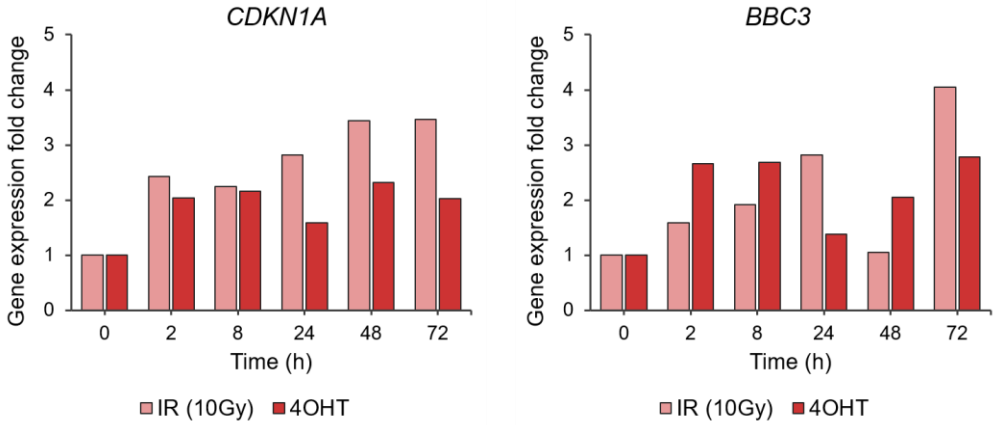


Figure 6.5 Time-course qPCR assessing gene expression of p53 targets. *We measured gene expression of two p53 targets (CDKN1A and BBC3) in DIvA cells, after induction of DNA damage by either γ-irradiation or 4OHT-induced activation of the AsiSI endonuclease.*

These preliminary results indicate that the DIvA is an appropriate model to study the molecular activity of p53, and we therefore decided to generate a DIvA knock-in cell line with fluorescently tagged p53. Also, according to these data, irradiation (10Gy) appears a stronger activator of p53 than AsiSI, and for this reason we selected IR as main method to induce p53 in the project.

6.3 Generation of p53 knock-in and knock-out clones

Following the preliminary validation of our cell model described above, we focused on developing the cellular material necessary for studying p53 search mechanism and selectivity in vivo. The SMT approach requires the protein of interest to be fused with HaloTag, a peptide developed from bacteria and engineered to be conjugated with small fluorescent ligands (England *et al*, 2015). The HaloTag ligands (e.g. JF549, TMR (Grimm *et al*, 2017)) are readily cell permeable, allowing for a rapid labelling of the chimeric TF-HaloTag protein in vivo, right before a SMT experiment. To express p53-HaloTag (p53-HT) at physiological level, we performed on our cells a gene editing knock-in (KI) by CRISPR-Cas9, to replace the endogenous p53 with the p53-HT version. After transfection of four CRISPR-Cas9 vectors (see details in Materials and Methods section, **9.2 Generation of p53-HaloTag and p53 KO cell lines**), cells were kept under antibiotic selection followed by a recovering period. As a result, we obtained mixed cell populations ('pools KI') including cells with a successful p53-HT KI and cells that still retain at least one copy of the endogenous p53 in their genome (Figure 6.6A, next page). To obtain a uniform population expressing only p53-HaloTag, we screened for the most promising knock-in pools and proceeded with isolation and expansion of ~30 single-cell clones (Figure 6.6B, next page). Both the pools and the potential p53-HT clones were screened by treating cells with Nutlin, a potent drug inhibitor of MDM2, which strongly enhances p53 protein level (Purvis *et al*, 2012).

Figure 6.6B shows selected clones that express p53-HT but not the endogenous p53, indicating the correct substitution with the tagged version. Of note, p53 expression was highly variable among positive clones, potentially reflecting that for some clones the gene editing strategy might have been only partially successful (e.g. only one allele being correctly edited). Importantly, the expression of the downstream protein p21, closely matches p53-HaloTag levels in these clones, suggesting that the knocked-in HaloTag p53 retains its capability of inducing the expression of its targets.

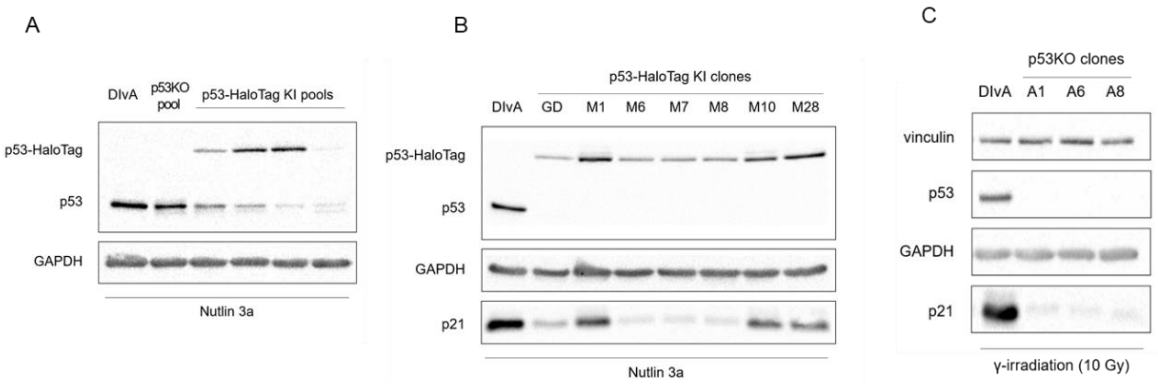


Figure 6.6 p53 and HaloTag-p53 expression in Crispr-Cas9 generated knock-in and knock-out clones in DIvA cells. *(A) p53-HaloTag knock-in pools (B) p53-HaloTag KI clonal populations (C) p53 knock-out clonal populations.*

In parallel, we also isolated single clones from the pool KO mentioned above (Figure 6.6C). In this respect, p53 KO cells will be used for several purposes. First, they allow to identify transcriptional targets specifically activated by p53. This is useful for the validation of our model and, later, for testing genome-wide the transcriptional response orchestrated by p53. Second, p53 KO clones provide a p53-null background that can be exploit for comparing different p53 mutants by transient transfection, without confounding effects caused by hetero-oligomers.

All the isolated KO clones do not retain a p53 band and they do not display a significant level of p21 expression, compared to the parental DIvA.

Next, in a selected group of p53KI and p53 KO clones, we measured the mRNA level of *CDKN1A*. As expected, in p53-null cells the expression of *CDKN1A* was kept at basal level both before and after DNA damage (Figure 6.7, left). Moreover, all p53-HT clones appear transcriptionally functional, although to a variable extent (Figure 6.7, right).

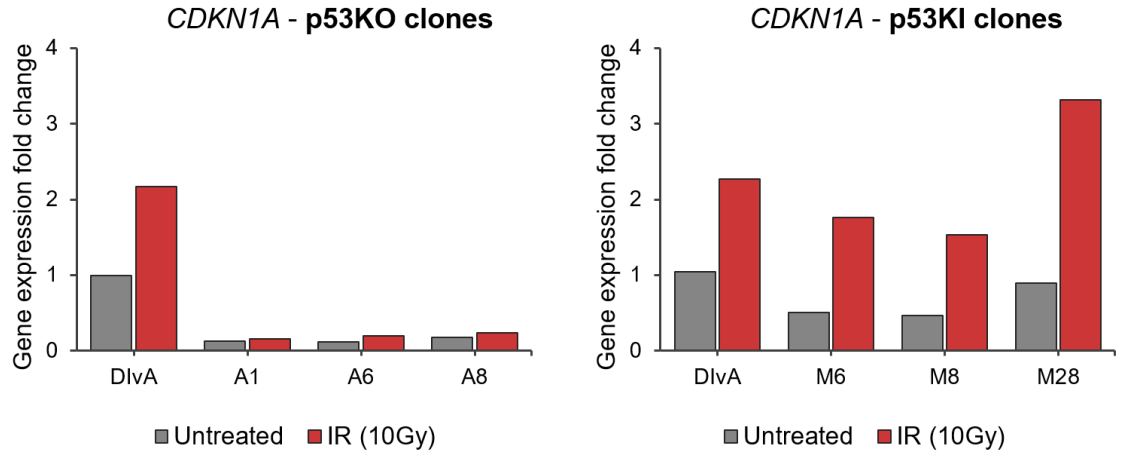


Figure 6.7 mRNA expression of the p53 target CDKN1A. *We compared CDKN1A expression DIvA in cells to KI and KO clones, at 24 hours post IR.*

Finally, we identified two clones -A6 (p53 KO) and M28 (p53-HT)- as the most promising ones to carry out our work. First, we tried to recapitulate the activation of p53 with both γ-irradiation and 4OHT, 24h after treatment (a time point broadly used to study the p53 transcriptional response (Nguyen *et al*, 2018)). The western blot (Figure 6.8) shows that the two treatments induce an accumulation of p53-HT, consistent with the increase of the endogenous p53 in the parental DIvA. Similarly, the level of p21 is enhanced in both cell lines after DNA damage, whereas it appears stably low in the p53 KO cells. A further analysis performed by RT-qPCR confirms that the expression of *CDKN1A* (i.e. the gene encoding for p21) increases after DNA damage in both the DIvA and the p53-HT clone, but not in the p53 KO one.

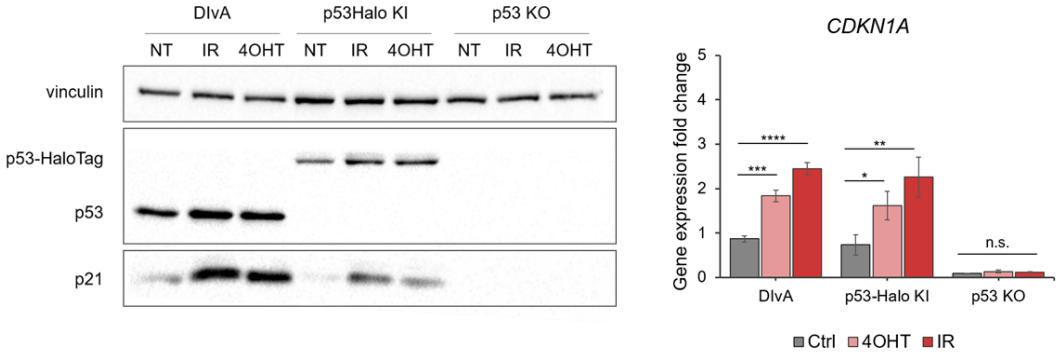


Figure 6.8 Characterization of p53 activation by IR and 4OHT in our p53-HT knock-in and p53 KO clonal cell lines. *Left, protein levels by western-blotting. Right, mRNA levels by RT-qPCR. Statistical T test (Paired-two-tailed; * p< 0.05, ** p< 0.01, *** p< 0.005, **** p< 0.0001, n=3).*

Last, we verified the correct localization of p53-HT in the cell nucleus (Figure 6.9, next page). As expected, p53-HT displays a clear nuclear localization, and its abundancy appears to increase after DNA damage.

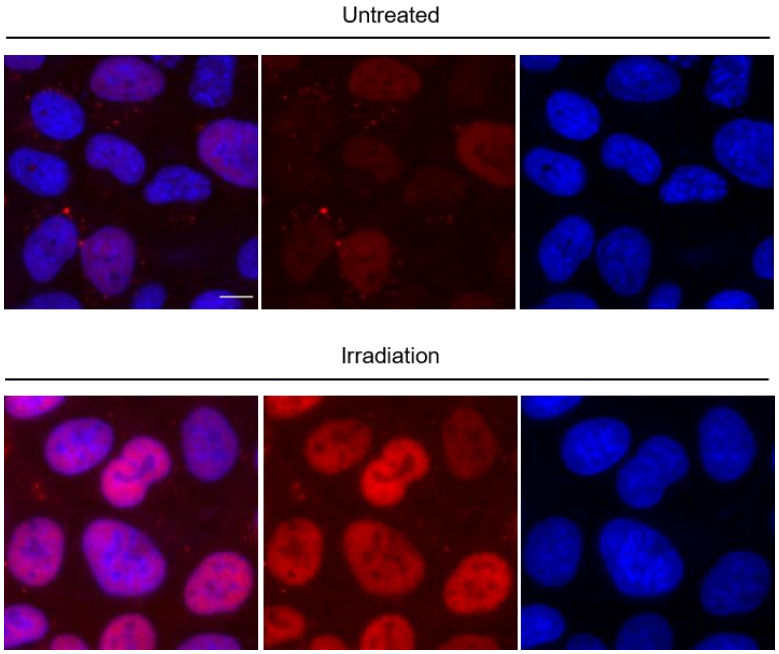


Figure 6.9 Localization of p53-HT in our knock in clonal cell-line. *p53-HT (red) correctly localizes in cell nuclei (blue). After DNA damage by irradiation, p53-HT nuclear level increases as expected by p53 activation.*

Overall, the results highlight that both IR and OHT effectively trigger a transcriptional response (although at different extent) mediated by p53 in the parental DIvA and in the derived p53KI clone. Importantly, the functionality of the p53-HT cell line provides us with a model to study the p53 search mechanism. The p53 KO clone, will be employed to investigate p53 selectivity, defining the genes that p53 specifically induces in the DIvA cell line.

6.4 Identification of the early p53 transcriptional response

The final step in the characterization of our model, was to identify the ideal time-point after DNA for studying the p53 search. We decided to look specifically for an early time point, to avoid confounding effects in the late hours following DNA damage, where several pathways -activated by p53 itself- intersect to each other, leading for example to the activation of genes that are only indirectly regulated by p53. As previously anticipated, we decided to use irradiation as DNA damage agent.

To identify an appropriate early time-point, we performed a time-course analyses on the expression level of *CDKN1A* and *MDM2* following DNA damage (Figure 6.10). The two p53 targets follow a similar pattern over time, peaking at 5h.

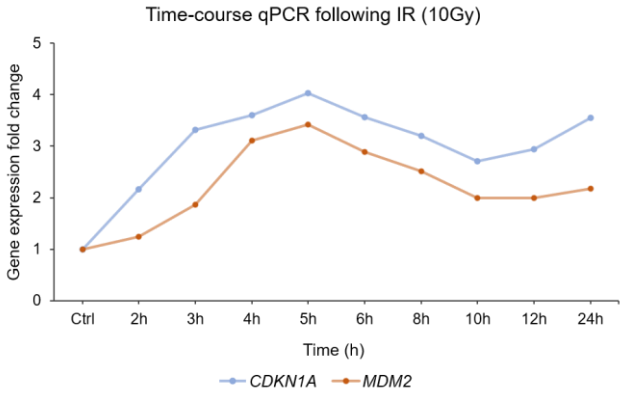


Figure 6.10 Time-course following the expression of p53 target genes in DIvA cells. We measure the expression of CDKN1A and MDM2 by RT-qPCR (n=2).

This suggests that in the time window of 4-5 hours after DNA damage, p53 reaches the maximum of its transcriptional potential. To confirm that also p53-HT is transcriptional active at 5h after DNA damage, we repeated several tests by qPCR. Both DIvA and the p53-HT display a significant transcriptional regulation on a panel of p53 target genes (Figure 6.11).

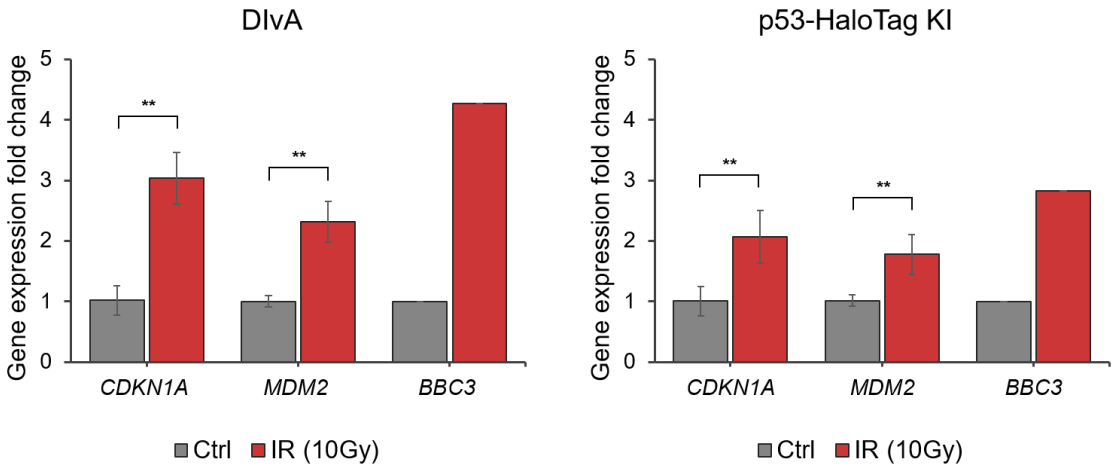


Figure 6.11 Gene expression of p53 targets at 5 hours post-irradiation. *The expression of CDKN1A, MDM2 and BBC3 was measured by qPCR in DIvA cells and in the p53-HT KI clonal population. Statistical T-test (Paired-two-tailed; ** p< 0.01; n=4 for CDKN1A and MDM2, n=1 for BBC3*).

Taken together, the data indicate that at 5h after DNA damage p53 is fully active. We chose this time point to investigate the p53 search mechanism.

7 p53 search mechanism in the chromatinized cell nucleus

In this section we describe the characterization of the p53 search mechanism. We first focus on the analysis of p53 kinetic features, related to different p53 subpopulations such as diffusing molecules (the main players of the search process) and chromatin-bound molecules. We will present data on how these populations are influenced by the organization of chromatin, what p53 protein domains are important for the search process and what is the impact of the search on gene expression.

p53 searches its target sites by a combination of fast and slow diffusion, with slow diffusing molecules that perform compact exploration in sub-micron sized domains and spatially segregate with p53 chromatin-bound molecules.

7.1 p53 search its target sites by a combination of fast and slow diffusion, and slow diffusing molecules spatially segregate with p53 chromatin bound molecules.

To investigate how p53 explores the mammalian cell nucleus searching for target genes, we characterized the diffusion properties of p53-HaloTag (p53-HT) in living cells by SMT.

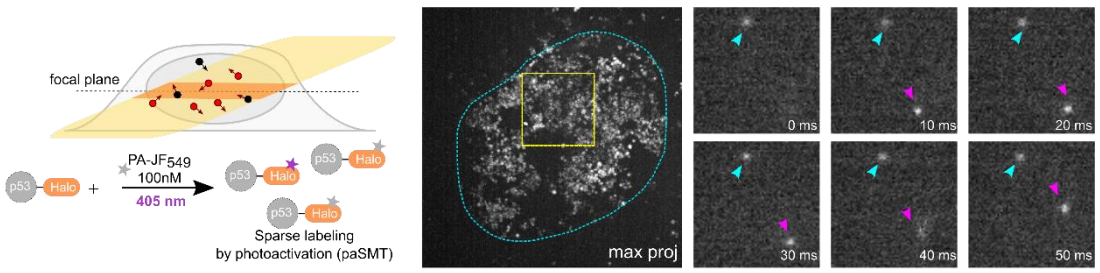


Figure 7.1 Live imaging paSMT of the endogenously tagged p53 in DIvA cells. *p53-HT is labelled with a cell permeable dye (PA-JF549, bottom left) which allows to regulate the number of visible molecules per frame (2-3). Molecules are illuminated through an inclined laser beam to reach a high signal-to-noise ratio (a technique known as highly inclined laminated optical sheet, HILO; see the method section). On the right boxes, a frame sequence taken from a SMT movie (with framerate of 100fps) highlights a quasi-immobile molecule (cyan arrow; possibly a molecule bound to chromatin) and a diffusing molecule (purple arrow, changing position over time), highlighting the heterogeneity of p53 mobility.*

By SMT, individual molecules appear as bright spots moving one frame after another. To reconstruct the paths followed by each molecule, the position occupied by a molecule at frame *t* has to be connected with its position at frame *t*+1, and so on, in a process called 'tracking'. To minimize tracking errors, it is convenient to limit the number of molecules to 2-3 per frame, while collecting enough tracks to

strengthen the statistical power. To this purpose, we labelled p53-HT DIvA cells with a photoactivatable ligand (JF-594) that -by applying a photoactivating laser (405 nm)- allows to switch on only a small fraction of molecules per time (Figure 7.1, left; (Grimm *et al*, 2016)). Tuning the power of the photoactivating laser, we can regulate the number of visible molecules obtaining few molecules per frame. The resulting photoactivated Single Molecule Tracking (paSMT) movies that I collected throughout the project are composed of 10,000 frames and were acquired at a fast frame-rate (100 frame per second, fps, corresponding to 10 ms per fame). This ensures to capture also molecules rapidly diffusing in the nucleus (Figure 7.1and **Movie 1**[1]).

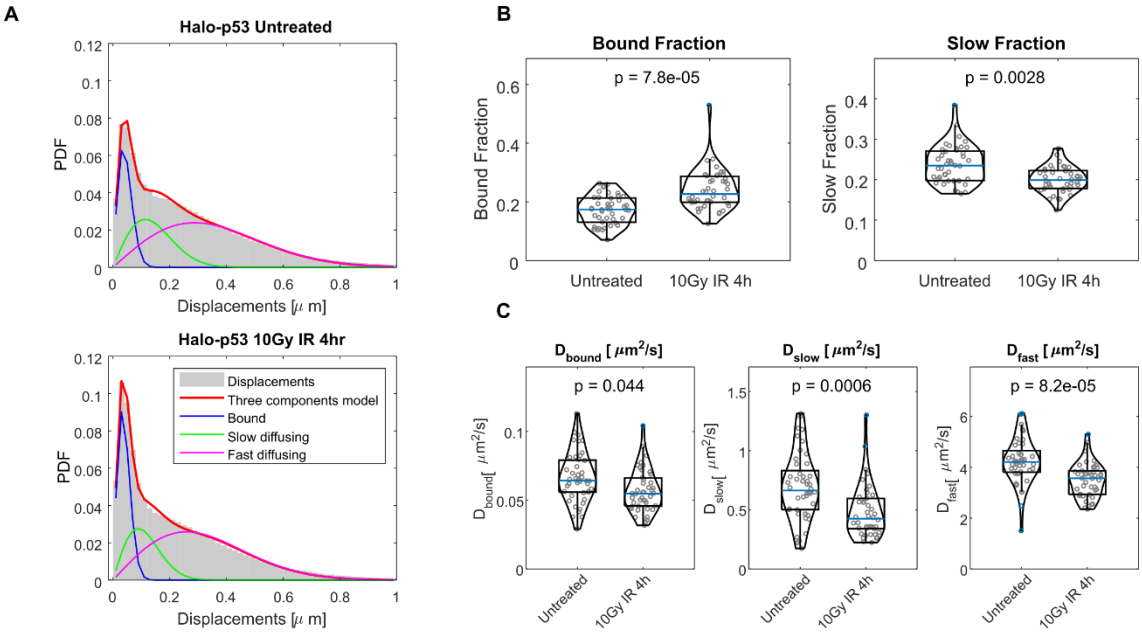


Figure 7.2 Kinetic parameters characterizing p53-HT and extracted from the analysis of the single-molecule displacements. *(A) Single-molecule displacements before and after DNA damage induced by IR. The displacement distributions are described by three kinetic subpopulations, including a bound state (blue line) and two diffusing states, slow and fast (green and magenta lines, respectively). Fitting the distribution of displacements allows quantifying the abundancy and the diffusion coefficients of the subpopulations. After activation through IR, the p53 subpopulations are modulated, with an increase in the bound fraction, indicating that more molecules are engaged in the bound state. Concomitantly, (C) diffusing molecules slow down (n. cells = 45, 43 for untreated and 10Gy IR respectively. Statistical test Kolmogorov-Smirnov, with Bonferroni correction for multiple testing).*

We firstly analyzed the p53 tracks collected from our movies by the SMT distribution of displacements (Figure 7.2), an approach largely employed for the analysis of single molecules (Mazza *et al*, 2012a). A single displacement is the distance that an individual molecule covers between two consecutive frames, from the position at frame *t* to the position at frame *t*+1. Molecules bound to chromatin

[1] Movies to support this thesis are available at: SMT videos

move very little (with small displacements and an apparent diffusion coefficient < 0.1 $\mu m^2/s$), while diffusing molecules move more rapidly (longer displacements; (Mazza *et al*, 2012a)). The frequency of all displacements from a SMT experiment is then used to build the afore-mentioned distribution (or histogram) of displacements, which is next fit to a mixture of multiple diffusing components (Figure 7.2A).

In untreated cells, the experimental distribution of displacements appears to be best fit by assuming three distinct kinetic sub-populations, confirming previous studies on the mobility of p53 (Mazza *et al*, 2012b; Loffreda *et al*, 2017b).

These sub-populations display an apparent median diffusion coefficient of 0.07, 0.7 and 4.2 $\mu m^2/s$ respectively (Figure 7.2B). Consistently with previous observations on a broad range of TFs (Mazzocca *et al*, 2021a), the slowest state involves molecules engaging binding interactions with DNA, and for this reason it is referred to as the 'bound fraction'. Indeed, as shown in Figure 7.3, a p53 DNA binding (p53-DNAb) mutant carrying 7 amino acid substitutions display a reduced bound fraction compared to the p53 WT.

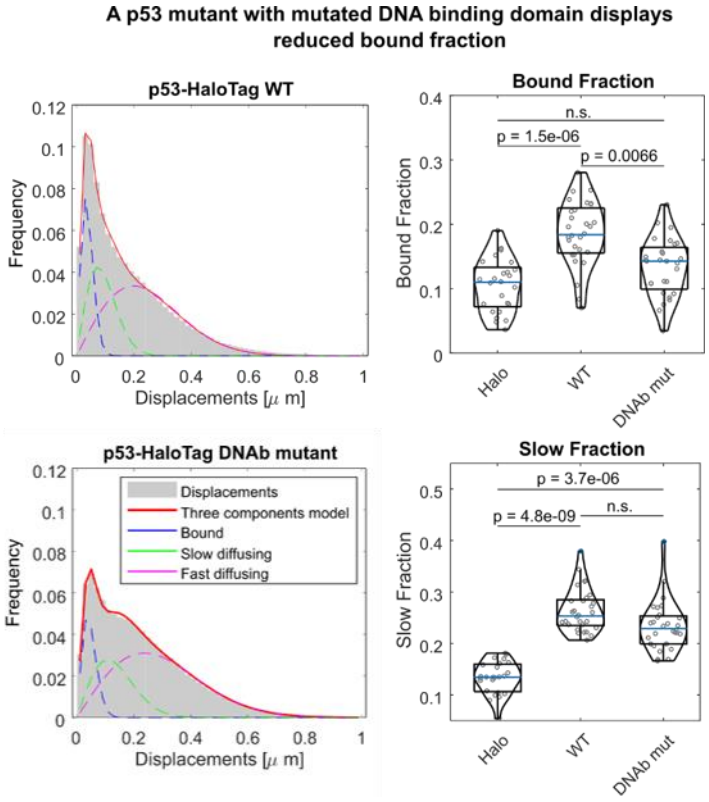


Figure 7.3 Kinetic parameters characterizing p53 WT and a p53-DNAb mutant extracted from the analysis of the single-molecule displacements*. p53KO cells were transfected with plasmids encoding for the two tagged p53 versions, or with unconjugated HaloTag and analyzed as above (n_{Cells} = 25, 29, 28 for Halo, p53 WT or p53-DNAb mutant respectively. Statistical test Kruskal-Wallis, with Bonferroni correction for multiple testing).*

We named the two remaining diffusive populations 'slow diffusing' and 'fast diffusing' based on their mobility. As previously noted (Mazza *et al*, 2012b), these populations may incorporate many more diffusive states (e.g. molecules diffusing in different complexes with coactivators, or transiently interacting with different nuclear substructures, etc.) which -however- converge to two main diffusive behaviors with defined kinetic properties on average. Whether -on average- they also play a different role in p53 search mechanism, is unknown.

To investigate this possibility, we next monitored how p53 mobility is affected upon its activation by DNA damage (IR 10Gy; Figure 7.2A). In turn, the fraction of molecules that binds chromatin increases (Figure 7.2B), as previously reported in a study from our lab on a different cellular model (the breast cancer MCF-7 (Loffreda *et al*, 2017a)). This time however - probably helped by the fast acquisition rates and the 'clean' tracking ensured by the paSMT approach- we also observed a modulation of the diffusing populations (Figure 7.2C), whose diffusion slows down, of ~40% for the slow population and ~20% for the fast one.

To investigate more in detail how p53 searches for its nuclear targets, we focused on the p53 diffusive sub-populations. For studying their kinetic properties (e.g. the anisotropy), they need to be separated from the bound population. This requires tracks segments to be classified into bound, slow or fast (Figure 7.4). For an unbiased classification, we applied the vbSPT method (Persson *et al*, 2013), a Hidden Markow Model-based approach (HMM) that is widely applied to categorize single molecule tracks (Hansen *et al*, 2020; Kim *et al*, 2021). This method provides also useful parameters as the transition rates (i.e. the likelihood for a molecule to change between the different kinetic states).

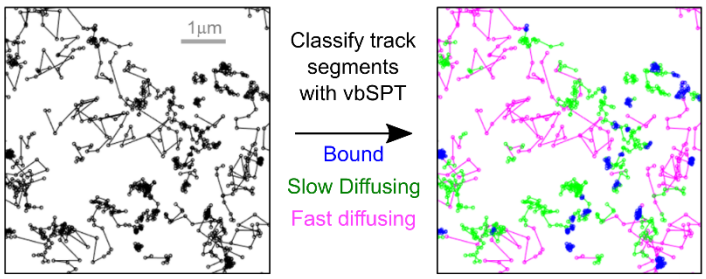


Figure 7.4 Classification of the track segments into bound, slow (diffusing) or fast (diffusing). *All the segments of a single molecule track are analyzed through the vbSPT algorithm (a Hidden Markow Model-based approach), which unbiasedly assigns the segments to a specific kinetic state (bound, slow or fast).*

For each sub-population, the diffusion coefficients resulting from the vbSPT classification are in line with those previously derived by the distribution of displacement, confirming the quality of the classification (Figure 7.5).

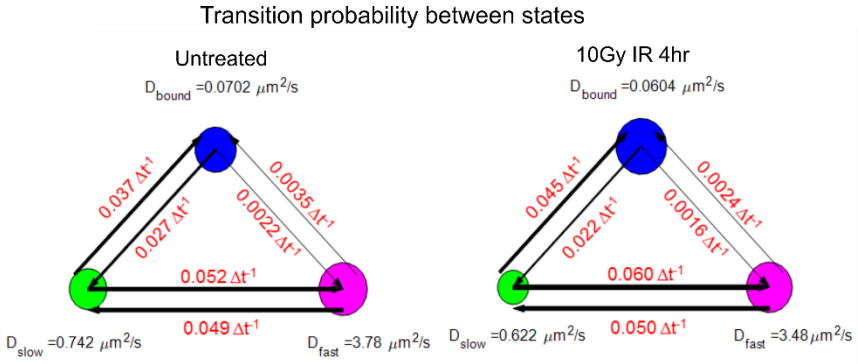


Figure 7.5. Quantification of the frequency by which molecules change kinetic state. *vbSPT allows to derive the probability transition for single molecules between the three subpopulations (bound, slow, fast). Slow diffusing molecules are 10 times more likely than fast molecules to reach the bound state.*

We also verified that bound tracks were not misclassified into diffusing molecules. To this end, we took the diffusing tracks (slow and fast) and we re-analyzed their displacements (Figure 7.6).

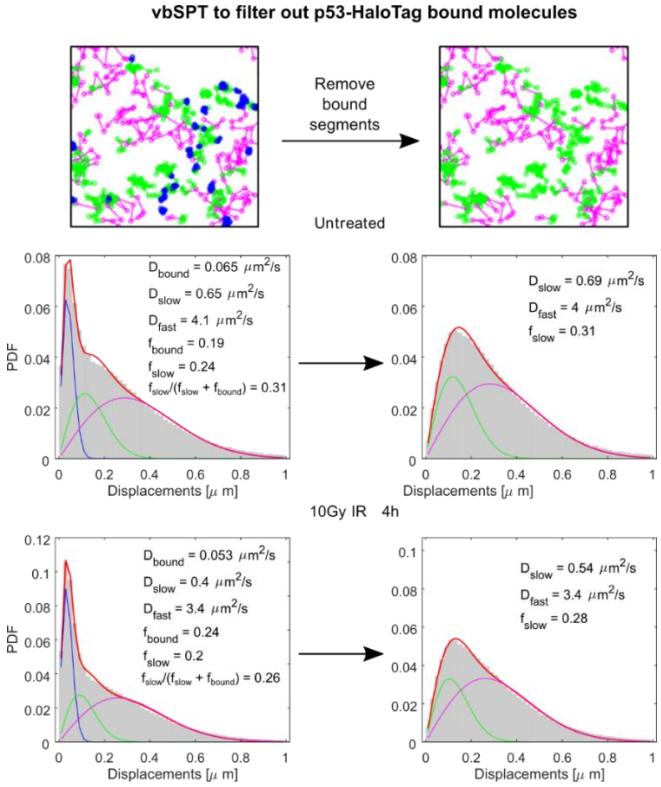


Figure 7.6 Displacement analysis before and after removal of bound track segment. *The removal of bound segments and the reanalysis of the resulting distribution of displacement –show no cross-contamination of bound molecules into track segments classified as slow or*

fast. The resulting diffusion coefficients and fractions are in agreement before and after the removal of bound segments.

As observed, they resemble the distribution that one would expect when p53 bound molecules are absent, with comparable estimates for diffusion coefficients and fractions before and after filtering out bound molecules.

Intrigued by having two diffusing states which may play a different role in the search process, we interrogated the vbSPT model for the transition rates of each population, that is -for example- how frequently a fast molecule or a slow molecule reaches the bound state. The vbSPT estimates that slow molecules become bound ten times more frequently than fast molecules (Figure 7.5). Qualitatively, this can be appreciated also by visual inspection of the track segments (Figure 7.4).

Further evidence that slow and fast populations have not equal probability to get bound is given by measuring how frequently the two classes 'visit' the regions occupied by bound molecules. This can be evaluated by plotting the likelihood of finding slow or fast molecules as function of the distance from bound molecules (Figure 7.7A).

While fast molecules largely occupy random positions relative to bound molecules, the slow ones are frequently found close to bound molecules, with a probability peak at about 100 nm (Figure 7.7B).

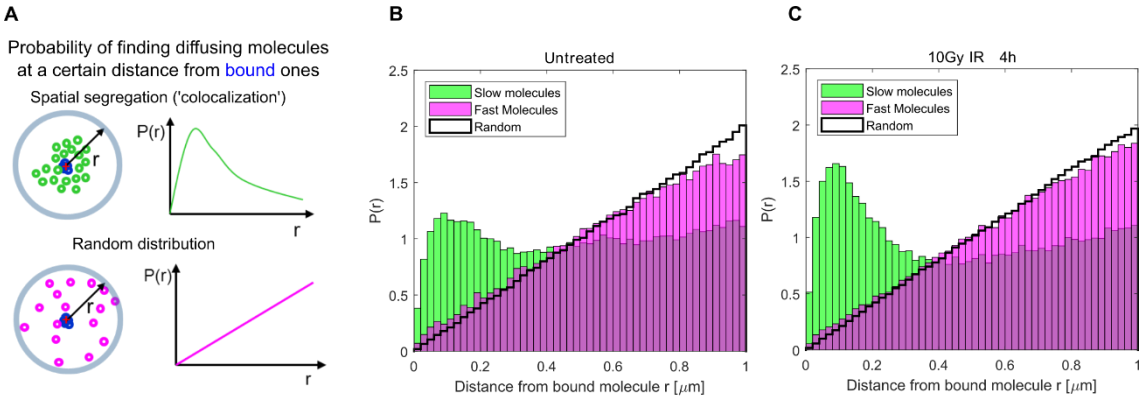


Figure 7.7 Co-localization between diffusing and bound molecules. *(A) To measure how likely p53 is to diffuse and visit regions occupied by bound molecules, we plot the frequency for free molecules (either slow or fast) to be observed at a certain distance from a bound one. For randomly distributed molecules, this probability scales linearly with the distance, while it should display a peak if diffusing and bound molecules co-localize.*
(B) Co-localization analysis performed separately for slow and fast molecules reveals that slow-diffusing molecules co-localize with bound ones, while fast molecules are randomly distributed and segregate from the other subpopulations. (C) The co-localization of slow and bound molecules is further increased after p53 activation by DNA damage.

This highlights that the slow population preferentially explores nuclear regions where p53 engages binding interactions. Importantly, after p53 activation by IR, the ~100 nm-probability peak increases further (Figure 7.7C).

Our data overall indicate that p53 searches for its targets by alternating between two kinetic states, slow and fast. The two diffusing sub-populations explore separated regions of the nuclear space, with the slow molecules that frequently locate at sites where p53 undergoes binding interactions with DNA. We therefore wondered what is the mechanism responsible for such compartmentalization and whether this behavior can facilitate the targeting of p53 to its binding sites.

7.2 Bound, slow diffusing and fast diffusing p53 molecules occupy nuclear regions with distinct chromatin density.

Hinted by the IC-CDs model that describes the cell nucleus to be compartmentalized by chromatin, we asked whether different chromatin compartments may be responsible for the segregation of p53 fast and slow molecules. In its theorization, the IC-CDs model (**4.3.1 Chromatin spatial distribution in the cell nucleus: the Cremer's model**), postulates that chromatin poor regions (or interchromatin compartment IC) provide the nucleus with routes for the rapid diffusion of molecules (e.g. TFs), that reach their targets -positioned at denser chromatin domains (CDs)- through a capillary system. To test the possibility that p53 fast and slow molecules segregation reflects the organization of chromatin, we use a novel microscope (see Materials and Methods, **10.2 SMT/mSIM**), custom designed in our lab. The microscope combines the possibility of performing SMT with a super-resolution approach (multifocal structured illumination microscopy, mSIM), allowing for the simultaneous acquisition of up to three colors (the SMT and two additional super-resolution channels). This live imaging technique therefore allows to capture on the same cell single molecule dynamics and high-resolution maps of the nucleus.

Here, in order to map chromatin regions at high-resolution based on their density, we label DNA with Hoechst 33342, a DNA-binder compatible with live imaging. In Figure 7.8 (next page), we show a comparison between a cell nucleus imaged in widefield and the corresponding mSIM counterpart, highlighting the increased resolution and optical sectioning of the mSIM image. Qualitatively, the chromatin map displays regions of different density, with poor-chromatin channels branching throughout the nucleus and being interspersed with denser chromatin domains, resembling the proposed IC-CDs model.

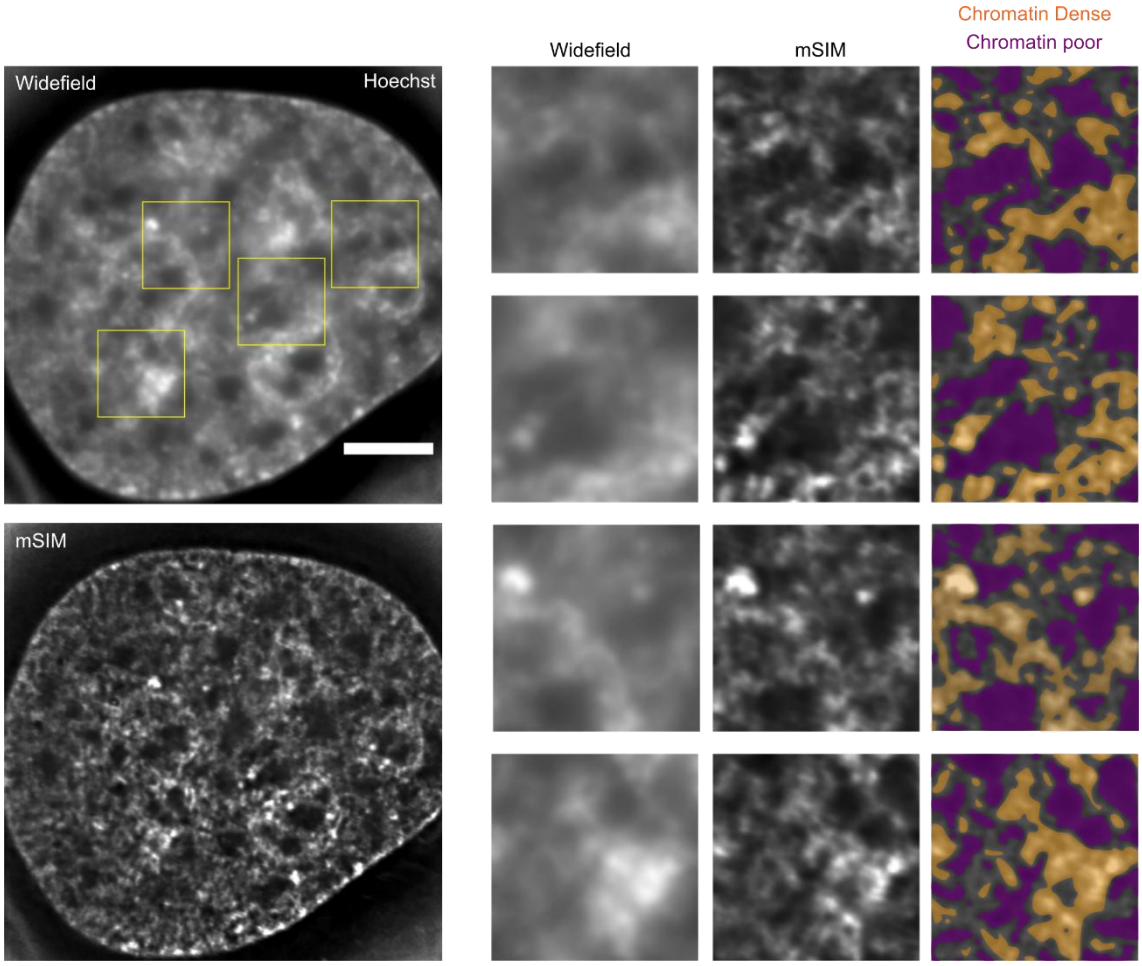


Figure 7.8 Exemplary mSIM acquisition of DNA density. *Conventional widefield and mSIM imaging of a Hoechst 33342-labeled living nucleus. mSIM allows to better map regions of different Hoechst intensity, interpreted by us as different chromatin densities. Chromatin dense and chromatin poor regions have been identified as pixels with the 30% top and 30% bottom fluorescence intensity in the defined ROIs of the mSIM image.*

To investigate how the compartmentalization of chromatin influences the distribution of different p53 populations, we overlay the p53-HT SMT movies to the normalized chromatin maps. Looking at the positions occupied by individual molecules, we find that p53 binds very often within regions dense of chromatin or in their proximity, potentially reflecting its capability of acting as a pioneer factor (Figure 7.9A). Interestingly, the diffusing molecules are frequently excluded from such dense domains (see Figure 7.9A and **Movie 2**). The classification of p53 tracks into bound, slow diffusing and fast diffusing as above described, allows substantiating this observation. Moreover, slow molecules seem to diffuse often close to the most DNA dense nuclear regions (Figure 7.9B).

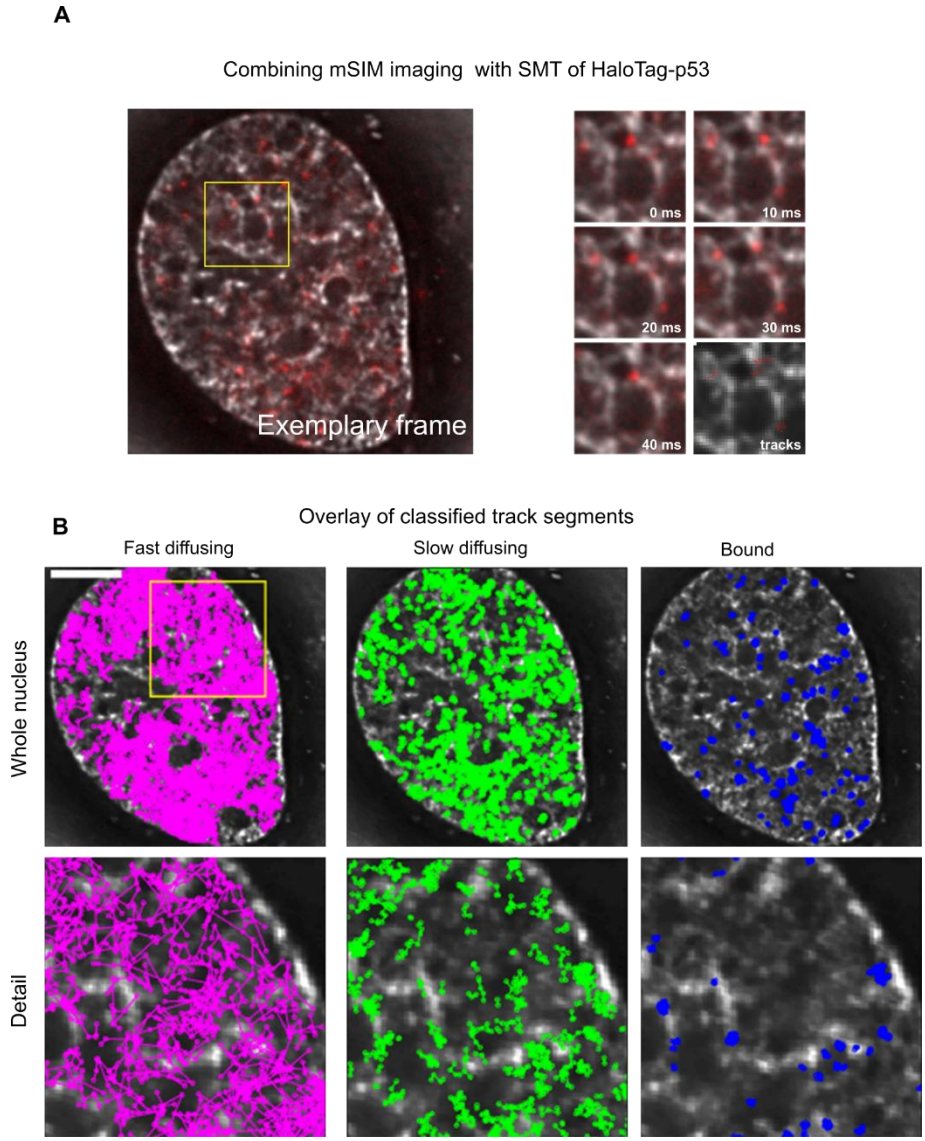


Figure 7.9 SMT/mSIM microscopy to investigate the kinetics of p53 molecules relative to chromatin organization*. We apply live imaging SMT/mSIM to follow p53 search process within chromatin. Labelled with Hoechst 33342, chromatin appears spatially organized in compartments of different density. Visual inspection of single molecules indicates that p53 subpopulations occupy distinct region of the nucleus depending on chromatin compartmentalization. (A) On the left, a representative frame acquired during a SMT/mSIM movie, where single molecule of p53-HT (acquired by SMT, red spots in the figure) diffuse through chromatin (imaged by mSIM, in gray). On the right, a frame sequence capturing the movement of slow molecules while diffusing along the surface of chromatin dense domains. (B) We overlay the segment tracks classified by the vbSPT algorithm into the three kinetic states, on the chromatin map obtained by mSIM. Viisual inspection of different subpopulations highlights that bound p53-HT molecules are positioned within or on the edges of chromatin dense domains, while fast diffusing molecules are preferentially observed in DNA poor-compartments. Scale Bar 5µm.*

To quantitatively evaluate the relationship between the compartmentalization of chromatin and p53 segregation, we first computed the distributions of Hoechst at the location sampled by bound, slowly diffusing and fast diffusing molecules. On average,

p53 diffusing molecules were found at regions with lower Hoechst intensity and lower local gradient than p53 bound molecules (Figure 7.10A). This preferred positioning of bound molecules appears to be peculiar for p53, since repeating the same analysis for the NF-kB sub-unit p65 (upon specific induction by TNF treatment), display bound and free molecules in regions that are undistinguishable from randomly positioned molecules (Figure 7.10B). Notably instead, the chromatin component histone H2B shows accumulation of both bound and free molecules at DNA dense regions (Figure 7.10C).

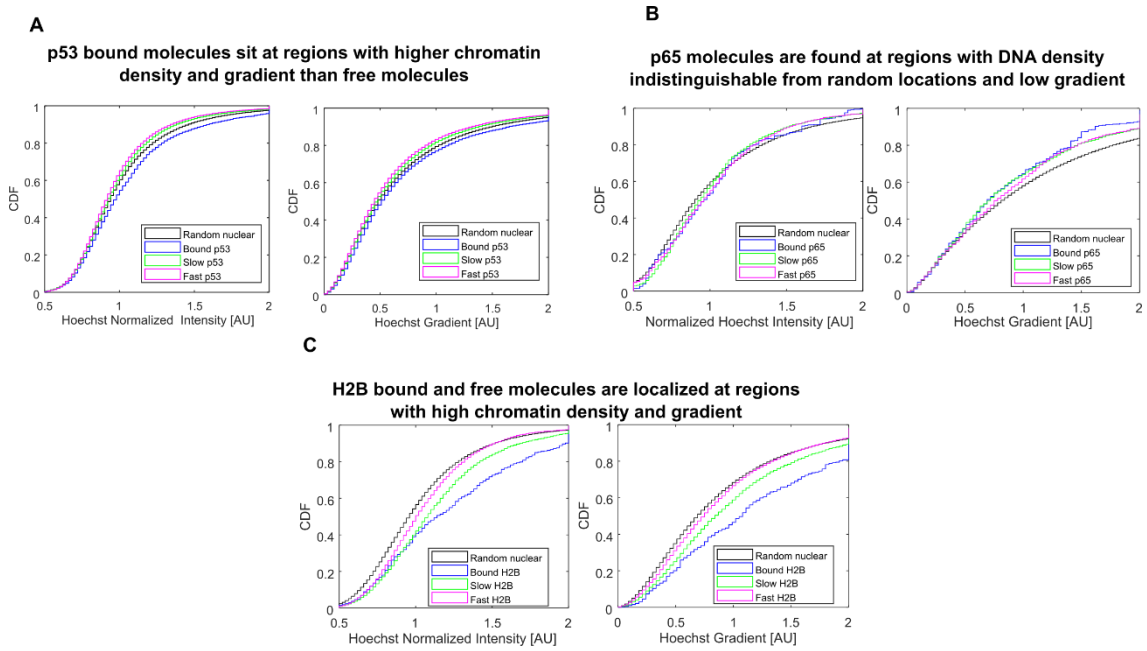


Figure 7.10 Hoechst intensity and Hoechst gradient at regions visited by different nuclear proteins*. (A) p53 bound molecules are located at regions with higher chromatin density and gradient than random locations, while diffusing molecules are found at regions with low Hoechst intensity. (B) The molecules of the NF-kB subunit p65 are found at regions with chromatin density comparable to what would be observed for randomly positioned molecules. (C) Histone H2B molecules are found at regions with high Hoechst intensity and gradient (CDF, cumulative distribution function;* n_{Cells} *= 29, 5, 15* for p53-HaloTag, p65-HaloTag and HaloTag-H2B respectively).

The peculiar localization of p53 molecules is also highlighted by averaging the normalized distribution of Hoechst signal around each classified p53 single molecule (Figure 7.11A). By performing this analysis, we found that fast molecules preferentially occupy chromatin poor regions, surrounded by regions at higher DNA density -such as the channels that compose the interchromatin compartment (IC)- while bound molecules are enriched in denser chromatin domains (CDs), surrounded by DNA poor regions.

Of note, chromatin around slow diffusing p53 molecules, display an intermediate profile, highlighting once again that slow molecules might facilitate the targeting p53 to its binding sites (Figure 7.11A). Finally, the targeting of p53 to the CDs appeared enhanced following its activation by IR (Figure 7.11B).

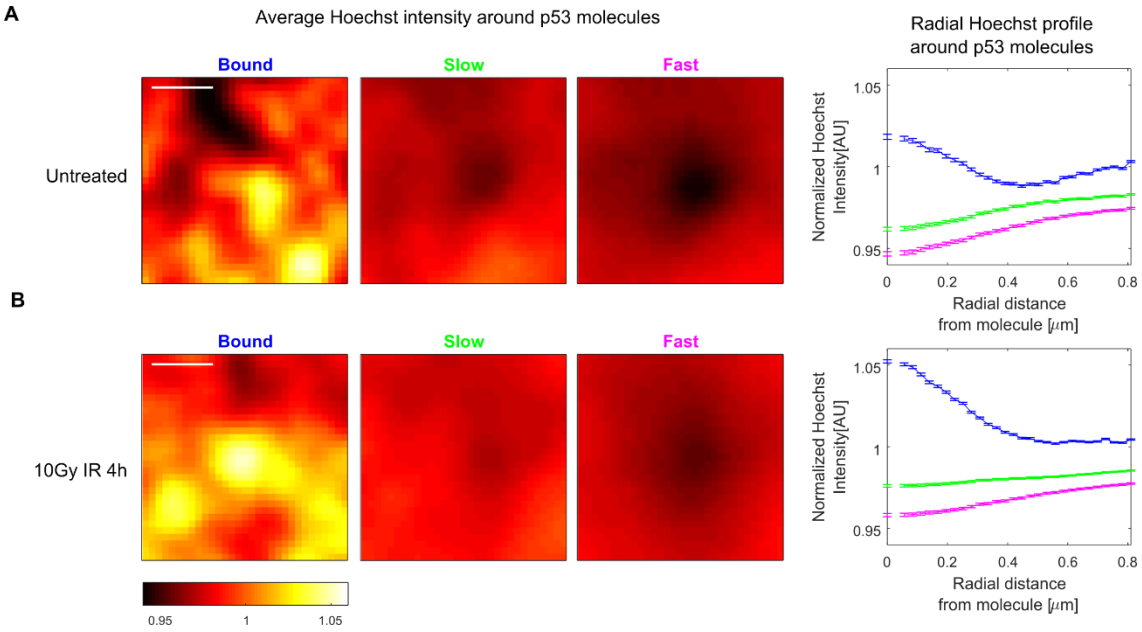


Figure 7.11 Average chromatin radial profile observed around bound, slow diffusing and fast diffusing molecules*. (A) In untreated conditions bound p53 molecules sit on chromatin dense regions surrounded by chromatin regions at lower density, while free molecules are enriched in chromatin poor regions surrounded by denser ones, with slow diffusing molecules displaying an intermediate profile between the ones observed for bound and fast diffusing ones. (B) The targeting of p53 bound molecules to chromatin dense regions is enriched upon the induction of DNA damage (error bars, s.e.m.* n_{Cells} *= 29, 29 for untreated and 10Gy IR respectively).*

In sum, populations of p53 with different mobility preferentially occupy regions with different chromatin density: while being generally excluded from CDs, p53 bound molecules are preferentially positioned at the edges or within these domains, and slow molecules diffuse at regions with intermediate chromatin density, possibly on the surface of CDs.

7.3 Slow and fast diffusing p53 molecules display a different degree of anisotropic nuclear diffusion.

As detailed in the introduction (see **3.1 Principles of the TFs search mechanism**), the search mechanism of TFs can be classified into two universal categories: compact and non-compact exploration. By compact exploration, TFs

densely sample the nuclear space, and -when exploring a certain region- they are likely to visit all the inside sites before going away. If the TF is located in region enriched with its targets, the dense sampling enhances the probability to find them. On the contrary, since compact exploration is time consuming, if there are no targets the compact search will be unproductive. For this reason, the τ_{search} derived from the compact exploration is distance-dependent (i.e. short in case the target is close to, long in case the target is far from the TF). By non-compact exploration instead (e.g. 3D Brownian motion), TFs freely move and -on average- they rapidly abandon a certain region, moving to another one. Here the τ_{search} is distance independent: the TF have equal probability to meet a target close to its real-time position or one far apart. Interestingly, recent data point out that some eukaryotic TFs do not adopt a unique exploration, but instead they switch between compact and non-compact exploration, a behavior named 'guided diffusion' (Hansen *et al*, 2020). In the guided diffusion, a TF molecule exploring non-compactly the nucleus, sometimes get entrapped in local regions and switch to the compact behavior. Even though the molecular details behind this switching are not fully understood, this behavior somehow resembles the 'facilitated diffusion' of bacteria, where TFs combine DNA sliding (compact exploration) and 3D Brownian motion (non-compact).

Since fast and slow p53 molecules together diffuse in distinct compartments as defined by the chromatin architecture, we speculated that they may cooperate to enhance p53 targeting efficiency. Through chromatin channels, fast molecules could cover large distances in regions devoid of targets (indeed bound molecules reside in denser compartments). On the other hand, slow molecules appear important to finalize the search, since they more frequently reach the bound state than fast molecules. However, the property of 'being slow' does not explain either their targeting efficiency nor why they frequently explore the chromatin domain's surface.

To assess whether the transition from slow to fast, and vice-versa, is not merely related to diffusional speed, but underlies a compact/non-compact switching as observed in the guided diffusion, we quantified the 'fold anisotropy' $f_{180/0}$ of diffusing molecules (Hansen *et al*, 2020). The $f_{180/0}$ metric quantifies the likelihood for diffusing molecules to walk back on their steps, and it can be inferred by the direction angles of their movements. More precisely, all the angles θ between three consecutive positions in a track are collected (Figure 7.12). $f_{180/0}$ measures how many times a step backward (θ=180°±30°) is more likely than a step forward (θ=0°180°±30°).

Because molecules moving of Brownian motion change directions randomly over time, all angles are equally probable, and $f_{180/0}$ will be 1 (isotropic diffusion). On the contrary, molecules entrapped in a confined region tend to step backward and have

$f_{180/0}$ >1 (backward anisotropy diffusion). A TF that searches for its targets uniquely by compact or non-compact exploration, will display backward anisotropy or isotropic diffusion, respectively, at all length scales. Differently, in guided diffusion, the anisotropy is not the same at all scales. Because guided exploration is defined by local compact-exploration within 'trapping zones' of a defined size, this induces a peak of backward anisotropy at spatial scales comparable to the size of the trap. At longer scales instead, diffusion becomes more isotropic.

We therefore collected all the angles from our p53-HT diffusing molecules (fast and slow molecules together) -while excluding the bound ones- and applied the $f_{180/0}$ metric. Similar to guided diffusion, in non-treated conditions, p53 displays a slight backward anisotropy at small length scales, indicating a local confinement, followed by a gradual decrease at longer displacement distances. Interestingly, when p53 gets activated by DNA damage, the small-scale backward anisotropy is further increased, reaching a peak at about 100 nm (Figure 7.12).

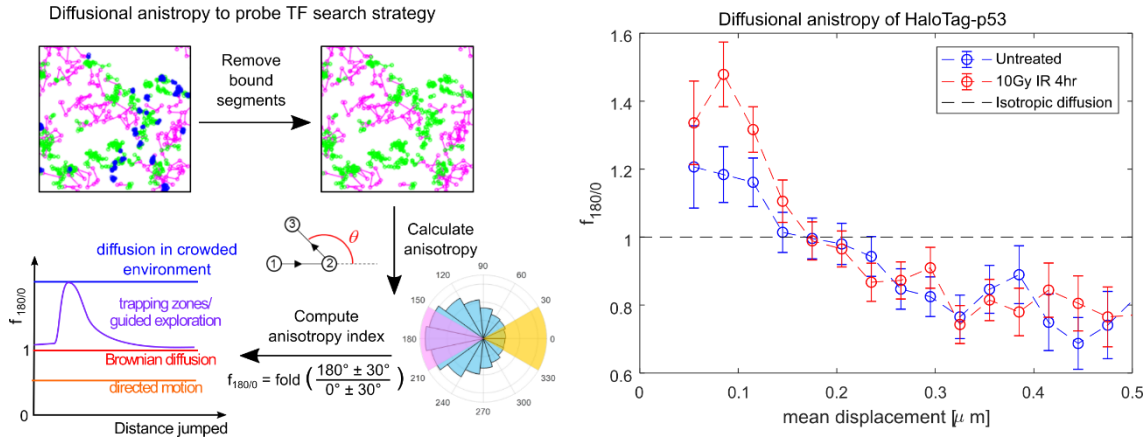


Figure 7.12 Analysis of p53 diffusional anisotropy. *(Left) After filtering out p53 bound molecules, we analyzed diffusional anisotropy of mobile molecules, by calculating the anisotropy index $f_{180/0}$ as function of the mean displacement. Such analysis allows to discern among different exploration strategies ranging from diffusion in crowded environment to isotropic diffusion and guided exploration. (Right) Upon activation by DNA damage the diffusional anisotropy profile of HaloTag-p53 is compatible with guided exploration (error bars: s.e.m, n_{cells} = 45, 43 for untreated and 10Gy IR respectively).*

This dual behavior with local backward anisotropy followed by a decay is compatible with a guided exploration mechanism. Moreover, the increase of backward anisotropy after DNA damage indicates that p53 guide exploration could be finely regulated during of p53 activation. As control of our approach, we verified the anisotropic diffusion of the architectural protein CTCF, recapitulating the results generated by others (Hansen *et al*, 2020). As expected, while CTCF displays the signature of guided exploration, the unconjugated HaloTag peptide diffuses more

isotropically (Figure 7.13). Finally, the NF-kB sub-unit p65 (an unrelated transcription factor) also shows nearly isotropic diffusion (Figure 7.13).

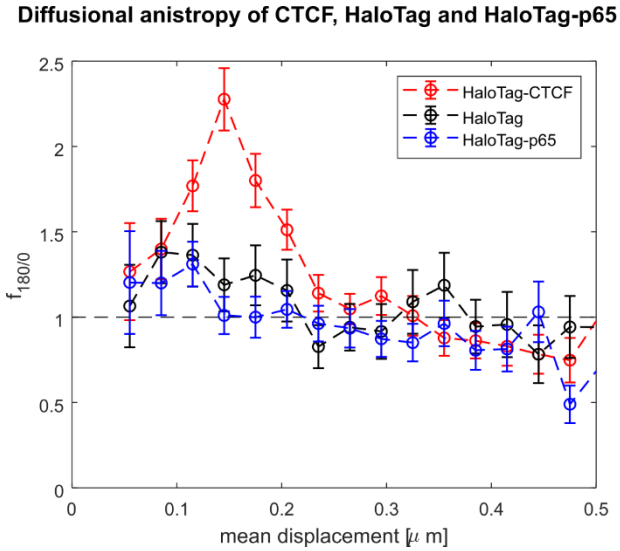


Figure 7.13 Diffusional Anisotropy of CTCF, unconjugated HaloTag and p65-HaloTag. CTCF display an anisotropy profile compatible with guided exploration, while unconjugated HaloTag and p65-HaloTag display more isotropic diffusion (n_{cells} = 14, 11, 8 for CTCF, unconjugated HaloTag and p65 HaloTag, respectively).

Our results highlight a peculiar search mechanism for p53, with an intermittent 'guided' search, where fast and slow states interchange depending on chromatin compartmentalization. Diffusing through interchromatin channels, fast molecules rapidly reach denser CDs. At the interface between channels and CDs, p53 slows down and get locally confined, allowing for an accurate compact sampling.

7.4 The anisotropic behavior of slow p53 molecules is controlled by the protein IDRs.

The data described so far indicate that p53 somehow 'perceives' different chromatin compartments modulating the type of exploration. Previous studies on nuclear proteins displaying 'guided exploration' (Hansen *et al*, 2020) or transient confinement (Garcia *et al*, 2021b), linked this behavior to the capability of their IDRs to mediate weak interactions with the nuclear environment. A recent report also indicates that swapping the IDRs between TFs can have a strong impact on their mobility (Chen *et al*, 2021).

To investigate the mechanism underlying p53 guided exploration, we reasoned that p53 IDRs may be responsible for its intermittent compact exploration. Both p53 N-terminal and CTD are reported to be intrinsically disordered, observation that we

further confirmed by the PONDR algorithm (http://www.pondr.com/; Figure 7.14**,** left).

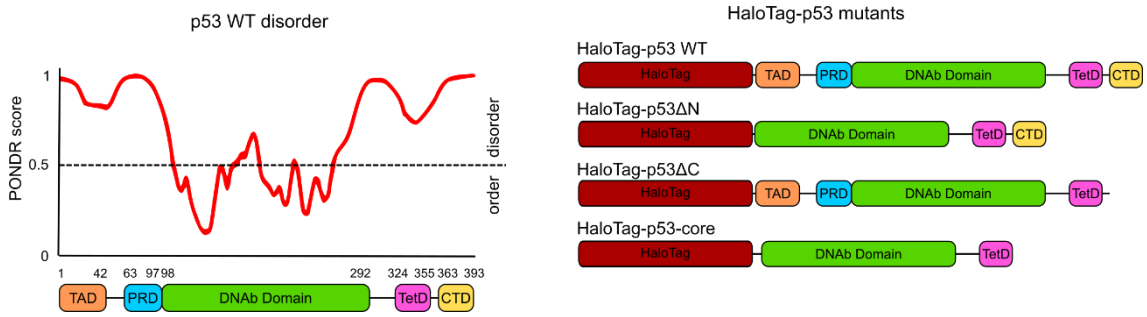


Figure 7.14 p53 Disorder and scheme for the HaloTag-p53 mutants lacking IDR regions used in this thesis. *On the left, protein disorder measured by PONDR, highlighting the p53 N-terminal and the C-terminal intrinsically disorder regions. On the right, p53 mutants without IDRs.*

To test our hypothesis, we designed three p53-HT mutants lacking the IDRs (see Figure 7.14, right). Specifically, (i) an N-terminal deletion mutant, without both the p53 transactivation domain and its adjacent proline-rich sequence, therefore incapable of inducing transcription; (ii) a CTD truncated mutant, previously reported to display altered DNA nonspecific interactions both in vitro and in vivo and a reduced DNA speed in vitro; (iii) a p53-Core mutant, consisting only of the DNA binding domain and the p53 tetramerization region.

We measured their diffusion properties and compared their behavior to wild-type p53-HaloTag. To avoid the formation of hetero-oligomers that could hamper the analysis, we transfected the different p53 constructs in our DIvA p53 KO cells (**Movie 3**). The overexpression of p53 WT-HaloTag in a p53-null background induces its activation, as previously demonstrated in other cell lines in our lab, and as we further verified by RNA-FISH in the current settings (see section **7.6 The p53 C-terminal IDRs controls the burst frequency of its target genes.**).

Reassuringly, by transient transfection we could recapitulate the same guided exploration mechanism we observed for knocked-in p53-HaloTag, with the overexpressed p53 WT-HaloTag showing a peak of backward anisotropy at short displacement distance. Importantly, all the tested mutants displayed a reduced backward anisotropy peak, and an unchanged anisotropy profile at long displacement distances (Figure 7.15A). This highlights that p53 guided exploration is at least partially mediated by its IDRs. Of note, when analyzing their displacement distributions, the diffusive populations of all mutants display higher diffusion coefficients and the mutant lacking the CTD shows a reduced fraction of molecules

engaged in chromatin binding interactions (Figure 7.15B). Remarkably, for all mutants the slow diffusing populations co-localize less frequently with bound molecules compared to p53 WT (Figure 7.15C).

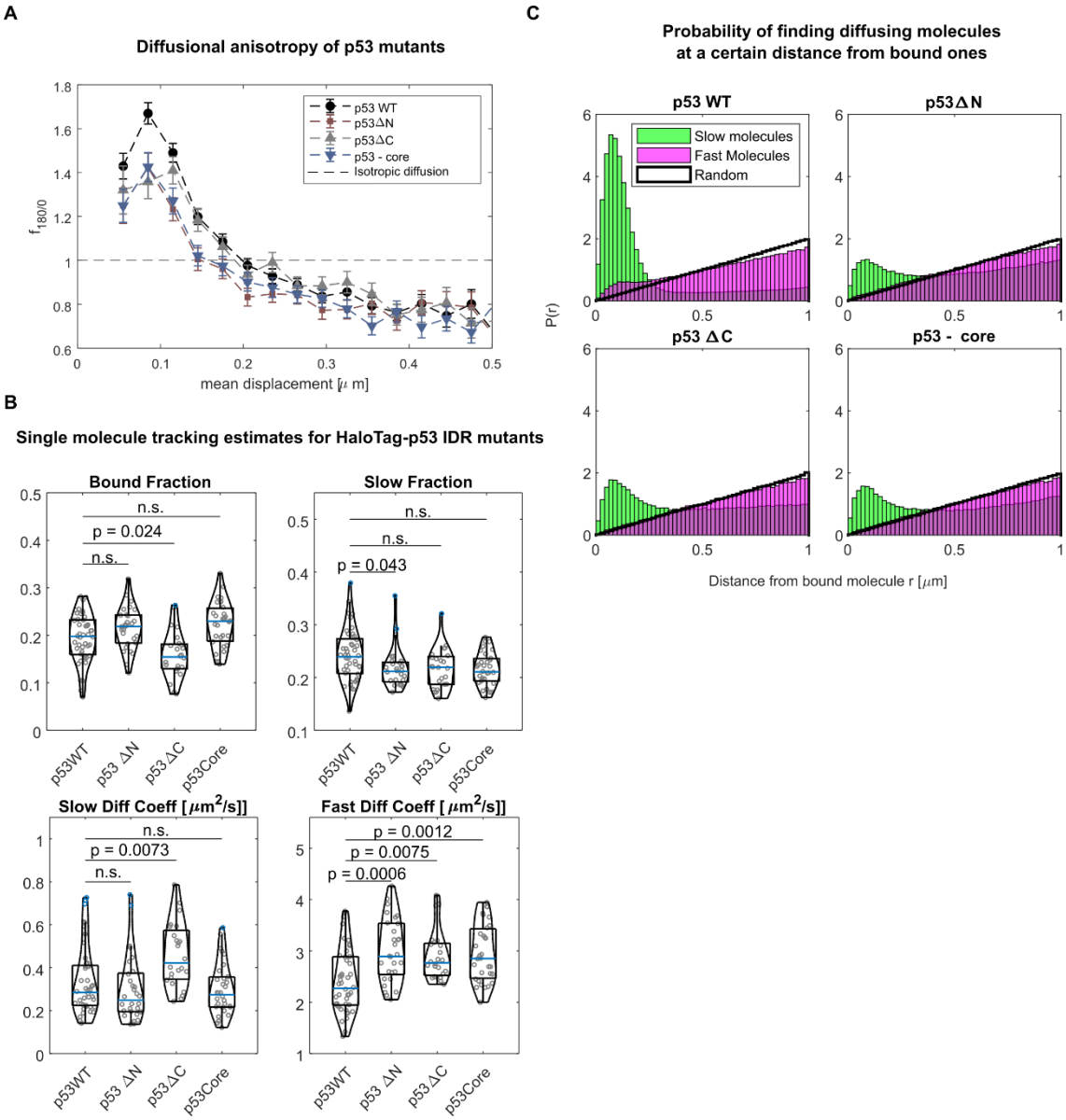


Figure 7.15 p53 exploration is governed by its IDRs. *The HaloTag-p53 mutants were transfected in p53 KO cells and analyzed by paSMT. (A) p53 mutants lacking IDRs display reduced anisotropy at short displacements, (B) increased diffusion speed for their fast components and (C) less co-localization than p53 WT between bound and slowly diffusing molecules (n_{cells} = 45, 29, 24, 31 for p53 WT, p53ΔN, p53ΔC and p53-core respectively).*

This corroborates further that slow molecules -by getting 'trapped'- might facilitate the targeting of p53 to its binding sites, and that such trapping is mediated by IDRs interactions. Surprisingly -though- mutants lacking the N-terminal domain (i.e. p53ΔN and p53-core) show an amount of bound molecules comparable to p53 WT

(Figure 7.15B). This apparent discrepancy suggest that the N-terminal domain may modulate several aspects of p53 binding. Indeed, TF bound fractions arise from two kinetic processes: the search mechanism (regulating the search time for DNA target sites) and the stability of TF-DNA interactions (determining for how long a TF stays bound, a parameter known as the residence time; (Mazzocca *et al*, 2021a)). Despite mutants without N-terminal domain display the signature of an impaired search mechanism, suggesting that they reach DNA targets less frequently, this could be counterbalanced by staying longer on DNA once bound, with overall unchanged bound fraction. Further SMT experiments estimating the residence time may clarify this possibility.

We finally analyzed whether the lack of functional IDRs might influence the targeting of p53 on chromatin domains, by applying our combined SMT/mSIM approach. The IDR mutants displayed a reduced capability than p53 WT to penetrate in chromatin dense domains, and are more likely to sit outside (Figure 7.16).

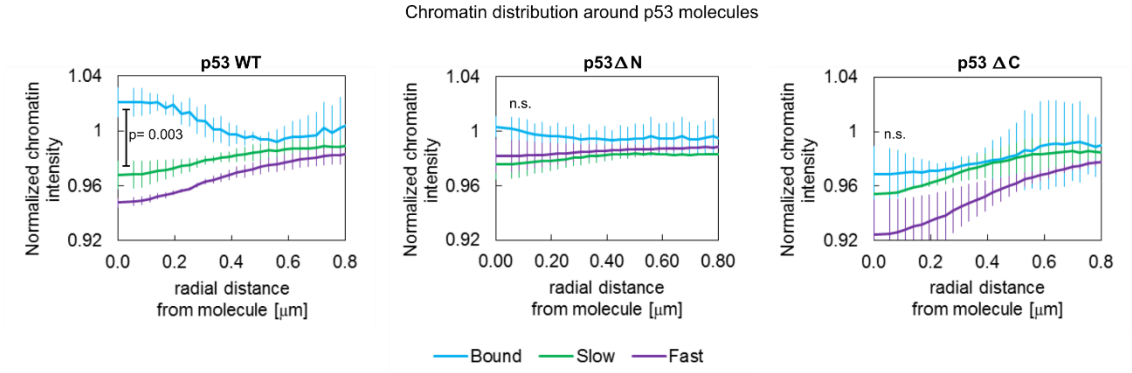


Figure 7.16 Chromatin density around p53 single molecules. *The HaloTag-p53 mutants were transfected in p53 KO cells and analyzed by the combined SMT/mSIM approach, showing that differently from p53 WT, mutants lacking IDRs are less efficiently targeted to chromatin-dense regions.*

Together, our data allows us to derive a model in which the capability of p53 to reach its binding sites in chromatin dense regions is regulated by its IDRs, that induce p53 to slow down and to explore the environment more exhaustively when approaching to these regions, ensuring a more efficient targeting.

7.5 Regulated p53 target genes reside on highly-connected chromatin regions.

Our data suggest that p53 IDRs guide p53 within chromatin dense binding sites, where its binding is enriched. We therefore hypothesized that p53 target genes could locate at dense chromatin regions. We first performed a DNA fish on the *CDKN1A*

gene, measuring its position relative to chromatin (Figure 7.17). Interestingly, the *CDKN1A* locus appears to have a slight preference of being positioned at dense chromatin sites, but close to regions at lower compaction, in agreement with our hypothesis. This result, however, suffers from two major drawbacks: first, DNA FISH requires the denaturation of the DNA double helix, presumably perturbing chromatin structure. Indeed, our SIM images of DNA density appear to display a more uniform chromatin distribution than what was observed above in live cells (compare Figure 7.17 with Figure 7.9). Second, DNA FISH is limited to observe only few loci of interest at a time.

To overcome these limitations, we decided to investigate whether p53 target genes are located at dense chromatin positions also genome-wide. Since p53 typically regulates only a fraction of all its potential target genes (Nguyen *et al*, 2018; Fischer, 2017), we first identified the subset genes successfully regulated by p53 in our cells.

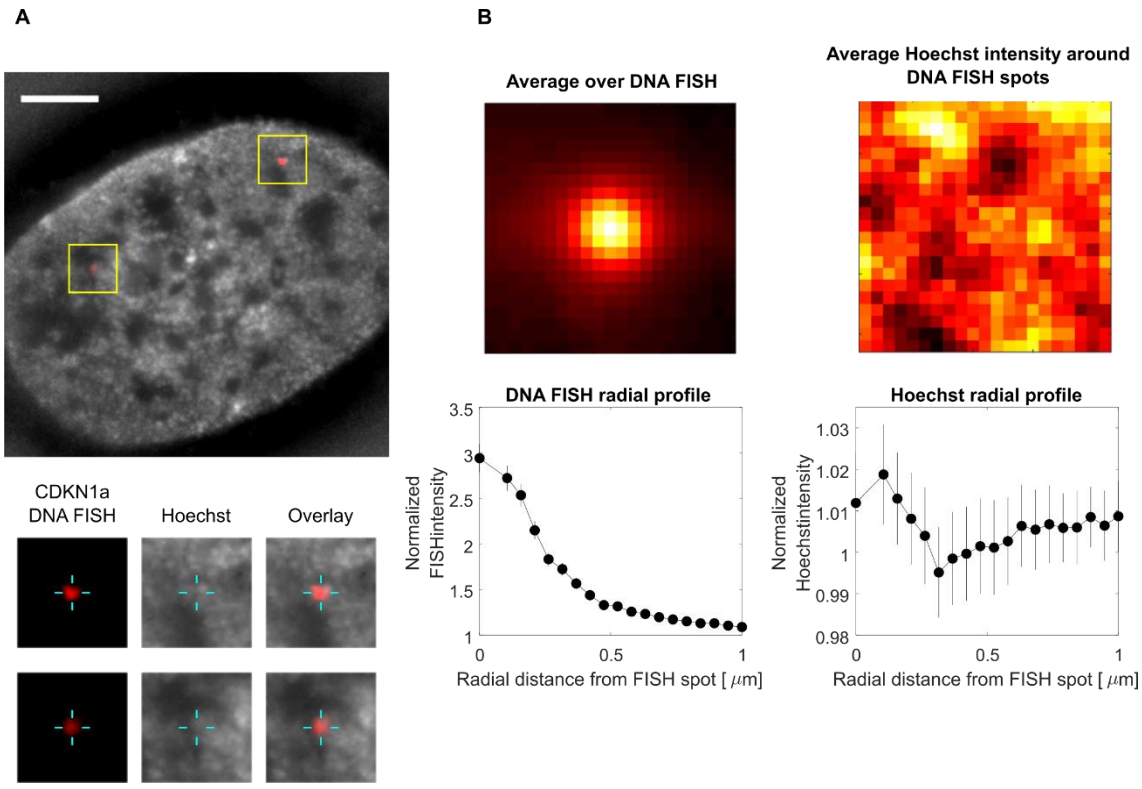


Figure 7.17 Chromatin density around the CDKN1A locus assessed by combining DNA fish with mSIM imaging of Hoechst-labeled DNA. *(A) Exemplary DNA FISH acquisition (red: CDKN1A locus, grey: Hoechst 33342, scale bar 5 µm). Insets show the localization of CDKN1A spots. (B) We averaged together 2 µm ROIs centered around DNA FISH spots from 31 cells and computed the average radial profile of DNA density. On average, the CDKN1A locus appears to be positioned in regions with higher chromatin density than its surroundings.*

To this end, we generated RNA-sequencing (RNA-seq) datasets in untreated conditions and upon DNA damage by 10Gy IR, and considered for our analysis only

upregulated genes, since p53 is largely accepted to be a transcriptional activator (as discussed in the paragraph **1.1 p53 as a transcription factor involved in tumor suppression**). We identified 2,681 upregulated genes (with fold change ≥ 2 and FDR < 0.2) in our DIvA p53 WT compared to p53 KO cells, and further 99 genes upregulated after irradiation in p53 WT cells only.

Reassuringly, enrichment analysis on the overall 2,780 genes reveals a strong dependence on p53 transcriptional network (Figure 7.18).

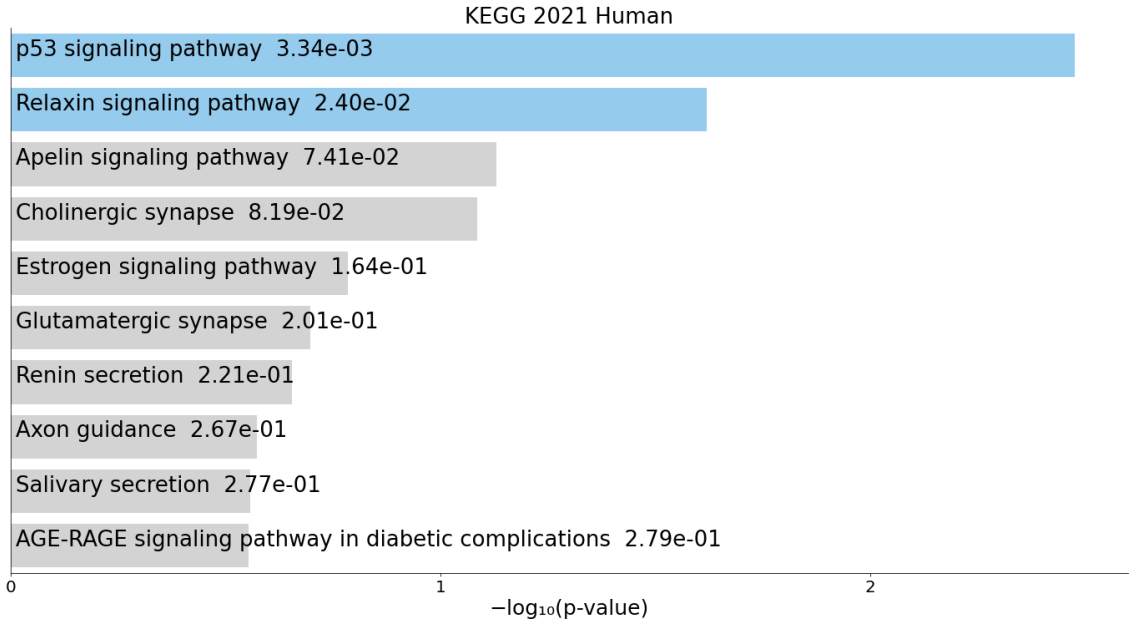


Figure 7.18 Enrichment analysis of genes specifically upregulated in p53 WT but not in p53 KO cells. *Enrichment analysis highlighting the pathways related to genes upregulated in DIvA cells (p53 WT) compared to p53KO cells (analysis performed with Enrichr).*

To exclude genes regulated only indirectly by p53, we decided to restrict our analysis on a core of high-fidelity genes, disregarding all the potential targets that still lack of a general consensus. To this end, we referred to two recent meta-analyses that collected data from independent genome-wide studies on p53 (Nguyen *et al*, 2018; Fischer, 2017). In brief, the two meta-analyses independently identified high-confidence p53 targets, which are defined as genes regulated by p53 across multiple datasets. To further reduce the risk of 'false positives' targets, we considered only the genes commonly identified by both studies. This defines a core of 175 genes that we call the 'p53 cistrome' according to a common nomenclature.

Comparing the p53 cistrome with our RNA-seq datasets we identified 113 high-confident target genes regulated by p53 in our cells (Figure 7.19).

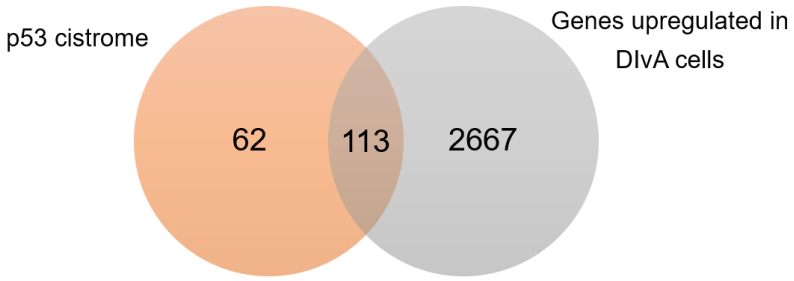


Figure 7.19 p53 target genes regulated in DIvA cells. *Genes upregulated in DIvA p53 WT compared to p53 KO, were matched to a core of high-fidelity target genes (p53 cistrome). The analysis identifies 113 genes upregulated by p53, 62 cistrome genes not regulated in DIvA and 2667 genes upregulated that are not part of the above-defined p53 cistrome list.*

To investigate whether these genes are located at dense chromatin sites, we next applied a method estimating the compactness of genomic regions from Hi-C data (Cortini & Filion, 2018), using a recent Hi-C dataset generated in DIvA cells (Figure 7.20A; (Arnould *et al*, 2021)). This approach quantifies the amounts of genomic interactions for any desired sequence in the genome and provides a chromatin 'compactness' index. Higher values of the compactness index correspond to more compact/dense chromatin (Cortini & Filion, 2018). For instance, genomic regions highly occupied by TFs and characterized by open chromatin, displayed lower compactness than average random sites in the genome.

By applying the same concept, we quantified the compactness index of p53 targets in DIvA cells, and compared our genes to both open chromatin regions identified as sites occupied by multiple TFs (known as 'extremely occupied targets' (xHOTs); (Boyle *et al*, 2014)) and dense heterochromatin sites, provided by the lamina-associated domain (LADs; (Guelen *et al*, 2008); Figure 7.20B). Interestingly, promoters of p53 regulated genes sit in regions at higher compaction than sites normally occupied by TFs, consistent with our observation that p53 search and reach its targets in dense chromatin domains. We next wondered if chromatin compaction could explain p53 selectivity, with non-regulated genes that occupy regions less accessible. We compared the compactness index of p53 targets regulated and non-regulated in our cells (Figure 7.20C). To our surprise, we found that promoters of non-regulated targets occupy lower density regions on average, suggesting that indeed regulated p53 target sites sit – on average – at regions with higher chromatin density than both non-regulated p53 target genes and regions accessed by multiple transcription factors.

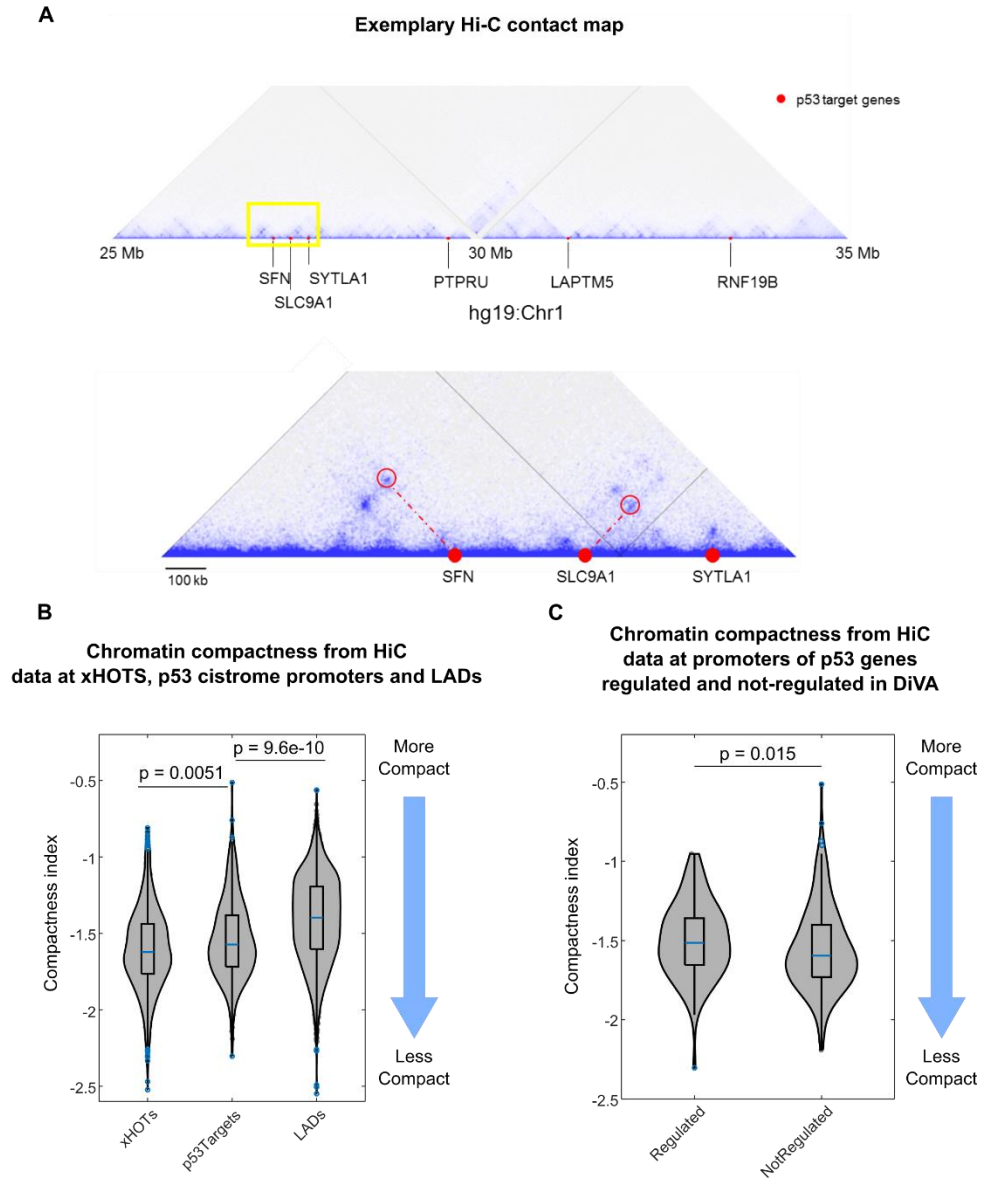


Figure 7.20 Analysis of chromatin compactness at p53 target genes from Hi-C*. (A) Visual inspection of Hi-C contact maps from DIvA cells, reveals that p53 target genes are often positioned in highly connected chromatin regions such as at the TAD boundaries (e.g. PTPRU, LAPTM5), within TADs (RNF19B, STYLA1) or at chromatin loops (red circles, e.g. SNF, SLC9A1). (B) We analyzed chromatin compaction by the same strategy used in (Cortini & Filion, 2018). Promoters of the p53 cistrome are found at regions of intermediate compactness compared to Highly Occupied Targets (xHOTs) and highly inaccessible Lamina associated domains (LADs). (C) Further, by stratifying the p53 cistrome into those targets that appear regulated in DIvA cells in our RNA-seq dataset, we observed that p53 regulated target genes are on average in chromatin regions at higher compactness than non-regulated genes.*

7.6 The p53 C-terminal IDRs controls the burst frequency of its target genes.

Guided exploration has been described as a means to speed up the search mechanism of nuclear proteins. In case of TFs, a faster search mechanism should result in an increase of the frequency by which target genes are activated (i.e. the more frequently a TF find a given gene, the more frequently the gene will be turned on). To test this hypothesis, we expressed WT or IDR mutants over a p53-null background and measured the expression of the p53 target gene *CDKN1A* at the single cell level. Using single molecule fluorescence in situ hybridization (smFISH, Figure 7.21A), we quantified the number of expressed mRNA molecules per cell.

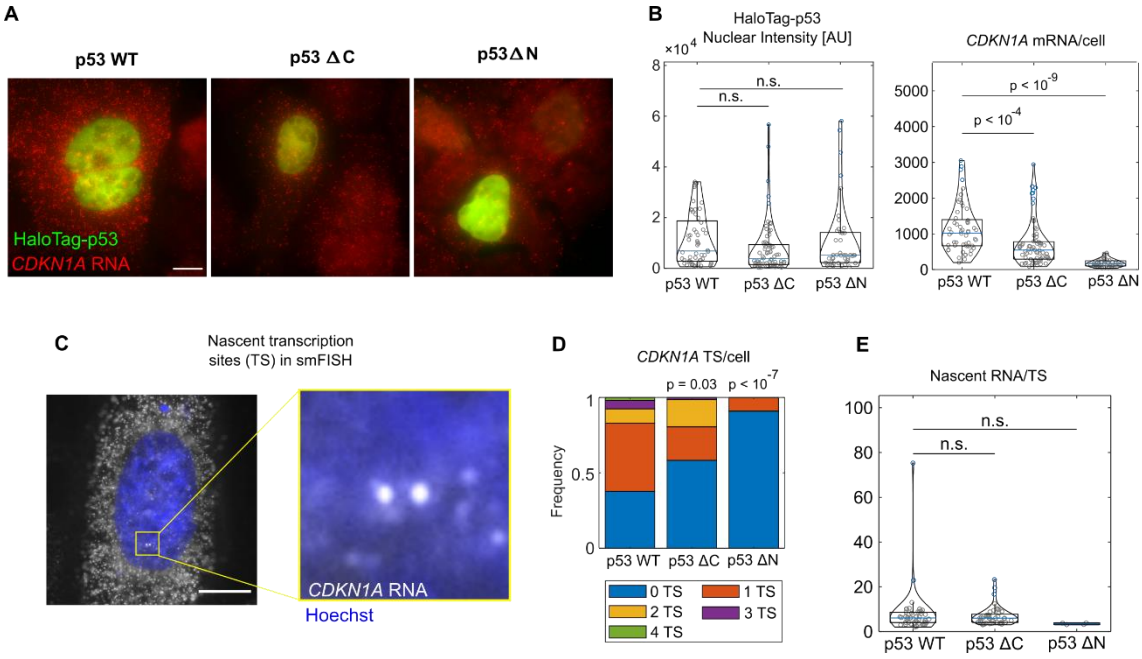


Figure 7.21 p53 CTD controls the expression of CDKN1A mRNA. *(A) Exemplary smFISH acquisitions. Shown is maximal projection of 3D stack. Scale bar 5 µm. (B) p53 expression levels (left) and CDKN1A mRNA expression at the single cell level (n_{cells}=53, 72, 42 for p53 WT, p53ΔC, p53ΔN respectively. Statistical test Kolmogorov-Smirnov, with Bonferroni correction for multiple testing). (C) smFISH allows to count nascent RNA at transcription sites (TS) that appears as bright dots in the nucleus. (D) p53-ΔC results in less frequent activation of TS (statistical test: Fisher-exact test) than p53 WT, but the number of nascent RNAs per TS (E) is similar in cells expressing WT or mutant p53 (statistical test: Statistical test Kolgomorov-Smirnov, with Bonferroni correction for multiple testing).*

Unsurprisingly, cells transfected with p53-ΔN -lacking the TF transactivation domain- showed lower mRNA counts than cells bearing p53 WT (Figure 7.21B). Notably, however, cells transfected with p53 WT also displayed higher mRNA counts than the CTD mutant (p53-ΔC, Figure 7.21B), despite the two constructs being

expressed at similar levels. Such difference across p53 mutants was also reflected at the nascent transcription level, by quantifying the number of active transcription sites (TS) per cell, that appear as 1 to 4 bright foci in the nucleus (Figure 7.21C). p53 WT results in a higher number of *CDKN1A* TS/cell than both p53 truncated mutants (Figure 7.21D), while the number of nascent transcripts per TS does not change significantly (Figure 7.21E), – in agreement with the hypothesis that the different p53 constructs result in a different frequency of activation of transcription. To further corroborate this hypothesis, we extended the smFISH analysis to two other p53 target genes (*MDM2* and *BCC3*, encoding for the PUMA protein) and we analyzed the distribution of mRNAs in single cells with a simple two-state model. The model (see Material and Methods, **11.4 Analysis of smFISH data**) describes a gene switching stochastically between an off and an on state. This model computes the 'relative burst frequency' (i.e. how frequently the gene switches between the 'off' and the 'on' state divided for the mRNA degradation rate) and the 'burst amplitude' (i.e. how many RNAs are synthesized during an 'on' period on average). For the p53-ΔC mutant, the analysis indicates a reduced burst frequency of all the tested genes (Figure 7.22), in line with the hypothesis that this mutant reach less frequently its binding sites.

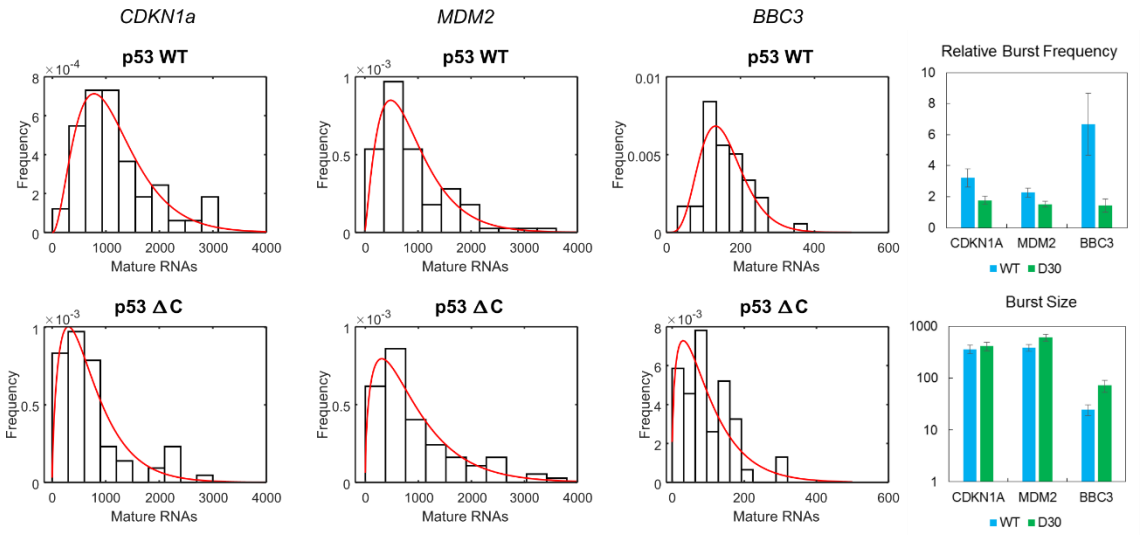


Figure 7.22 Modelling distribution of mRNAs of p53 target genes. *The mRNA expression of three different genes is analyzed using a bursting model, to compute the burst frequency (i.e. frequency of activation of each gene) and the burst size (i.e. how long a gene stays turned on). All three tested genes display a higher relative burst frequency upon p53 WT expression, than for p53-ΔC.*

In sum, these results suggest that p53 IDRs warrant an efficient targeting of p53 to its genes, modulating -in turn- the frequency by which the genes are transcribed.

7.7 A strong IDR fused to p53 modulates its search strategy.

To further confirm the role of IDRs in controlling the p53 search mechanism, we fused an exogenous (IDR), derived from the N-terminal domain (aa 1-214) of the FUS protein, to the full p53 coding sequence (Figure 7.23A). The so-generated HaloTag-FUS-p53 construct, displayed nuclear localization, and differently from p53 WT, the capability of forming visible condensates in vivo at expression levels higher than a certain critical concentration (located around the 60-th percentile of our transfected cell population; Figures 7.23 B-C).

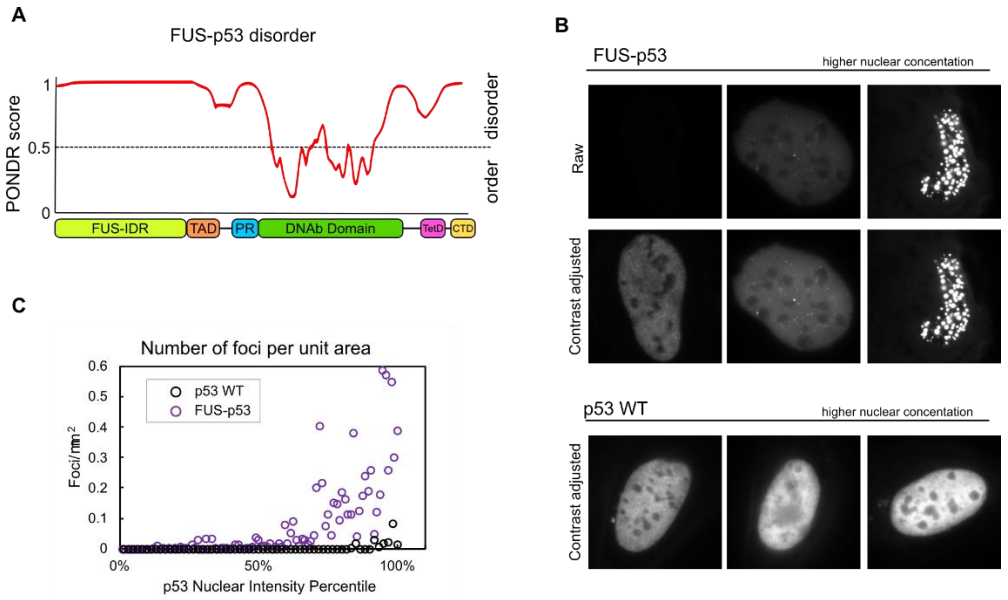


Figure 7.23 Characterization of the FUS-p53 construct. *(A) PONDR analysis to estimate the intrinsic disorder of the FUS-p53 construct. (B) FUS-p53 forms nuclear condensates above a certain critical concentration (approximately at the 60th percentile of our transfected cell population (C)).*

When analyzing FUS-p53 mobility at the single molecule level by paSMT, this mutant displayed a reduced mobility compared to the WT. The diffusion coefficients of FUS-p53 were lower and a larger fraction of FUS-p53 molecules were classified as bound and slowly diffusing (Figure 7.24A), with almost a two-fold drop in the fraction of fast diffusing molecules compared to p53 WT. Moreover, diffusing FUS-p53 displayed a stronger signature of guided exploration with a more prominent anisotropy peak at spatial scales below 150nm (Figure 7.24B). Taken together, these data indicate that FUS-p53 spends more time in the slow, highly anisotropic, diffusion state, and therefore samples the nuclear environment in a more compact fashion,

supporting the notion that TFs IDRs play a major role in determining the search strategy of TFs. We next employed our SMT/mSIM approach to monitor whether this higher compact exploration improves the capability of targeting FUS-p53 to dense chromatin domains. To our surprise, however, chromatin dense regions appear -on average- less permeable to FUS-p53 than to WT-p53 (Figure 7.24C).

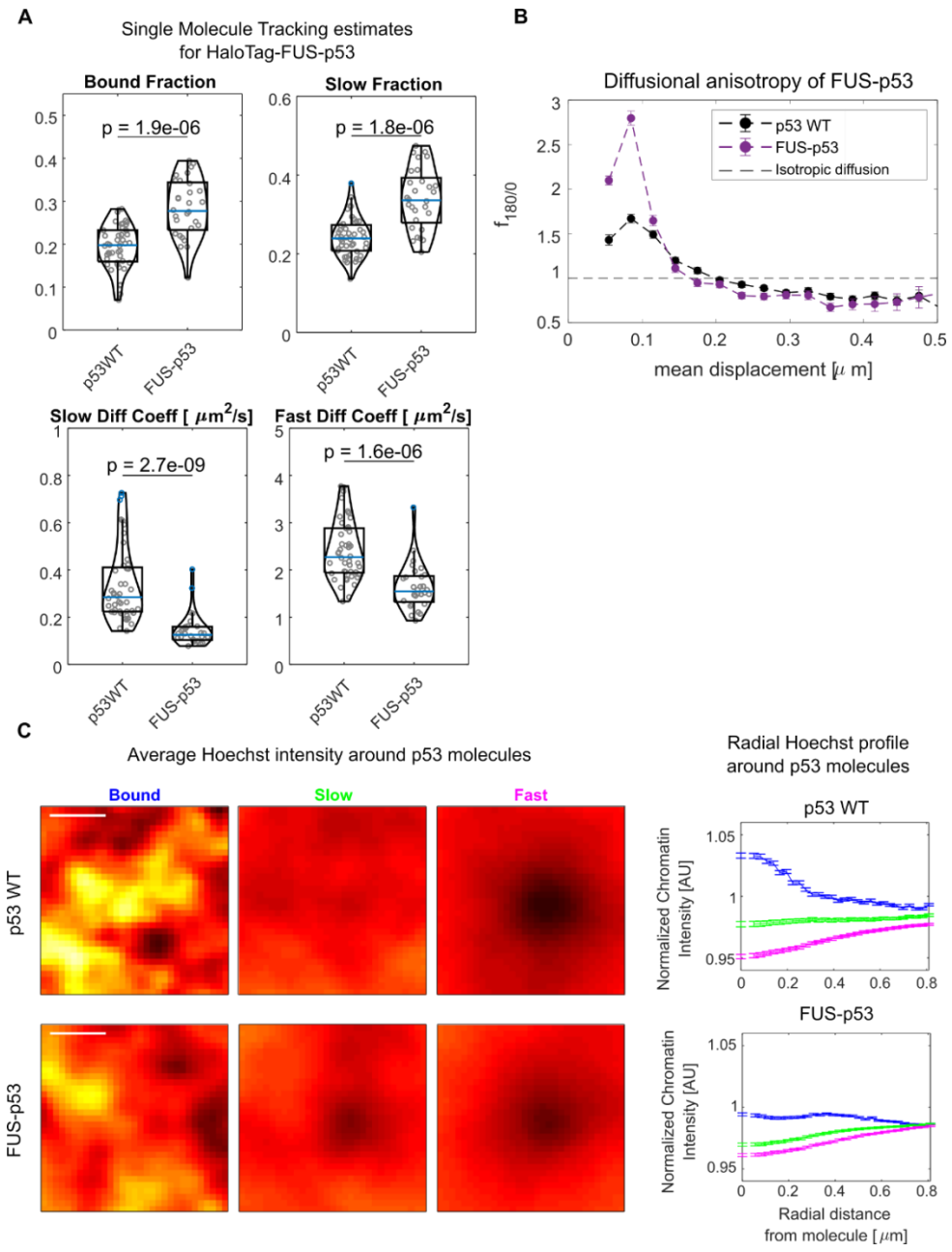


Figure 7.24 Fusing a strong IDR to p53 renders its diffusion more compact, but interferes with its targeting to CDs. *(A) Analyzed by paSMT, FUS-p53 displays a higher fraction of molecules in bound and slow diffusing states, slower diffusion coefficients and (B) a more prominent diffusional anisotropy peak (n_{cells} = 45, 29 for p53 WT and FUS-p53 respectively, statistical test Kolmogorov-Smirnov, with Bonferroni correction for multiple testing). (C) SMT/mSIM analysis of chromatin distribution around p53 molecules shows that*

FUS-p53 is less capable of penetrating chromatin dense regions (n_{cells} = 18, 35 for p53 WT and FUS-p53 respectively).

This unexpected last finding, with FUS-p53 that -despite higher anisotropy- is less frequently recruited to dense chromatin regions, led us to probe the transcriptional potential of FUS-p53. By smFISH, we quantified the expression level of *CDKN1A* and *MDM2*. On average, cells expressing FUS-p53 displayed comparable levels of *CDKN1A* transcripts and more *MDM2* mRNAs to those expressing p53 WT (Figure 7.25).

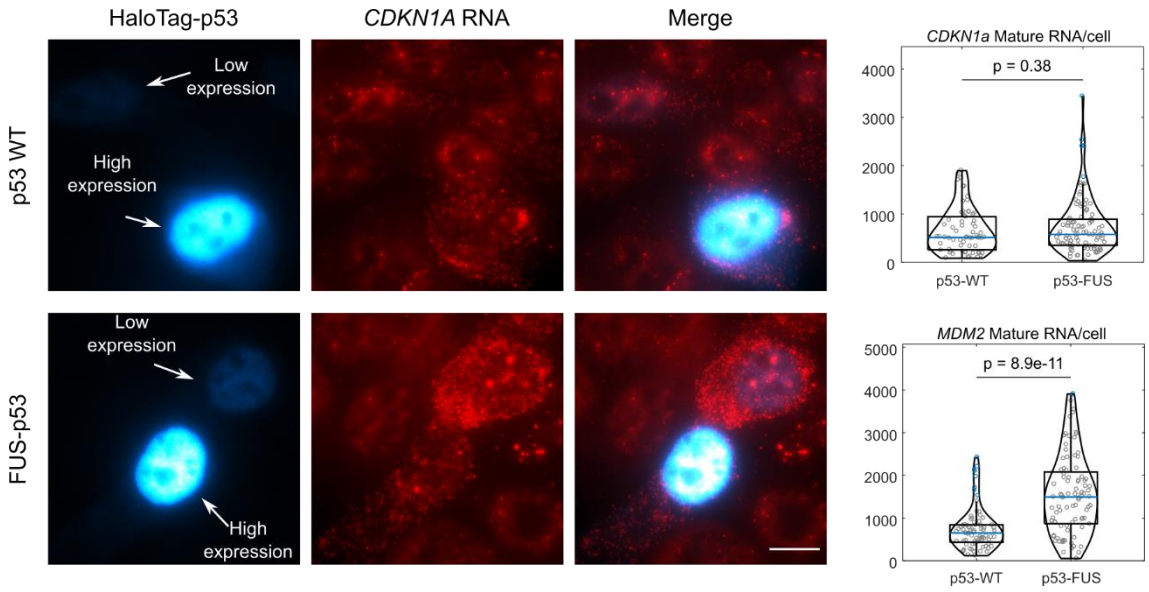


Figure 7.25 smFISH for p53 targets in cells transduced with p53 WT and FUS-p53. *On average, FUS-p53 results in similar activation of the target gene CDKN1A but 2-fold higher expression of MDM2, compared to p53 WT (n_{cells} = 65, 102, 83, 96 for p53 WT/CDKN1A, FUS-p53/CDKN1A, p53 WT/MDM2, FUS-p53/CDKN1A respectively, Statistical test Kolmogorov-Smirnov, with Bonferroni correction for multiple testing).*

At the single cell level, however, p53 WT and FUS-p53 resulted in dramatic differences in the transcriptional output as function of the TF expression levels. Transcripts of *CDKN1A* and *MDM2* were found to increase with p53 WT levels, as expected for a transcriptional activator (Figure 7.26 and Figure 7.27).

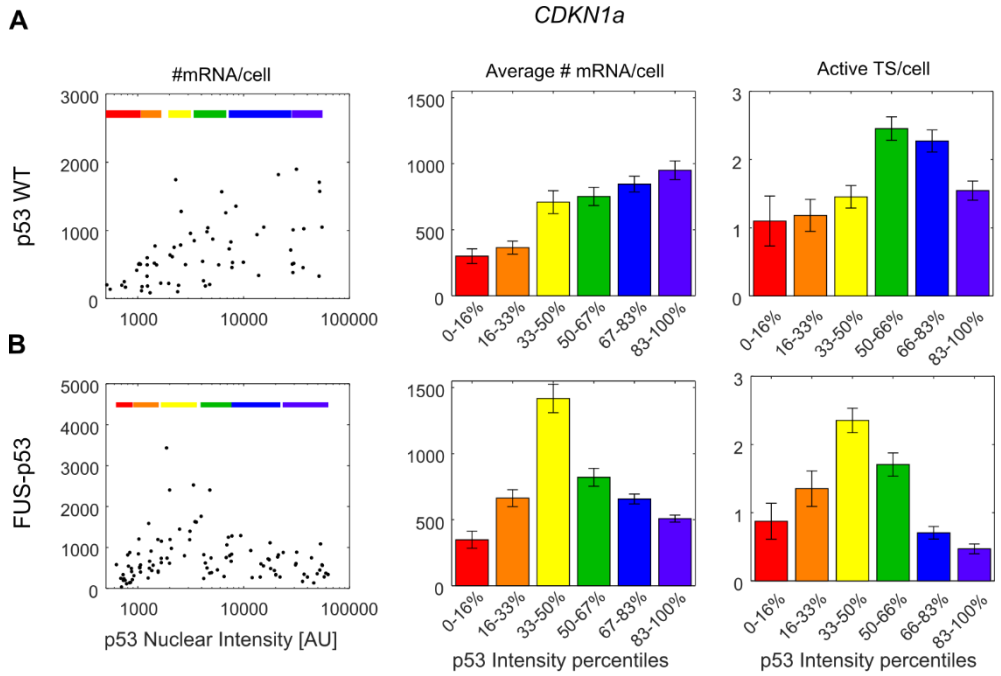


Figure 7.26 Dependence of CDKN1A RNA levels on p53 WT and FUS-p53 nuclear levels*. (A) CDKN1A shows a progressive increase in expression as a function of p53 WT nuclear levels (left: single-cell mRNA counts), center: data binned in sextiles for p53 nuclear levels, right number of active TS per cell, binned in sextiles). (B) FUS-p53 instead results in a stronger increase of CDKN1A at low expression levels, followed by repression at higher FUS-p53 levels. The maximum of gene expression is found between the 3rd and 4th sextiles of FUS-p53 nuclear levels and roughly matches with the value at which FUS-p53 forms visible condensates.*

The amount of nascent RNA in response to FUS-p53 displayed instead a more peculiar profile: at expression levels lower than a critical concentration, *CDKN1A* and *MDM2* mRNA counts increased with FUS-p53 more rapidly than with p53 WT (Figure 7.26 and Figure 7.27). This rapid increase, was also observed for the number of active TS observed per cell, indicating that in this regime of low FUS-p53 expression, the targeting of the activator to its regulated genes is enhanced. However, over a critical concentration of FUS-p53, we observed a decrease in mRNA counts and active TS per cell (Figure 7.26 and Figure 7.27). This observation was faithfully reproduced in another cellular model (MCF7 cells, data not shown), and only affected p53 target genes, as we verified by repeating the smFISH experiment on the housekeeping gene GAPDH (Figure 7.28).

Of note, the critical concentration at which gene expression starts to be inhibited with increasing levels of FUS-p53 roughly matches with the one at which visible FUS-p53 condensates appear.

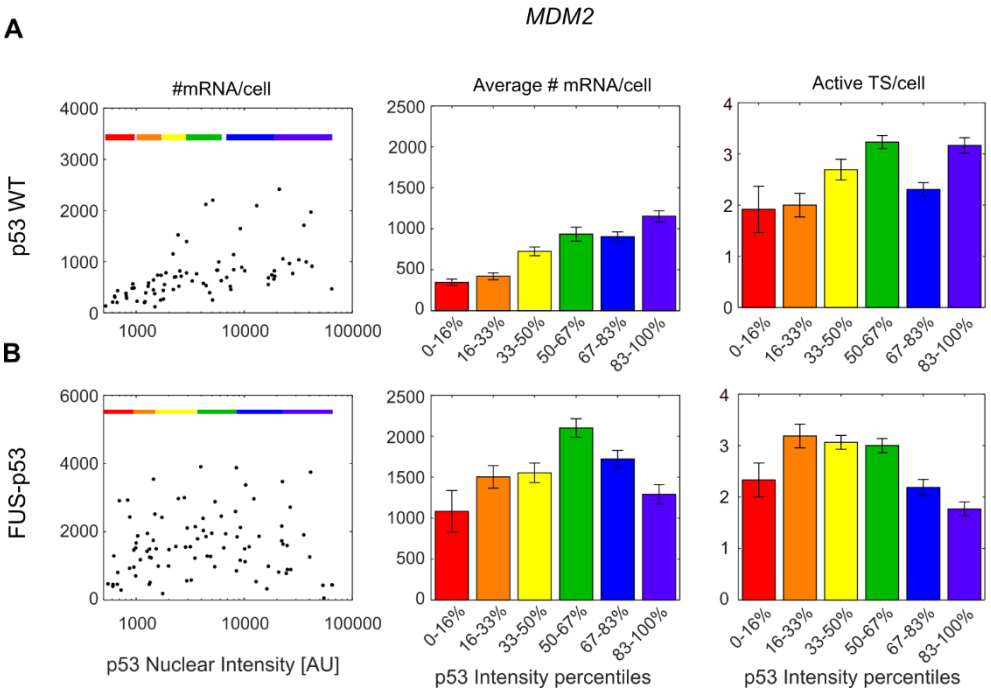


Figure 7.27 Dependence of MDM2 RNA levels on p53 WT and FUS-p53 nuclear levels. *(A) MDM2 shows a progressive increase in expression as a function of p53 WT nuclear levels (left: single-cell mRNA counts), center: data binned in sextiles for p53 nuclear levels, right number of active TS per cell, binned in sextiles). (B) FUS-p53 instead results in a stronger increase of CDKN1A at low expression levels, followed by repression at higher FUS-p53 levels. The maximum of target gene expression is found between the 3rd and 4th sextiles of FUS-p53 nuclear levels and roughly matches with the value at which FUS-p53 forms visible condensates.*

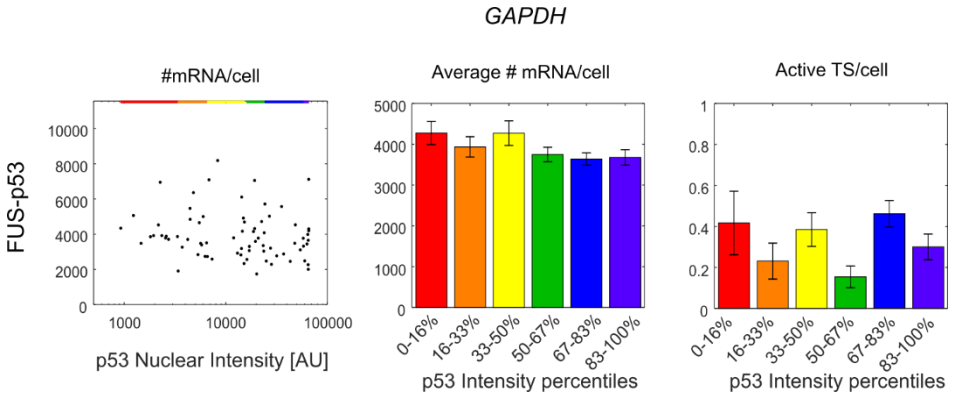


Figure 7.28 Gene expression of GAPDH related to p53-FUS nuclear level. *GAPDH expression does not depend on FUS-p53 levels, with comparable GAPDH counts in cells with different abundancy of FUS-p53.*

These data suggested us that IDR-mediated interactions need to be finely regulated in order to promote the targeting of the TF on its binding sites, as an excess of strong IDR-containing p53 could divert the search mechanism and inhibit transcriptional activation. To test this hypothesis, we therefore re-analyzed our

SMT/SIM acquisitions on FUS-p53 by splitting the dataset in two subgroups: cells expressing FUS-p53 at low levels, with no visible condensation, and cells expressing FUS-p53 at high levels and displaying two or more condensates (Figure 7.29). This analysis highlighted that while FUS-p53 is targeted to the correct chromatin environment when expressed at low levels, with bound molecules enriched at CDs, high expression of FUS-p53 derails this targeting, resulting in the localization of both diffusing and bound p53 molecules in chromatin poor regions.

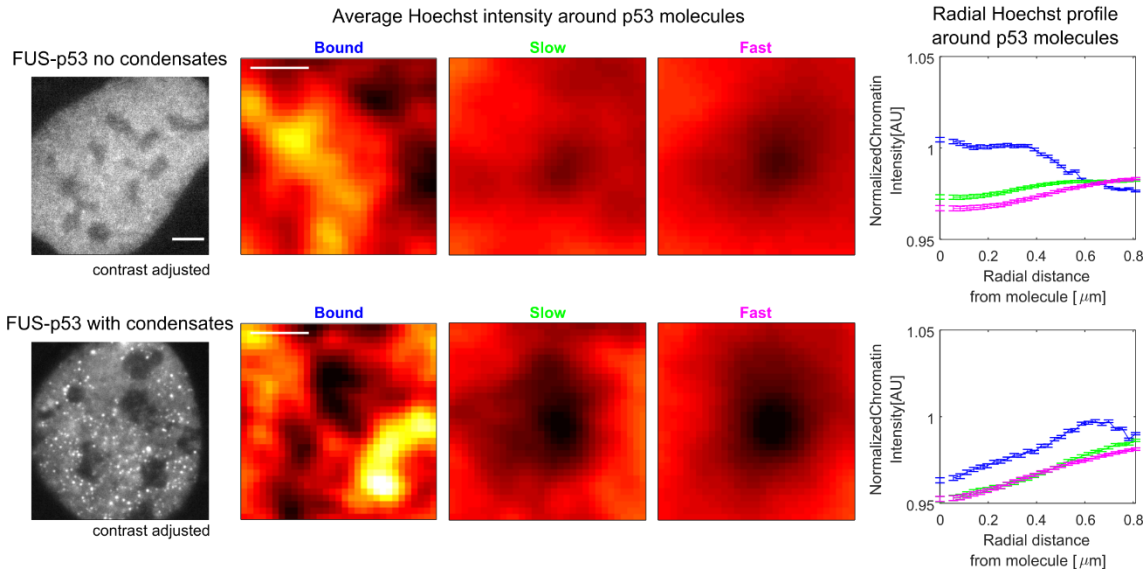


Figure 7.29 Effect of FUS-p53 levels and clustering on p53 target search. *Stratifying FUS-p53 SMT/mSIM data in cells without and with visible clusters, highlights that CDs are reached more frequently when FUS-p53 is expressed at low levels, with no visible clusters. (n_{cells} = 22,13 for "no-condensate" and "with condensates" datasets respectively).*

In sum, our data demonstrate that chromatin density controls the target search of p53, through interactions with the protein IDRs, and these interactions need to be finely regulated since both IDR deletions or potentiation can have detrimental effects on p53 targeting and on transcription of its target genes.

8 p53 search mechanism relative to nuclear bodies

8.1 p53 kinetics is influenced by regulatory PML bodies

The results presented in this section are preliminary and represent the base for future research lines.

As illustrated in the introduction, p53 interacts with a plethora of molecules, some of which capable of assembling distinct nuclear compartments. A fascinating example is provided by PML, a protein expressed under several isoforms which aggregate and form nuclear sub-structures called PML nuclear bodies (NBs). PML bodies have been reported to influence many cellular processes as senescence or cell death (Bernardi & Pandolfi, 2007). Even if the molecular mechanisms underlying such a variety of functions is poorly understood, PML NBs are suggested to provide sort of a melting-pot in the nucleus where many different proteins (including p53) flow into and acquire PTMs (Lallemand-Breitenbach & Thé, 2010).

A body of evidence indicates that PML modulates several p53 PTMs involved in its activation, including phosphorylation and acetylation (Alsheich-Bartok *et al*, 2008; Pearson *et al*, 2000; Rokudai *et al*, 2013; Hoffman *et al*, 2002), potentially through recruitment at PML bodies.

However, most of the studies on p53 recruitment to PML bodies have been performed in overexpression conditions that lead to μm-sized bodies (larger than what observed in physiological settings), and in these conditions p53 appears to be constitutively recruited to PML bodies, hampering the dissection of how the p53/PML interaction is coordinated at the spatiotemporal level (Figure 8.1).

We therefore decided to use our SMT/mSIM approach and to optimize transfection of PML-GFP to monitor the recruitment of HaloTag-p53 to PML-NBs (**Movie 4**).

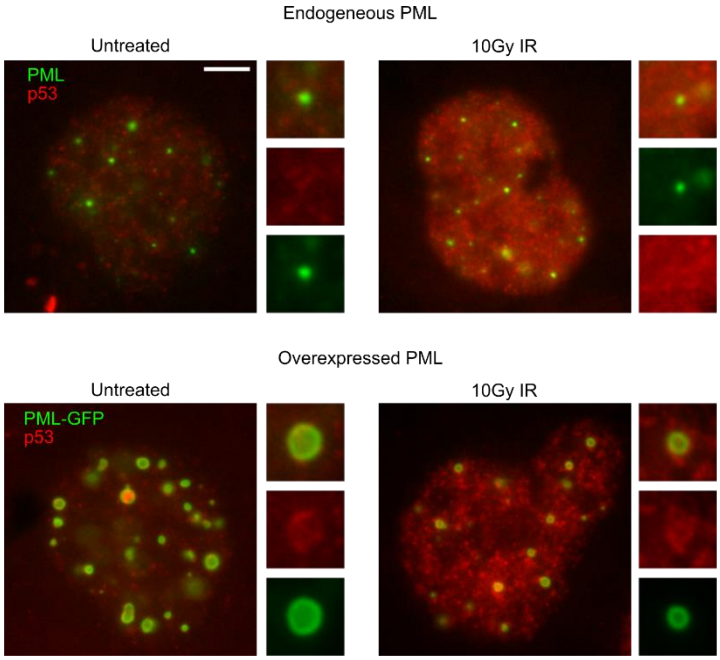


Figure 8.1 **PML overexpression leads to larger PML bodies and constitutive recruitment of p53.** (Top) Immunofluorescence on endogenous PML identifies PML bodies as sub-micron sized organelles. In these conditions, analysis of p53 localization at PML bodies is hampered by the fact that p53 molecules are scattered throughout the nucleoplasm. (Bottom) Overexpression of PML-GFP leads to larger PML bodies (see below) and constitutive recruitment of p53.

To this end, we transiently transfected PML3-GFP, a PML isoform reported to interact with p53 in the DIvA parental cell line U2OS (Fogal, 2000; Hoffman *et al*, 2002). To minimize the side effects of overexpression that may alter the PML body size and p53 recruitment, we used a tiny amount of PML3-GFP plasmid in combination with an empty vector (Material and Methods, **9.3 Plasmids and transient transfections**), to reduce the number of PML3-GFP copies incorporated by cells, while ensuring a good transfection yield.

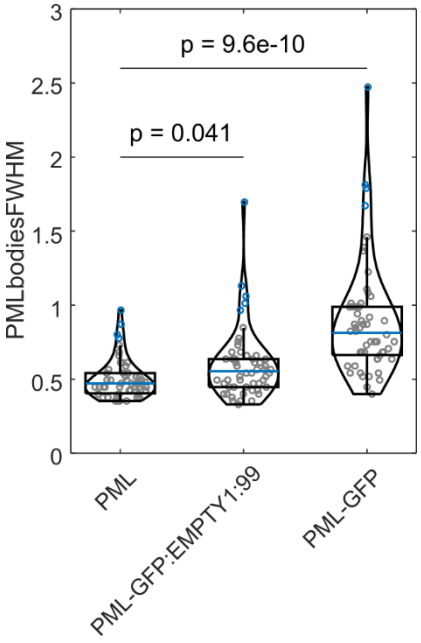


Figure 8.2 Optimization of PML-GFP transfection. *Transfection of PML-GFP mixed with an excess of empty vector leads to PML bodies that are just 20% larger than endogenous ones. Vice-versa, transfection with the PML-GFP vector leads to bodies that are approximately twice as large (n = 51 for each condition, statistical test Kruskal-Wallis).*

In these optimized conditions, we obtained PML bodies with a classical spherical shape, ~20% larger than endogenous ones as highlighted by comparing the diameter of PML3-GFP with PML endogenous bodies stained by immunofluorescence (Figure 8.2). We considered this to be a reasonable compromise for preliminary experiments. Indeed, when p53-HT single molecules were analyzed in untreated conditions for their distribution in the nucleus relative to PML3-GFP bodies (Figure 8.3), we did not find any significant co-localization, highlighting that the modest PML-bodies size increase obtained in our conditions seems not to cause, by itself, p53 recruitment. Interestingly, however, after DNA damage we noticed a significant enrichment of p53 bound molecules, consistent with the hypothesis that PML bodies may host p53 to modulate p53 modifications/activation (Figure 8.3).

Importantly, the fraction of p53 molecules found at PML represents a minority of the whole p53 bound population, which may be hardly detected through standard imaging technologies (e.g. immunofluorescence), where all molecules are labelled.

We speculate that this low but significant percentage of molecules may indicate a distinct p53 subpopulation, potentially representative of those p53 molecules needing PTMs refueling.

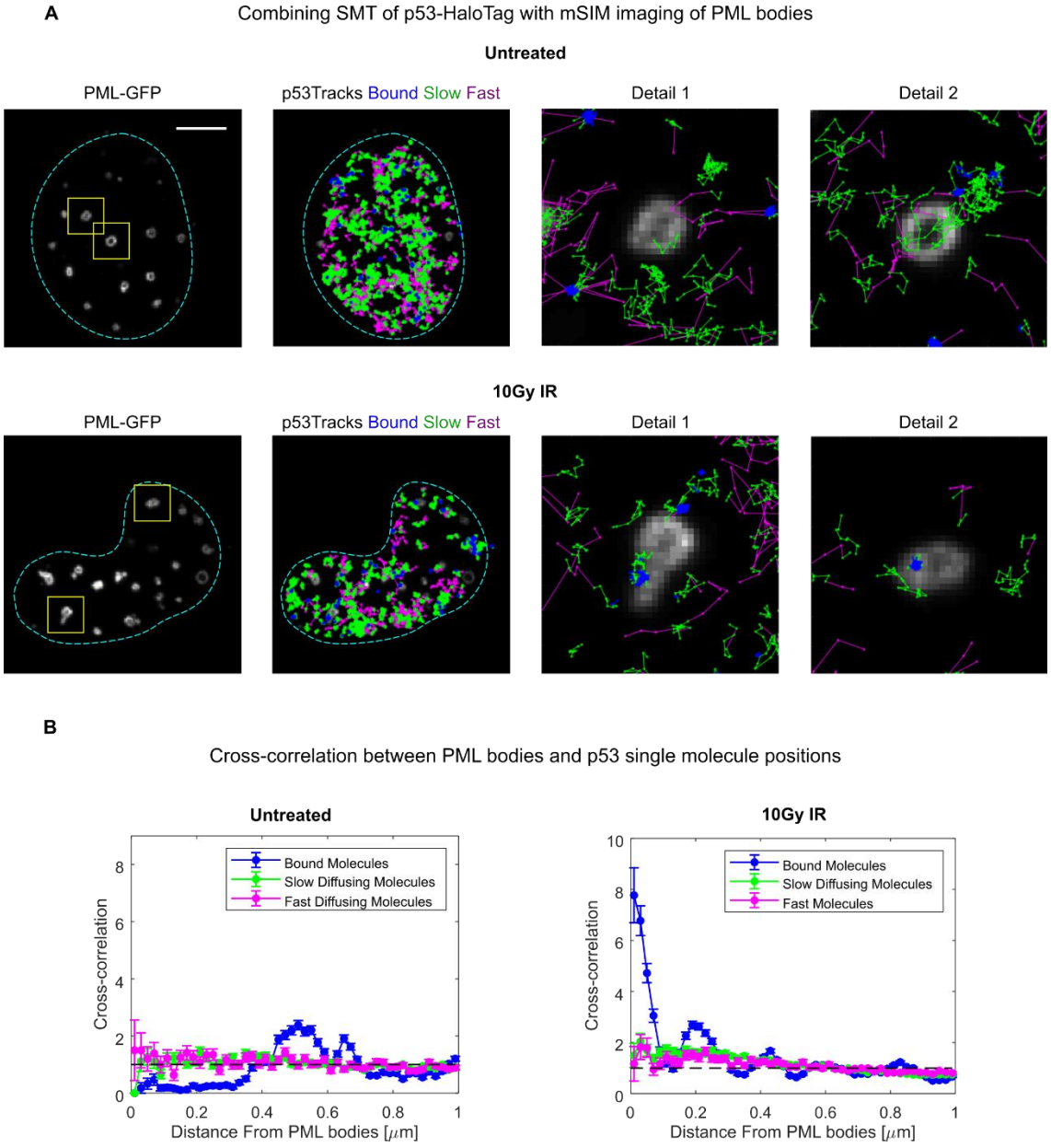


Figure 8.3 Immobilized p53 molecules are selectively targeted to PML bodies upon irradiation. *(A) We combined mSIM imaging of PML bodies with single molecule tracking in DIvA cells. The details (right) show example zooms on nuclear bodies. p53 molecules appear to be selectively targeted at PML bodies only after activation by irradiation. (B) To quantify this observation, we computed the cross correlation between PML body positions and the positions of p53 bound, slow diffusing and fast diffusing molecules, finding a significant cross-correlation only between bound molecules and PML bodies upon the induction of DNA damage (n_{cells} = 15, 17 for untreated and 10Gy IR respectively).*

To further evaluate the role of PML in regulating p53, we silenced PML in our DIvA cells by a short hairpin RNA. After silencing, cells displayed a considerably lower amount of the PML protein (Figure 8.4). Confirming a cross-talk between PML and

p53, we noticed a reduced level of p53 both at basal conditions and after DNA damage, in PML-silenced cells.

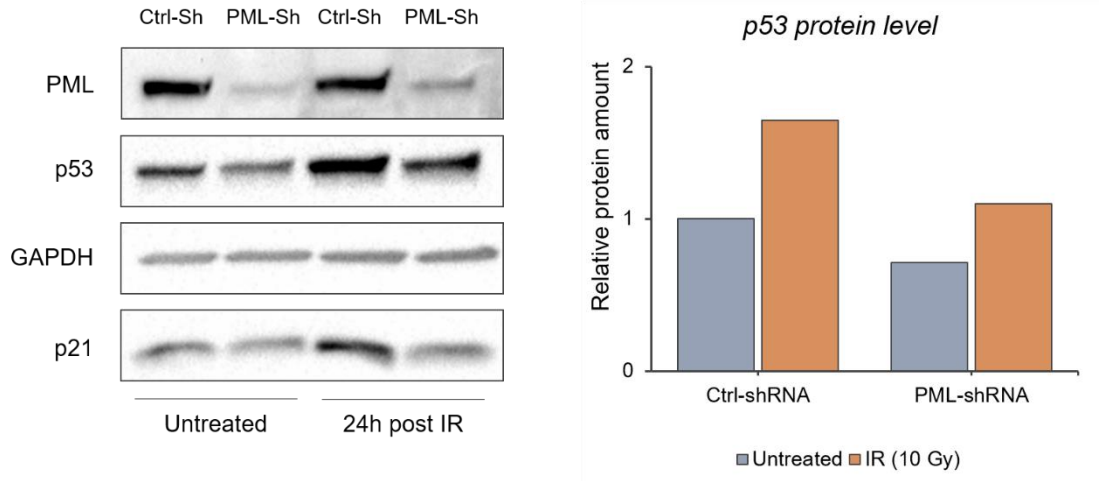


Figure 8.4 Western blot analysis on cells silenced with a shRNA against PML. *PML silencing reduces the amount of p53 (left and right panel). Similarly, p21 (transcriptional target of p53) does not significantly accumulate in PML-silenced cells (left panel).*

Consistently, qPCR analysis on p53 target genes indicate that in PML-silenced conditions, p53 poorly induces its targets (Figure 8.5).

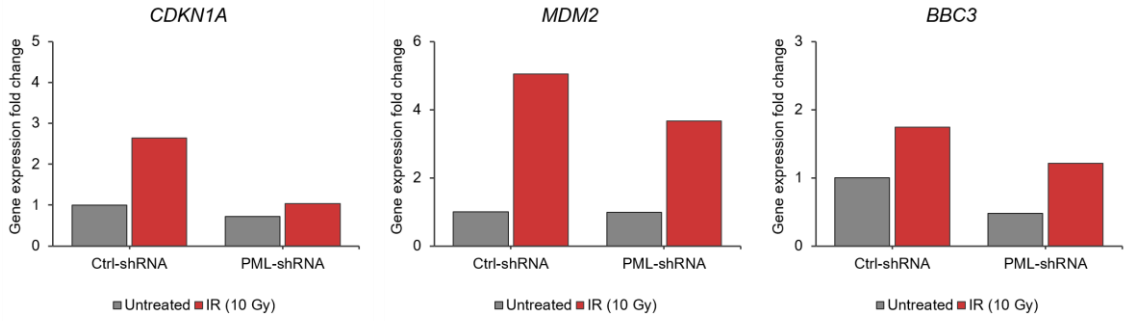


Figure 8.5 qPCR analysis on DIvA cells silenced for PML. *Canonical p53 target genes are poorly expressed in silenced cells, confirming the role of PML as a positive regulator of p53 activity.*

Together, these preliminary observations indicate that in silenced cells, the efficiency by which p53 finds its modifiers (usually encountered at PML NBs) may be reduced. By acquiring less activating PTMs, p53 is kept at reduced protein levels and may interact with its cofactors at a lower extent. We aim to extend our preliminary results in the near future, by monitoring by SMT/mSIM live imaging the recruitment of p53 at PML bodies over time, and by performing additional experiments on the modifications status of p53 and mutants preventing these modifications.

DISCUSSION

How mammalian TFs efficiently find their genes throughout the genome and how this process is modulated by the compartmentalized cell nucleus, is poorly understood. The nucleus appears organized in compartments that regulate relevant molecular processes as transcription, DNA repair and replication. It is not clear whether this spatial organization also facilitates the recruitment of TFs to their promoters/enhancers.

Of particular interest for the recruitment of molecular factors to their target of action is the nuclear organization as proposed in the Cremer's model, which postulates a different permeability of the nuclear space depending on chromatin density that may strongly shape the search process. Chromatin not uniformly distributed in the nucleus generates a compartmentalized environment, with interchromatin poor regions (or interchromatin channels) ensuring a rapid diffusion of molecules towards their target sites in the denser chromatin domains.

In this study we show that the tumor suppressor p53 -required for an effective response against DNA damage- reaches its target sites within chromatin through a search process that appears modulated by the nuclear environment, generating distinct exploration strategies depending on chromatin compartmentalization.

p53 displays different subpopulations of diffusing molecules that can be grouped into two main categories (slow and fast). Slow molecules show higher likelihood to reach the bound state, suggesting that the two populations could play different roles in the targeting process, with the slow molecules promoting binding. We thus wondered whether slow molecules visit regions that may facilitate their recruitment to binding sites. The spatial distribution of diffusing molecules relative to bound ones, reveals that slow and fast populations explore separated regions, with slow molecules visiting sites densely occupied by bound molecules, and the fast subpopulation more randomly distributed throughout the nucleoplasm. This observation led us to investigate whether p53 subpopulations segregate in distinct local environments, depending on the nuclear organization. We focused on the distribution of chromatin, which provides compartments of different permeability as described by Cremer's model. Using live imaging SMT/mSIM we investigated the position of molecules depending on chromatin density. The results indicate that fast molecules diffuse preferentially within chromatin-poor spaces, compared to slow molecules (found at compartments of intermediate density) and bound molecules (positioned at higher density). This confirms a molecule partitioning based on chromatin distribution.

To obtain a complete picture of the p53 search process, we investigated whether the p53 slow-to-fast transition (and vice-versa), underlies a switch in the exploration strategy (e.g. compact vs non-compact) by diffusive molecules. Indeed, the efficiency of the targeting process is regulated by the capability of TFs to densely sample nuclear regions enriched by its targets. Our data indicate that p53 adopts an intermittent search process, switching from non-compact exploration at long distances to compact exploration at short scales, as highlighted by a peak in p53 anisotropic diffusion at ~100nm, which increases following p53 activation by DNA damage. This intermittent behavior is compatible with guided exploration, a search mechanism emerging as potential common strategy to accelerate the search of nuclear proteins: for example, the chromatin regulator CTCF and the Polycomb subunit CBX2 display similar anisotropic diffusion as we measured for p53 (Hansen *et al*, 2020; Kent *et al*, 2020).

Together, the data can be interpreted under the Cremer's model (Cremer *et al*, 2020): fast molecules traverse the interchromatin channels by free diffusion which is an efficient strategy to rapidly escape from regions devoid of genes (indeed, our DNA FISH and Hi-C data suggest genes to be located in regions with relatively high DNA density). At the interface with chromatin domains (regions highlighted by intermediate density), p53 slows down and switches to compact exploration, scanning nuclear sub-compartments (100-nm in size), and finding its targets at denser positions.

What does induce this switch? Moving from the observation that eukaryotic TFs often undergoes compact exploration through interactions mediated by their IDRs (Hansen *et al*, 2020; Garcia *et al*, 2021b), we performed an analysis on several p53 mutants lacking the disordered p53 C-terminus, N-terminus or both IDRs. All mutants displayed a reduced degree of anisotropy of their slow components together with a reduced co-localization with bound molecules, highlighting that IDRs lacking mutants are less capable of compact explorations and visit less frequently regions potentially occupied by p53 targets. Consistently, bound molecules of p53 IDR mutants occupy chromatin regions at a lower density on average compared p53 WT. This suggests an impaired search process for the tested mutants, leading to localize with less efficiency targets in dense chromatin regions. Indeed, genes regulated by p53 in our cells appear to occupy dense chromatin positions genome-wide, compared to open chromatin regions, as revealed by our analysis on a Hi-C dataset.

A final gene expression analysis by single molecule RNA-FISH, shows that p53 WT induces a higher burst frequency than IDR mutants, consistently with the current interpretation that the burst frequency is linked to the binding frequency of TFs.

In recent years, the role of IDRs in transcription have been mostly linked to the capability of disordered proteins to form condensates and -according to some evidence- enhance transcription (Wei *et al*, 2020). Emerging data however indicate that IDRs may increase transcription independently of the formation of condensates (Trojanowski *et al*, 2021). Our data extend further this observation. Indeed, when fusing p53 to a strong IDR (derived from FUS), at low level of FUS-p53, the strong IDR potentiates guided exploration and the transcription of target genes. However, higher levels lead to the formation of FUS-p53 condensates and to remarkably lower transcription, when compared to similar nuclear level of p53 WT. We wondered if the IDR-mediated interactions leading to FUS-p53 condensates may sequester p53 from the right target sites. Consistent with this hypothesis, FUS-p53 bound molecules in cells with condensates do not localize any longer on dense chromatin regions, suggesting that in these conditions FUS-p53 molecules are not targeted at right genomic positions. These results suggest a model in which IDR-mediated interactions in the nuclear environment need to be finely balanced to warrant efficient targeting. According to this model, p53 mutants deficient in IDR-mediated interactions explore the nuclear environment in a less compact fashion and are unlikely to find their binding sites, but strong IDR interactions can result in an excessive compact search, that sequesters the TF away from its REs. Interestingly, a similar trade-off has been described for facilitated diffusion, the most studied accelerated search strategy in prokaryotes. Here, optimal search is achieved when the factor spends 50% of the search time by sliding on DNA, and shorter or longer sliding times result in inefficient targeting (Mirny *et al*, 2009). To further validate the model, we are currently generating cell lines stably expressing either WT -p53 or FUS-p53 under an inducible promoter, to be tested by ChIP-seq analysis in presence and absence of condensates.

This new model raises some novel and important questions that we would like to address in the next future.

First, it will be important to identify the mechanism regulating p53 guided exploration at the interface of chromatin domains. One compelling hypothesis is that this behavior is the macroscopic manifestation of facilitated diffusion as p53 is one of the few eukaryotic TFs that have been shown to be capable of sliding on DNA in vitro, and the sliding is mediated by its disordered C-terminal domain. Indeed, even if DNA sliding is supposed to occur at very small length scale in vivo (with the TF dissociating from DNA after scanning few hundreds of nucleotides), hampering the direct visualization of this process, facilitated diffusion combines discontinuous but multiple steps of sliding, on separate DNA stretches. If DNA fibers are abundant in a sub-compartment, a TF after dissociating from one site has good chances to encounter a

new DNA sequence (e.g. by 3D diffusion or hopping) nearby in space, giving rise to sequential steps of compact exploration in a confined region. This scenario depends on the amount of DNA, since isolated chromatin fibers might be unlikely to intercept the TF. We speculate that this may be in line with the observation that p53 regulated genes lie within chromatin dense regions, and with the capability by p53 to readily associate to nucleosomal DNA (being a pioneer), without the necessity to encounter open sequences.

A further level of complexity is given by the lack of information about whether TFs can slide on nucleosomal DNA. Indeed, two scenarios appear reasonable: (i) TFs slide on nucleosome-poor DNA dissociating when encountering nucleosomes, or (ii) TFs scan also nucleosomal DNA by displacing DNA chains form the histone octamer surface, possibly facilitated by coactivators or architectural proteins (e.g. p300, HMGB1). On this, experiments that perturb the interaction between p53 and its coactivators (interacting themselves with p53 IDRs), may provide valuable information.

Recent evidence highlights that nascent RNAs preferentially localize at the interface of chromatin domains. In this context, RNAs might represent potential candidates that could mediate p53 compact exploration. Indeed, recent studies indicate that RNAs may mediate the targeting of TFs and p53 itself has been found to interact with RNAs both at enhancer regions and promoters (Melo *et al*, 2013; Link *et al*, 2013; Schmitt *et al*, 2016; Statello *et al*, 2021). Experiments to characterize the p53 search after transcriptional inhibition to reduce the amount of RNAs at chromatin domains are however challenging, since drugs interfering with transcription have been shown to perturb chromatin organization at the micron scale.

Interestingly, p53 guided diffusion appears modulated after DNA damage, displaying an increased compact exploration. Given the strong influence of the local environment on the search process, a key question is whether the higher compact exploration reflects acquired modifications on p53 (e.g. PTMs) following DNA damage or on the cell nucleus itself. Experiments measuring the mobility of proteins different from p53 (e.g. unrelated TFs, HaloTag peptide) after DNA damage, or perturbing p53 modifications (e.g. by using p53 mutants) will clarify the role played by the nucleus and p53 PTMs.

A key point is how guided exploration may contribute to p53 selectivity, regulating the activation of a subset of targets. The observation that -on average- targets not regulated occupy positions at lower chromatin density, on one hand is surprising since Cremer's model inherently predicts that dense regions are more difficult to be reached by TFs. On the other side, however, it can indicate that p53 target genes

should locate in the 'correct' chromatin environment to be effectively regulated. A possibility is that regions at low density are poorly attractive for p53: if the persistence of compact exploration in a sub-compartment (and with it, the likelihood to encounter a gene), is influenced by the availability of chromatin, regions at low density possibly do not retain p53 for a sufficient amount of time to locate its targets.

Alternatively, other molecular barriers (differently from chromatin density and nucleosomes) may prevent p53 to diffuse/associate with non-regulated targets.

Another important question, is whether the search mechanism that we describe for p53 applies also to other TFs. To a limited degree, data from other nuclear proteins seem to point against a common search mechanism. Factors as the oncogene cMyc and the NF-kB subunit p65 (in this work) apparently display non-compact search mechanisms, while P-TEFb diffuses through compact exploration at all spatial scales (Izeddin *et al*, 2014b). Other chromatin regulators display 'guided exploration' that depends on their capability to form clusters. TFs are therefore likely to explore the nuclear environment differently, by transiently interacting with different nuclear compartments and molecular players.

Finally, the dynamic activity of TFs -that can be switched on and off by different PTMs- requires the TF to interact with different nuclear structures depending on the 'stage' of its activity (Misteli, 2007; Galganski *et al*, 2017). To face this problem, we assessed whether p53 visits PML nuclear bodies, specialized compartments reported to modulate TFs functions and PTMs. Our preliminary data point in favor of this dynamic view. While at basal conditions we do not observe a significant interaction, after DNA damage p53 is recruited at PML bodies, possibly to accumulate activating modifications. To further explore this observation, we will investigate the molecular bases of this recruitment, for example assessing our IDRs mutants, since p53 IDRs are crucial domains for the accumulation of PTMs.

Taken together our data point out that TFs -in their spatiotemporal activity- need to 'see' and adapt to different nuclear structures, while disregarding others. We envision that by increasing the throughput of the SMT/mSIM approach to probing multiple nuclear sub-compartments, it will be possible to generate maps of the cell nucleus that act like a 'cartography', and study how different nuclear environments are sensed by different TFs.

MATERIALS AND METHODS

9 Lab procedures

9.1 Cell culture

DIvA cells (originated from the human osteosarcoma U2OS) were a kind gift from the Gaëlle Legube lab (Aymard *et al*, 2014). The U2OS-derived line with endogenous expression of CTCF-HaloTag (used by us for control experiments), was a kind gift from the Tjian and Darzacq group (Hansen *et al*, 2020).

All cell lines, including those that we derived by CRIPR/Cas9 from DIvA, were cultured in DMEM (Thermo-Fisher, cat. 31053044) supplemented with 10% heat-inactivated Qualified Fetal Bovine Serum (Gibco), 100 units/mL penicillin-streptomycin (Thermo-Fisher, cat. 15140122) and 2 mM L-Glutamine (Gibco, Thermo-Fisher, cat. 25030081). For DIvA and DIvA-derived cell lines, DMEM was further supplemented with 1µg/mL puromycin (Thermo-Fisher, cat. A1113803) to maintain the stable expression of the AsiSI-RE enzyme.

Cells were grown at 37°C in a humidified incubator with 5% CO_2. For routine maintenance, cells were passaged every 2-3 days by trypsinization, plated on 75-cm flasks, and tested once a week for mycoplasma contamination (PCR-based assay).

9.2 Generation of p53-HaloTag and p53 KO cell lines

CRISPR/Cas9 gene editing was applied to create p53-HaloTag knock-in (KI) and p53 knock-out (KO) cell lines.

p53 KI DIvA cells were generated using four plasmids (kind gifts from Edouard Bertrand lab, Figures 9.1 and 9.2) encoding the Cas9-D10A nickase (Cong *et al*, 2013), two guide RNAs (sgRNA1 and sgRNA2) and a repair vector (sequence in supplementary material) carrying HaloTag and a neomycin resistance gene, flanked by two p53 homology arms (~800 bp).

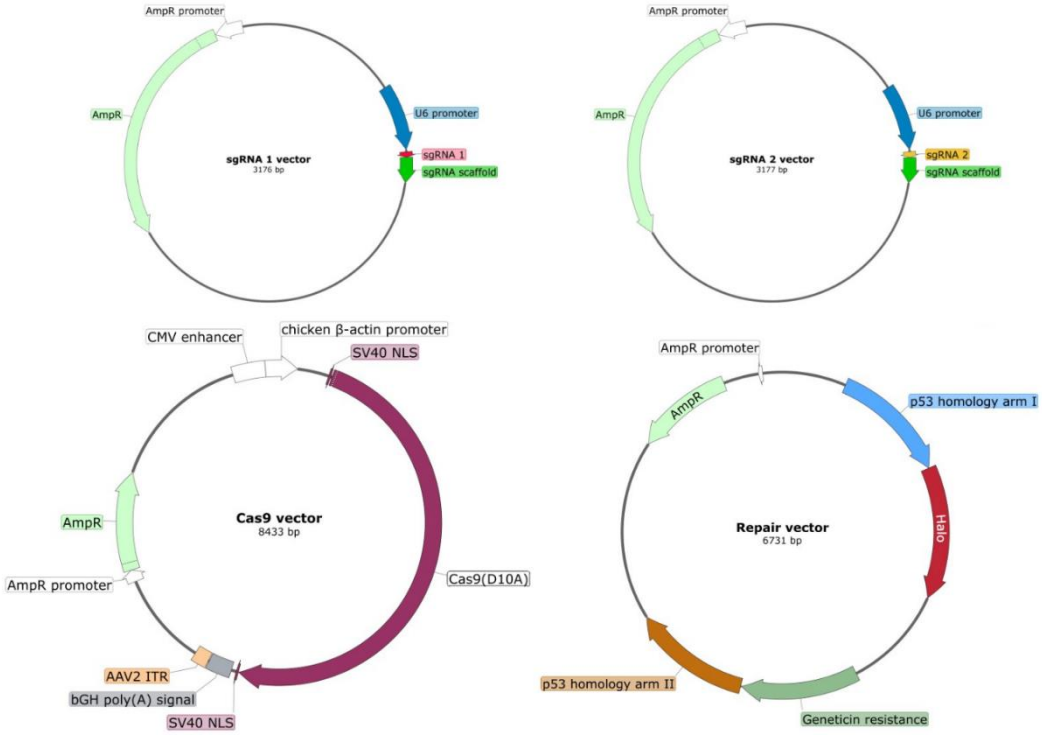


Figure 9.1 Maps of vectors used to generate p53-HaloTag cells by CRISPR/Cas9. *To induce p53-KI gene editing, we transfected our cells with four plasmids. Three of them (expression vectors) encode for single guide RNAs and the Cas9 nickase. Finally, a repair vector with homology arms for p53 provides the template to insert HaloTag in the Tp53locus. Image created with SnapGene.*

The sequences targeted by sgRNA1 and sgRNA2 on the p53 genomic locus (TP53) were 5' GATGACATCACATGAGTGAG 3' and 5' CAGCCACCTGAAGTCCAAAA 3', respectively, located at the TP53 3' end. DIvA cells seeded on a 6-well plate were co-transfected with the four plasmids (625 ng of each vector per well) using Lipofectamine 3000 (Thermo-Fisher, cat. L3000008) and following the manufacturer's protocol. After transfection, cells were left to recover few days before two rounds of antibiotic selection (mild, 4 days and long, ~15 days) with geneticin (G418 Sulfate, Thermo-Fisher, cat. 10131027, used at 800 µg/mL). Afterwards, cells were screened by western-blots and the promising populations (defined by a strong p53-HaloTag band compared to p53 endogenous, as indicated in **Figure 6.3**, Results section) were plated at single cell per well (96-well plate) to isolate and expand individual clones. Clones were functionally tested for nuclear localization of p53-HaloTag, for its protein abundancy and the capability to induce canonical p53 target genes (both at the RNA and protein level).

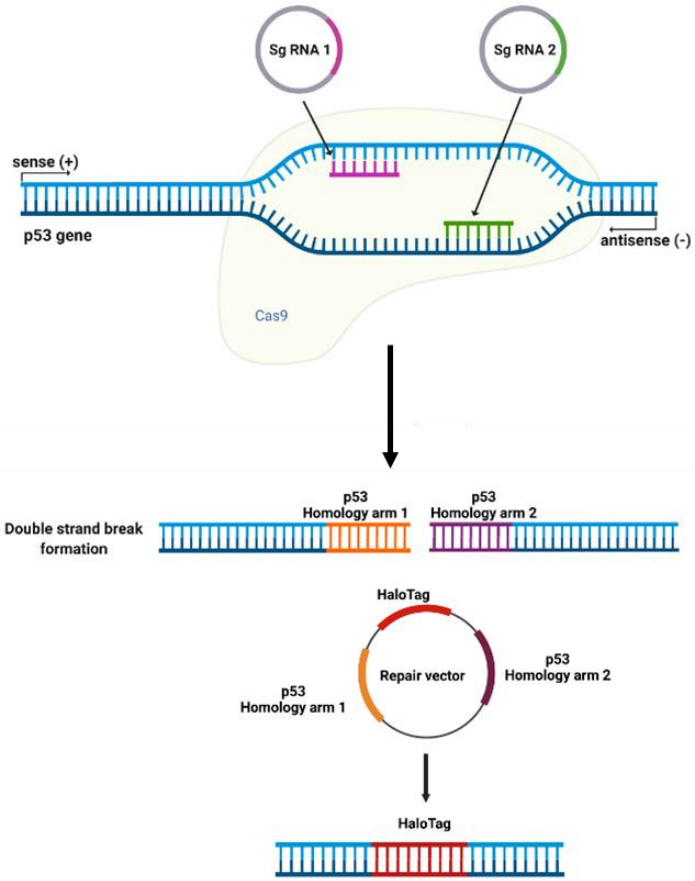


Figure 9.2 Scheme of CRISPR/Cas9 genome editing to induce the p53-HaloTag knock-in. *(A) Single guides RNAs target Cas9 on TP53 locus, generating a double strand break (DSB). DNA is repaired by homologous recombination, using the repair vector as a template, and inserting HaloTag. Image created with BioRender.com.*

To generate p53 KO DIvA cells, a commercial CRISPR/Cas9-based gene editing kit was employed (Santa Cruz Biotechnology, cat. sc-416469-NIC), including constructs expressing Cas9-D10A nickase, two sgRNAs and a GFP marker transiently expressed for selection. DIvA cells were transfected with UltraCruz Transfection Reagent (sc-395739) following the manufacturer's instructions and sorted for GFP the next day. Single clones were isolated and expanded from the sorted population. Successful knock-out was confirmed by absence of p53 (no western-blot bands) and inactivation of p53 target genes (both at the RNA and protein level).

9.3 Plasmids and transient transfections

Plasmids expressing p53-HaloTag mutants were synthesized and sequenced by the Genewiz Company from a pFN22A vector (Promega) containing the HaloTag-p53 WT (Figure 9.3). In particular: deletions of p53 N-terminus (aa 1-94), C-terminus

(aa 363-393), or both domains (aa 1-94 and 363-393); insertion of the FUS N-terminal domain (aa 1-214).

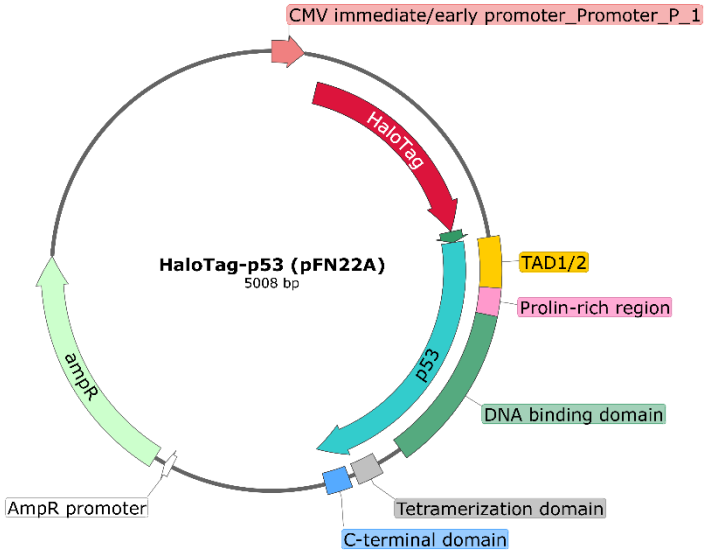


Figure 9.3 HaloTag-p53 WT vector used to generate p53 IDR mutants. *The search properties of WT-p53 were compared to deletion mutants (TAD and CTD) and to the insertion mutant FUS-p53, by transient transfection. Image created with SnapGene*

Several vectors expressing HaloTag-conjugated proteins were used in control experiments: H2B (pFC15A backbone, Promega; (Mazza *et al*, 2012a)), the Nf-κb subunit p65 (pFN22K, Promega) and HaloTag itself (pHTN, Promega).

Plasmids were transiently transfected in p53-null cells (DIvA p53 KO). Briefly, DIvA cells were seeded at ~30% of confluence, and 24h later (~60% confluence) were transfected with Lipofectamine 3000 (Thermo-Fisher, cat. L3000008), according to manufacturer's protocol.

A plasmid expressing PML3-GFP (Weidtkamp-Peters *et al*, 2008) and a lentiviral vector encoding for PML short-hairpin (sh) RNA, were a kind gift from Rosa Bernardi's lab. PML3-GFP was transfected in p53-HaloTag KI cells with Lipofectamine 3000 (Thermo-Fisher, cat. L3000008). To obtain a low and more physiological expression of PML-GFP, cells were transfected with a small quantity of PML-GFP mixed with a vector expressing no proteins (pCDNA4/TO, Addgene) at a ratio of 1:99, with a total DNA amount as suggested by the Lipofectamine protocol. Silencing of PML was performed by lentiviral infecting the shRNA against PML, and using a scrambled shRNA (Addgene) in control cells. Cells were kept under antibiotic selection for 96h and analyzed by western blot and qPCR analysis.

9.4 Immunofluorescence

DIvA cells were grown on coverslips, treated (with IR or 4OHT, see below **9.5 Treatments to activate p53**) or left untreated, washed once in PBS and fixed in 4% paraformaldehyde for 10 min at room temperature (RT). After fixation, cells were quenched with 50 mM NH_4Cl, permeabilized with 0.1% Triton X-100 (Sigma Aldrich) and blocked for 45 min at RT in PBS supplemented with 20% FBS, 0.05% Tween 20 (Sigma Aldrich, cat. P2287) and 5% BSA (Sigma Aldrich, cat. A7906). Next, cells were incubated for 1h at RT with primary antibodies (diluted in 2% BSA) to detect the AsiSI-ER-HA tagged enzyme (mouse monoclonal HA antibody, Abcam, cat. ab130275, 1:500 dilution) and the DNA damage marker γ-H2AX (rabbit polyclonal γ-H2AX antibody, Abcam cat. ab11174, 1:200 dilution). After three PBS washes, cells were incubated in the dark with secondary antibodies diluted 1:1000 in 2% BSA (goat anti-mouse IgG Alexa Fluor 488, Themo-Fisher, A11029, Themo-Fisher, cat. A21070 goat anti-rabbit IgG Alexa Fluor 633, Themo-Fisher, cat. A21070), for 1h at RT. Cells were washed three times in PBS and incubated in the dark with a DNA staining solution (Hoechst 33342, Thermo-Fisher, cat. H3570, 1 µg/mL in PBS), for 10 min at RT. After two PBS washes, coverslips were mounted with Vectashield (Vector Laboratories) and sealed.

Images were collected with a custom-build widefield microscope (described below, see **10 Live-Cell Imaging** section) equipped with a sCMOS camera (physical pixel size=108.3 nm) and using a 60x oil-immersion objective (N.A.=1.49NA).

9.5 Treatments to activate p53

To stabilize p53 and induce a p53-mediated transcriptional response, DIvA cells were exposed to DNA damage either by 10 Gy γ -irradiation (using a ^{137}Cs source, Biobeam 2000) or through the nuclease enzyme AsiSI-RE, activated upon treatment with 300nM 4- Hydroxytamoxifen (4OHT, Sigma-Aldrich, cat. H7904). Alternatively, p53 was activated by treatments with 10µM Nutlin-3a (Sigma-Aldrich, SML0580), an inhibitor of the p53 negative regulator MDM2.

9.6 Western blot

Cells cultured on 10-cm dishes were washed one time in cold PBS and lysed in 300 µL RIPA buffer (25 mM Tris HCl pH 7.6, 150 mM NaCl, 1% Sodium deoxycholate, 1% Triton X-100, 2 mM EDTA dihydrate) supplemented with protease inhibitors (Sigma-Aldrich, cat. 4693124001). Samples were next incubated at 4 °C for 20 min under

constant rotation, centrifuged at 12000 rcf and the supernatants were then collected. Protein lysates were quantified by BCA assay (Thermo-Fisher, cat. 23225), loaded on 8% or 12% SDS-polyacrylamide gels and run at 100 V for ~2-3 hours. Proteins were transferred to nitrocellulose membranes in cold transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol) via run at 100 Volts for 2 h at 4°C. Next, membranes were blocked in 5% non-fat dried milk in TBS-T solution (0.1% Tween20 in TBS: 20 mM Tris base, 137 mM sodium chloride, pH 7.6) for 1h at RT while shaking. Membranes were incubated with primary antibodies, all diluted in 5% non-fat dried milk in TBS-T solution. Depending on the experimental settings, the antibodies employed were: mouse monoclonal anti-p53 DO-1 (Santa Cruz Biotechnology, cat. sc-126; 1:3000 dilution, incubated 1h at RT), rabbit monoclonal anti-p21 (Abcam, cat. ab109520; 1:1000 dilution, incubated overnight at 4°C), rabbit monoclonal anti-GAPDH (Abcam, cat. ab128915; 1:50000 dilution, incubated 1h at RT), mouse monoclonal anti-vinculin (Thermo-Fisher, cat. MA5-11690; 1:4000 dilution, incubated 1h at RT), rabbit polyclonal anti-PML (Novus Biologicals, cat. NB100-59787; 1:5000 dilution, incubated overnight at 4°C). Afterwards, membranes were washed three times in TBS-T (5 min each wash at RT, while shaking) and then incubated for 1h at RT with peroxidase-conjugated secondary antibodies (anti-mouse IgG, Cell Signaling, cat. 7076; anti-rabbit IgG, Cell Signaling, cat. 7074) diluted 1:5000 in 5% non-fat dried milk in TBS-T solution. The membranes were finally developed using ECL substrate (Bio Rad, cat. 1705061) and images were acquired with a CCD camera using ChemiDoc MP imaging system.

9.7 RNA extraction and RT-qPCR

Cells grown on 6-cm dishes were washed once in cold PBS and lysed in 750 µL of *TRIzol* reagent (Thermo-Fisher, cat. 15596018) to extract the total RNA. Lysates were purified using silica membrane columns (Machery-Nagel, NucleoSpin RNA Plus). The isolated RNA was quantified and tested for purity by *NanoDrop* fluorometer (Thermo-Fisher). For each sample, 2µg of RNA was reverse-transcribed to cDNA using the *High-Capacity cDNA Reverse Transcription Kit* (Thermo-Fisher cat. 4368814), following the manufacturer's protocol. Real-time qPCR analysis was performed to assess the expression of p53 target genes. Each reaction (20 µL final volume) was composed as following: cDNA (5 µL of 1:100 dilution), 150 nM primers, SYBR Green mix (Roche, LightCycler 480 I Master). To normalize the cDNA amount among different conditions, samples were run altogether using the constitutive gene GAPDH as internal standard.

9.8 RNA-Seq

DIvA p53 WT and DIvA p53 KO cells were plated on 6-cm dishes (12 x 10^4 cells/dish, in triplicate per condition, using different vials for each replicate) and grown in DMEM. Two days later, cells were irradiated (10Gy) to activate p53. Five hours after irradiation, irradiated and control cells were collected, and their RNA extracted as described above.

Sequencing libraries from the extracted mRNA were prepared with the *mRNA Stranded TruSeq* protocol (Illumina) and sequenced via *NovaSeq 6000 SP Reagent Kit* (Illumina, 100 cycles).

Sequencing reads were aligned to the reference genome GRCh37 (alias hg19). Differential Gene expression analysis was performed using DESeq2 (Bioconductor). The identification of differentially expressed genes (DEGs) was performed by calculating fold change (>2) and statistical significance relative to basal or p53 KO conditions, using a false discovery rate below 0.2 (Benjamini–Hochberg procedure).

9.9 RNA smFISH combined with p53-HaloTag staining

Single molecule Fluorescent In Situ Hybridization was performed to measure nascent and mature RNAs of three p53 target genes (*CDKN1A*, *BBC3,* and *MDM2*) using as control the housekeeping gene *GAPDH*. To compare the transcription dynamics of HaloTag-p53 WT and IDRs mutants, RNA expression was coupled with p53 abundancy. To this purpose, the smFISH was combined with the labeling of HaloTag. To detect single RNAs, we used *HuluFISH* RNA probes (PixelBiotech GmbH) -labeled with ATTO-647N fluorophore- for *CDKN1A*, *BBC3,* and *GAPDH* whereas for *MDM2* we employed *Design Ready Stellaris* probes (Biosearch Technologies) labeled with Quasar 670 dye.

p53-null cells (DIvA p53 KO) were grown on glass coverslips in a 6-well plate and transfected with p53-HaloTag constructs as described above. The day after transfection, cells were fixed in 4% PFA (10 min at RT), and washed with 135 mM Glycine in PBS for 10 min a RT. Next, coverslips were washed twice in PBS and permeabilized with Triton X-100 diluted in PBS at 0.1% (for cells later hybridized with *MDM2* probes) or 0.3% (for *CDKN1A*). After two PBS washes, cells were incubated with a fluorescent HaloTag ligand (2.5 µM TMR in PBS) for 1h at RT. To remove the unconjugated TMR, cells were washed with abundant PBS.

The following steps change depending on the probes (*CDKN1A*, *BBC3, GAPDH* or *MDM2*). I split the protocol in two parts.

CDKN1A, BBC3 and *GAPDH* (*HuluFISH* probes)

Cells were washed twice with HuluWash buffer (2xSSC, 2M Urea), 10 min each wash at RT. Cells were next hybridized by adding onto coverslips 0.5 µL of probes (*CDKN1A*, *BBC3* or *GAPDH*) diluted in 50 µL of HuluHyb solution (2xSSC,2M Urea, 10% dextran sulfate 5x Denhardt's solution). Hybridization was performed in a humified chamber at 30 degrees overnight. The following day, cells were washed twice in HuluWash buffer (30 min per wash, at 37 °C in the dark) and then incubated with a DNA staining solution (Hoechst 33342, 1 µg/mL in PBS) for 10 min at RT in the dark. Coverslips were mounted on glass slides with Vectashield (Vector Laboratories) and sealed.

MDM2 (*Design Ready Stellaris* probes)

Cells incubated with 1 mL of the washing buffer A, composed as following: 10% saline sodium citrate (SSC), 20% formamide solution (Thermo-Fisher, ca. AM9342), in RNase-free water (Sigma-Aldrich, ca. 95284). Next, cells were hybridized by adding onto coverslips 1 µL of 12 nM *MDM2* probe diluted in 100 µL hybridization buffer (10% dextran sulfate, 10% SSC-20× buffer, 20% formamide in RNase-free water). During hybridization cells were incubated overnight in a humidified chamber at 37°C. The following day, cells were washed twice in buffer A (30 min per wash, at 37 °C in the dark), once in 10% SCC-20×, and then incubated with a DNA staining solution (Hoechst 33342, 1 µg/mL in PBS) for 10 min at RT in the dark. Coverslips were mounted on glass slides with Vectashield (Vector Laboratories) and sealed.

Images were acquired with our custom-built widefield microscope (for details, see the **10 Live-Cell Imaging** section), using a 60x oil-immersion objective (N.A.=1.49), a sCMOS camera, and a led source for illumination. Z-stack images were collected with 0.3 µm step size.

9.10 DNA FISH

To evaluate the position of the *CDKN1A* locus relative to chromatin, we performed DNA FISH on genomic DNA.

9.10.1 Cell fixation, pre-treatment and permeabilization

DIvA cells seeded in a 24-well plate and grown on coverslip glasses, were washed twice in PBS and fixed using 4% PFA for 10 min at RT. Coverslips were next rinsed three times in PBS at RT (3 min for each wash) and permeabilized using 0.5% Triton

X-100 (diluted in PBS) for 10 min at RT. To remove RNAs that may potentially interfere with the procedure, coverslips were treated with 100 µg/mL RNase A diluted in PBS (Thermo-Fisher, cat. EN0531), for 1h at RT. Coverslips were then incubated for 1h with 20% glycerol (in PBS) at RT, followed by three consecutive rounds of freezing, thawing and soaking: 30 sec keeping coverslips on dry ice, gradual thawing in ambient air, and 2-min incubation of 20% glycerol at RT, respectively. Cells were thus washed three times in PBS (10 min each, at RT), and incubated in 0.1 M HCl for 5 min at RT. After a washing with 2x Saline-Sodium Citrate buffer (SSC, Sigma-Aldrich, cat. S6639), coverslips were incubated overnight at RT in a solution of 50% formamide (pH 7.0; Thermo Fischer, cat. AM9342) diluted in 2x SSC. The next day cells were washed in PBS for 3 min, and treated for 2 min at RT with pepsin (40 units/mL; Sigma-Aldrich, cat. P6887) diluted in 10 mM HCl, to break down cellular proteins and facilitate the entering of the probes. Pepsin was finally inactivated by two washes with 50 mM Mg2Cl2 in PBS.

The preparation of cells continued treating coverslips with 1% PFA diluted in PBS for 1 min, followed by a washing step in PBS (5 min), and twice washes (5 min each) in 2x SSC at RT. Finally, cells were incubated with 50% formamide in 2x SSC, for 1h at RT.

9.10.2 Probe preparation and DNA FISH hybridization

We used a DNA probe for the *CDKN1A* locus (CHR6: 36644236-36655116) obtained from Genomic Empire (cat. CDKN1A-20-OR), and labelled with 5-TAMRA (orange dUTP, 548nm). A probe mix was prepared by adding 2 µL of *CDKN1A* probe and 8 µL of hybridization buffer (Genomic Empire). The mix was denatured at 73 °C for 5 min, immediately after followed by 2 min in ice, and then left for 15 min at 37 °C. In parallel, coverslips with cells were warmed at 73 °C for 5 minutes, to denature genomic DNA. After denaturation of probe and genomic DNA, 10 µL/sample of probe mixture was added on a pre-heated slide (73 °C), and for each drop of probe mix a coverslip with cells was placed on top, with cells facing the probe. Coverslips were sealed with rubber cement probes, to avoid the hybridization mix to dry out. Cells were then placed into a humidified chamber, and incubated at 37°C for 16 hours (in the dark) to allow hybridization. The next day, rubber cement was removed from slides and the coverslips were washed with solution A, composed as following indicated: 0.3% Igepal (Sigma, cat. CA-630) diluted in 0.4x SSC. Cells were incubated with solution A at 73°C for 2 min (in the dark), while gently shaking. A second wash was performed with the solution B (0.1% Igepal in 2x SSC) for 2 min

(in the dark) at RT, while shaking. Cells were next incubated with a DNA staining solution (Hoechst 33342, 1 µg/mL in PBS) for 10 min at RT in the dark. After three PBS washes in PBS, coverslips were mounted on slides using Vectashield and sealed with nail polish.

Images of DNA FISH were collect using our custom-built microscope (see the next section), with a 60x oil-immersion objective (N.A.=1.49) and a sCMOS camera. Chromatinized nuclei (stained by Hoechst) were acquired with our high-resolution implementation (mSIM), using a 405 nm laser (49 ms/frame laser exposure, 50 ms/frame camera exposure, 20 fps), collecting 224 images per cell. The description of the mSIM imaging approach is reported in more detail in the next section.

10 Live-Cell Imaging

Single molecule imaging movies were acquired using a custom-built widefield microscope, designed to illuminate the sample by an inclined laser beam (Tokunaga *et al*, 2008). The microscope is equipped with a 60x oil-immersion objective (Olympus Life Science, N.A.=1.49NA), a *Hamamatsu Orca Fusion* sCMOS camera (Hamamatsu Photonics; physical pixel size=108.3 nm), and control systems for temperature (37°C), CO_2 (5%) and humidity, to maintain cells under physiological conditions during live-cell experiments.

10.1 Photo-activated Single Molecule Tracking (paSMT)

To acquire SMT movies at a low molecule density, cells expressing HaloTag-conjugated proteins were labeled with two fluorescent ligands, a photoactivatable (PA) JF549 for the SMT videos and the JF646 used for collecting a reference image of the nucleus by a red light led source (Excelitas Xcite XLED1, Qioptiq). Photo-activation allowing SMT was achieved through a 405 nm laser (Coherent, continuous-wave current), with the number of photoactivated molecules (2-3 molecules per frame) tuned by the laser power (0.5-2 mW at the microscope entrance). Photoactivated molecules were excited by a 561 nm laser (Cobolt 60-DPL, 20 mW power at the microscope entrance), adopting stroboscopic illumination to reduce the photobleaching rate and to follow individual molecules for prolonged times. The resulting time-lapse movies (10,000 frames per video) were acquired at 5 ms/frame of laser exposure (561 nm laser) and 10ms/frame of camera exposure, to obtain 100 frames per second (fps) movies.

10.2 SMT/mSIM

To acquire SMT movies while collecting reference images of the nuclear architecture, our custom-built microscope was upgraded with a multifocal structured illumination microscopy (mSIM) implementation. The scheme of our SMT/mSIM microscope is depicted in Figure 10.1A. Briefly the microscope is composed by two illumination arms. The first one, directs the laser light from a 200mW 561 nm laser (Cobolt 06-DPL, Hubner Photonics) and from a 200mW 647 nm laser (Coherent Obis, Coherent Inc) to perform SMT using HiLO illumination. Here a movable mirror (MM) in a conjugated plane of the back focal aperture of the objective allows to achieved the desired light beam inclination in the object plane (that is set to roughly 67°). The second line expands the collimated light from 405nm and 488 nm lasers (Coherent Obis, Coherent Inc) to 0.6 cm in size and directs it onto a digital micromirror device (DMD, Vialux v7000) that allows to project a pattern of diffraction-limited spots on the sample. The physical size of each DMD micromirror is 13.67 μm, and with the lenses used in our microscope (see scheme), this corresponds to approximately a projected image of each pixel on the sample plane of approximately 117nm in side. The chosen illumination pattern (an equilateral triangular lattice with side equal to 16 DMD pixels) is then scanned over the mSIM field of view (Figure 10.1B). 224 different images are necessary to completely scan the entire field of view (**Movie 5**). The two illumination arms are then combined through a dichroic mirror (DM1), reflecting wavelengths below 490 nm and transmitting longer wavelengths (Semrock Inc.).

We verified that upon reconstruction of the mSIM super-resolved image (see below), the lateral resolution of the microscope increased by a factor higher 1.4x (Figure 10.1C), and that minimal chromatic aberration (below the microscope resolution limit) was present when imaging fluorescent beads with the two different illumination arms of the microscope (Figure 10.1D).

Cells expressing HaloTag-conjugated proteins, were labeled with JF549 ligand and Hoechst 33342 (to stain DNA). Two mSIM images were collected before and after each SMT movie, projecting on the sample a series of 224 patterns of diffraction-limited spots (one pattern per frame) through the 405 nm laser (49 ms/frame laser exposure, 50 ms/frame camera exposure, 20 fps). The SMT movie were acquired using stroboscopic illumination by the 561 nm laser (5 ms/frame laser exposure, 10 ms/frame camera exposure, 100 fps), collecting 2000 frames per movie. Comparison of the two mSIM acquisitions for each cell allows to discard those acquisitions affected by cellular motion or stage drift.

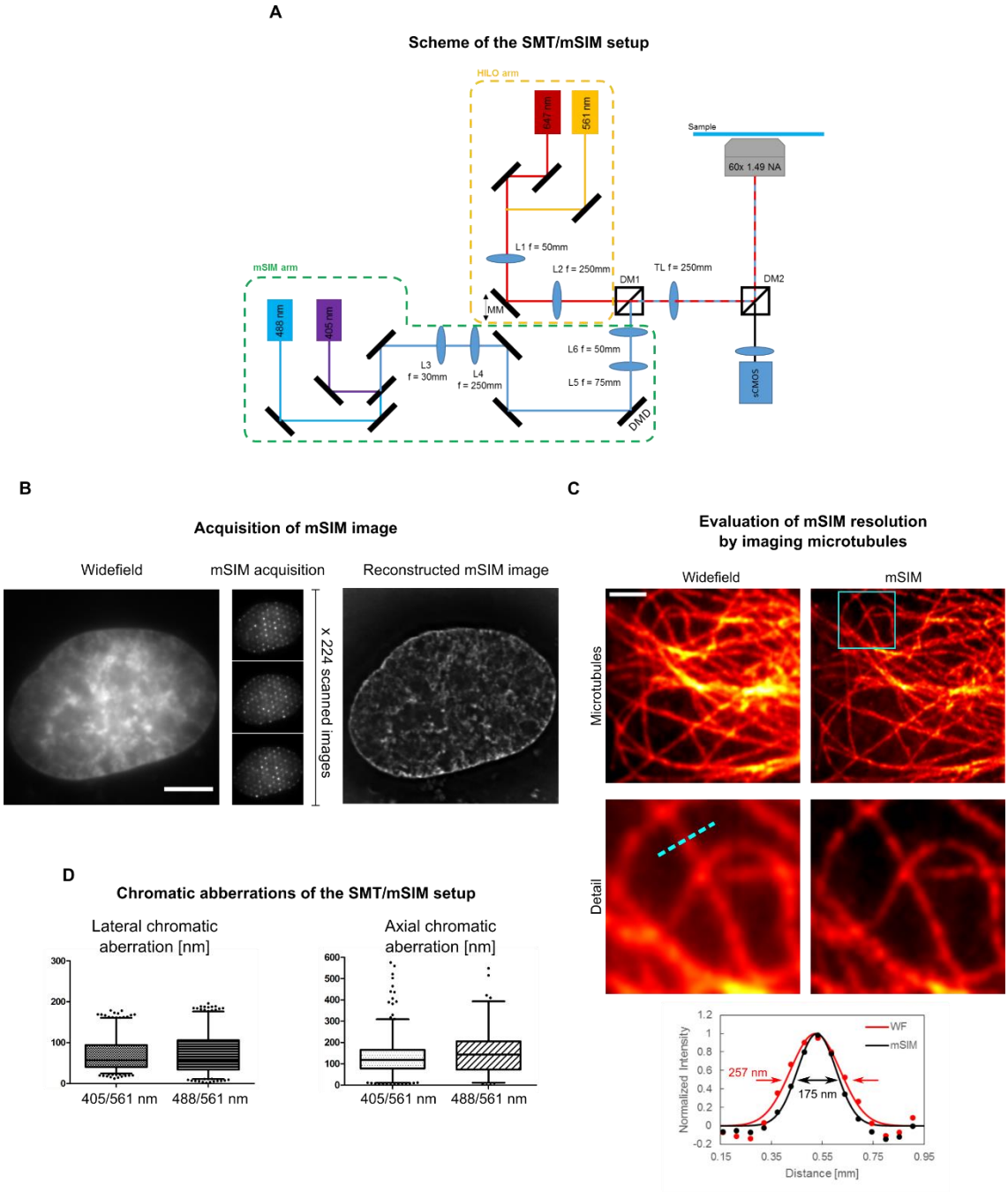


Figure 10.1 mSIM microscopy. *(A) Scheme of the SMT/mSIM setup. (B) Example of a mSIM acquisition. (C) Resolution increase of mSIM by imaging microtubules. (D) Chromatic aberrations of the SMT/mSIM setup, by imaging 100nm-sized fluorescent beads.*

10.3 Labelling of cells expressing HaloTag-conjugated proteins

Two days before paSMT or SMT/mSIM experiments, cells were seeded onto 4-well LabTek chambers (Thermo-Fisher, cat. 155382PK) at a ~30% confluence. One hour

before imaging, cells were labeled using slightly different protocols depending on paSMT or SMT/mSIM.

paSMT. Cells were incubated for 15 min at 37°C with 10 nM PA-JF549 ligand diluted in phenol-red free DMEM. Next, 10 nM JF646 ligand was added to the medium, incubating the cells for further 15 min at 37°C. After labeling, cells were thoroughly washed in PBS (two rounds of three washes separated by incubation at 37 °C for 15 min in phenol-red free DMEM).

SMT/mSIM. Cells were incubated for 30 min at 37°C with 1 nM JF549 ligand diluted in phenol-red free DMEM. After three washes in PBS, cells were incubated for 10 min at 37°C with Hoechst 33342 (2 µg/mL) diluted in phenol-red free DMEM. Cells were washed three times in PBS and fresh phenol-red free DMEM was added.

11 Data analysis

11.1 Tracking and analysis of SMT movies

SMT movies from paSMT and SMT/mSIM acquisitions were analyzed using the ImageJ plugin 'TrackMate' that detects individual spots in each frame of a movie and connects them into tracks (Tinevez *et al*, 2017). The diffraction-limited spots are identified by applying a LoG (Lapacian of Gaussian) filter to each frame to suppress noise and signal arising from large structures (e.g. cellular autofluorescence, out of focus background). Spot detection was performed by setting 0.8 µm as maximum molecule diameter. The identified spots are then connected into tracks, using the LAP algorithm (Jaqaman *et al*, 2008) that requires to specify the maximal frame to frame displacement for the observed molecular specie. The probability of a diffusing molecule to displace more than a certain distance r_{max} in a time Δt can be calculated as:

$$P(r > r_{max}, \Delta t) = e^{-\frac{{r_{max}}^2}{4D\Delta t}}$$

Where D is the diffusion coefficient of the observed molecule. The maximal frame-to-frame displacement in TrackMate was therefore set using the relationship above, so that less than 1% of the tracks would be missed for typical diffusion coefficients of nuclear proteins (smaller than $5\mu m^2/s$; (Lu & Lionnet, 2021)). This leads to a maximal displacement threshold $r_{max} = 1\mu m$ in a frame-to-frame time $\Delta t = 10\,ms.$ The control experiments on unconjugated HaloTag were instead analyzed with a maximal displacement $r_{max} = 2\mu m$, since its diffusion coefficient has been reported in the range of $15\ \mu m^2/s$ (Lu & Lionnet, 2021).

TrackMate generated tables providing lists with x, y and t positions for each track. The track tables were imported in Matlab and the tracks were analyzed using custom-written routines to: (a) Generate and model the distribution of single molecule displacements, in order to extract the abundancy of each kinetic subpopulation (indicated as bound, slow and fast fractions in the result section), and their respective diffusion coefficients.
(b) Classify individual track segments in one of the three subpopulations using a Hidden Markov model approach, vbSPT.
(c) Analyze the co-clustering between bound molecules and diffusing molecules.
(d) Analyze the diffusional anisotropy of unbound molecules.

11.1.1 Analysis of the distribution of single molecule displacements

The tracks generated by TrackMate were used to populate a histogram of frame-to-frame displacements, using a bin size Δr equal to 20nm. The histogram was then normalized (to a probability density function, pdf) to give the probability $p(r)\Delta r$ to observe a molecule moving a distance between $r-\frac{\Delta r}{2}$ and $r+\frac{\Delta r}{2}$ in the time Δt. This probability is then fit by a three-component diffusion model:

$$p(r)\Delta r = r\Delta r \sum_{i=1}^{3} \frac{f_i}{2D_i\Delta t} \exp\left(-\frac{r^2}{4D_i\Delta t}\right)$$

Where f_i is the fraction of molecules with diffusion coefficient D_i, so that $\sum_{i=1}^{3} f_i = 1$. In this formulation D_i assumes the value of apparent diffusion coefficient given by the sum of the true diffusion coefficient and a term due to the limited localization accuracy σ^2 in single molecule tracking: $D_i = D_{i,true} + \frac{\sigma^2}{\Delta t}$. Further, such analysis can underestimate the fraction of molecules in the fast diffusion state, given that faster diffusing molecules have a larger probability of escaping the observation volume. To correct the estimated fractions of molecules in each state, we normalized each of the fraction to the probability that a molecule will move less than 1μm along the optical axis:

$$f_{i,corr} = \frac{f_i}{\mathrm{erf}\left(\frac{1}{2\sqrt{4D_i\Delta t}}\right)}$$

Fitting of the distribution of displacements was performed at the single cell level, and the distributions of obtained parameters ($D_{bound}, D_{slow}, D_{fast}, f_{bound}, f_{slow}$) were plotted as violin-plots using the IoSR-Surrey Matlab toolbox, and compared using

Kruskal-Wallis non-parametric testing. P-values were Bonferroni adjusted for multiple comparisons on 5 different parameters ($D_{bound}, D_{slow}, D_{fast}, f_{slow}, f_{fast}$).

11.1.2 Classification of Track Segments using vbSPT.

To classify tracks into 'Bound', 'Slow diffusing' and 'Fast diffusing' segments, we accumulated data on different cells, isolated tracks lasting longer than 8 frames and gave them as input to the vbSPT algorithm (Persson *et al*, 2013), by imposing a maximum number of states equal to 3. The algorithm provides the diffusion coefficients for each of the states together the transition rates between these states. We verified that the diffusion coefficients extracted for each isolated component are in agreement with the ones obtained by the analysis of displacements. Further, we verified that no contamination of bound molecules is present in the diffusing components extracted by vbSPT, that is fundamental in order to correctly analyze diffusional anisotropy (see below).

11.1.3 Co-clustering between bound and diffusing molecules.

The distribution of distances between bound and free molecules were evaluated by first calculating the centroid of each of the track segments classified as bound and then accumulating an histogram of the distances between these centroids and all the particle positions belonging to 'slow diffusing' and 'fast diffusing' track segments. These histograms were then compared to the distribution of distances between bound molecules and randomly positioned molecules within the cell nucleus.

Cross-correlation between PML bodies and p53 molecule positions.

Cross-correlation between the positions of PML bodies and p53 molecules were computed in circular regions of $1\mu m$ in radius, centered on PML bodies, as described in (Stone & Veatch, 2015).

11.1.4 Analysis of diffusional anisotropy.

To analyze diffusional anisotropy, we removed bound track segments, and we calculated the angle between two consecutive displacements. We then calculated the fold anisotropy metric $f_{180/0}$, a measure of more likely is for a single molecule to jump in the backward direction, compared to taking a step forward – calculated as the probability of observing a backward jump (with an angle between jumps in the range $[180° - 30°, 180° + 30°]$), divided by the probability of observing a forward jump (with

an angle between jumps in the range $[0° - 30°, 0° + 30°]$). Plotting $f_{180/0}$ as function of the distance run by the molecule, allows to discriminate between compact, non-compact and guided exploration (Hansen *et al*, 2020). Error bars were calculated as standard deviation from 100 subsampling of the data using 50% of the original data.

11.2 Reconstruction of mSIM images

As described by York and colleagues (York *et al*, 2012), in order to reconstruct a super-resolved and optically sectioned image from the 224 individual frames acquired in our mSIM set-up, it is necessary to: (i) identify the position of the illumination spots in the image plane at each acquisition frame; (ii) Perform digital pinholing of the individual frames to get rid of out-of-focus blur; (iii) fuse together the images through pixel-reassignment that result in a $\sqrt{2}$ increase in the lateral resolution of the microscope. All these steps are performed with custom-written routines in Matlab.

11.2.1 Identification of the illumination spot position

The position at which illumination spots appear in the image, does not necessarily correspond to the points in which the DMD spots illuminate the sample, but rather they are the convolution product of the illumination pattern projected on the sample for the actual distribution of fluorescent labels in the sample. On average, however, the distance between the recorded spots should reflect the average distance between illumination points. Similar to what performed in York et al., we use this principle to find the lattice vectors that describe the 2D displacement between any two illumination spots.

The position of all illumination points in the acquisition stack can be defined by two set of vectors (that need to be found for each individual acquisition):

a) The lattice vectors that describe the displacement between an illumination spot and the two nearest neighbor ones.
b) The offset vectors, that specify the absolute position of the illumination spot closer to the top-left of the image in each of the images of the raw acquisition.

11.2.2 Identification of the lattice vectors

To identify the lattice vector, we first localize the positions at which illumination spots appear in the sample by using the ThunderStorm plug-in in ImageJ/FIJI. These coordinates are then used to generate a stack of binary images with ones at the pixels where the illumination spots have been localized and zero elsewhere. The

images are then Fourier transformed and the Fourier magnitude images are then averaged together, giving rise to a periodic lattice of peaks, spaced by the inverse of the average spacing between peaks in the real images. Next, we search for peaks in the Fourier dimension, verify that we can find harmonics of the peaks found at lowest spatial frequency, and identify three candidate peaks with the lowest spatial frequency (that show harmonics). We next verify that the vector sum and differences of the position vector of these identified peaks, also point to a detectable peak. If two vectors satisfy these conditions these are chosen as the lattice vectors in the Fourier space, and are then Fourier transformed to obtain real lattice vectors.

11.2.3 Identification of the offset vectors

Once we have the lattice vectors, we can find the position of any other illumination spot in one of the images of the raw stack by knowing the position of one of them (i.e. by finding the position of the top left illumination spot in each them image, the offset vectors). To this scope we translate the expected pattern of illumination over the x and y positions and maximize the autocorrelation between the expected illumination distribution and the experimental one.

11.2.4 Virtual Pinholing and pixel reassignment

To suppress out of focus blur, we next perform virtual pinholing, around the positions of illumination spots detected above. To this scope, we scale the raw images of a factor 10 in both dimensions, and we apply a gaussian mask of 130nm in standard deviation around the position of each detected illumination spot. Next, for every frame in the stack and for every spot position in the image the raw data from a squared region around the illumination spot are copied to the final image matrix in a squared region centered to the original coordinates multiplied by two. This procedure is analogous to move the information from each pixel to half of its distance from the illumination center, the principle on which Image Scanning Microscopy (ISM) is based (York *et al*, 2012). Summing the result of this procedure for each illumination spot generates the final super-resolved image that is then scaled down by a factor 10 and deconvolved using YY the Richardson-Lucy algorithm.

Given that the illumination pattern is uneven – dimmer at the FOV edge than at its center – and to provide a map for local variations in chromatin density - the resulting images are then normalized, by dividing the mSIM image for a blurred version (with a gaussian filter with a standard deviation of 2.16 μm).

11.3 Analysis of the SMT/mSIM acquisition

Analysis of the combined SMT/mSIM acquisition are carried out with custom written Matlab Routines. First, TF tracks are generated from the SMT movie using TrackMate as described above, and sent to the vbSPT algorithm to extract track segments belonging to the 'bound', 'slow diffusion' and 'fast diffusion' states. These classified track segments are then overlayed to the mSIM image of DNA density, and the normalized intensity and gradient of the Hoechst intensity signal at positions occupied by the p53 molecules are calculated, and used to populate cumulative histograms. Finally, the average Hoechst intensity in an $1\mu m \times 1\mu m$ pixel region surrounding each molecule are calculated separately for each of the three diffusion states of the TF.

11.4 Analysis of smFISH data

smFISH data was analyzed using the Matlab package FISH-quant. Mature RNA molecules are identified as 3D Gaussian spots with maximal intensity above an arbitrary threshold, that was kept constant for each of the target RNAs analyzed. Nascent RNAs at active transcription sites were identified by looking for high intensity nuclear foci setting the detection threshold so that no more than four actively transcribed loci could be found in each nucleus. This analysis allows to count the number of active transcription sites per cell and to count the number of pre-mRNAs present at the transcription site at the moment of fixation. Single cell abundance of mature and nascent RNAs were plotted as violin-plots using the IoSR-Surrey Matlab toolbox, and compared using Kruskal-Wallis non-parametric testing. The number of transcription sites per cell were plot as stacked bar plots and compared using the Fisher exact test for the probability of having 1 active TS or more.

To provide information about the kinetics of transcription, we fit the distribution of mRNAs in single cells to a model that accounts for a gene switching stochastically between an active (on) state, during which transcription can occur, and an inactive (off) state. This random telegraph model depends on two parameters, the burst size (i.e. how many mRNAs are produced on average during an 'on' period) and the burst frequency (i.e. how frequently the gene switches from the 'off' to the 'on' state).

In particular the model we used is given by (Raj *et al*, 2006):

$$P(n_{RNA}) = P(n_{RNA}|f_{rel}, b) = \frac{\Gamma(f_{rel} + n_{RNA})}{\Gamma(f_{rel})\Gamma(n_{RNA} + 1)} \left(\frac{b}{b+1}\right)^{n_{RNA}} \left(\frac{1}{b+1}\right)^{f_{rel}}$$

that provides the probability of observing a certain number of mRNAs per cell, $P(n_{RNA})$, as function of the size of the transcriptional burst b, and the relative burst frequency f_{rel}- given by the ratio between the actual burst frequency and the mRNA degradation rate, $f_{rel} = f/k$.

The equation described above is valid on a number of assumptions, namely the fact that (i) transcription is at 'steady-state', a condition that is verified upon long-term (>24h) transfection with HaloTag-p53 (Loffreda *et al*, 2017a); (ii) Transcription occurs from one allele at the time, a condition that is indeed verified for the *CDKN1A* and the *BBC3* genes, where the large majority of cells display 0 or 1 TS upon all the tested conditions. To relate transcriptional activation with the levels of expression of HaloTag-p53 (WT and FUS-p53), a custom-written routine in Matlab was used.

References

Ahn JH, Davis ES, Daugird TA, Zhao S, Quiroga IY, Uryu H, Li J, Storey AJ, Tsai Y-H, Keeley DP, *et al* (2021) Phase separation drives aberrant chromatin looping and cancer development. *Nature* 595: 591–595

Alberti S, Gladfelter A & Mittag T (2019) Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* 176: 419–434

Alexander JL & Orr-Weaver TL (2016) Replication fork instability and the consequences of fork collisions from rereplication. *Genes Dev* 30: 2241–2252

Allen MA, Andrysik Z, Dengler VL, Mellert HS, Guarnieri A, Freeman JA, Sullivan KD, Galbraith MD, Luo X, Kraus WL, *et al* (2014) Global analysis of p53-regulated transcription identifies its direct targets and unexpected regulatory mechanisms. *eLife* 3: e02200

Alsheich-Bartok O, Haupt S, Alkalay-Snir I, Saito S, Appella E & Haupt Y (2008) PML enhances the regulation of p53 by CK1 in response to DNA damage. *Oncogene* 27: 3653–3661

Amiad-Pavlov D, Lorber D, Bajpai G, Reuveny A, Roncato F, Alon R, Safran S & Volk T (2021) Live imaging of chromatin distribution reveals novel principles of nuclear architecture and chromatin compartmentalization. *Science Advances*

Arnould C, Rocher V, Finoux A-L, Clouaire T, Li K, Zhou F, Caron P, Mangeot PE, Ricci EP, Mourad R, *et al* (2021) Loop extrusion as a mechanism for formation of DNA damage repair foci. *Nature* 590: 660–665

Aubrey BJ, Kelly GL, Janic A, Herold MJ & Strasser A (2018) How does p53 induce apoptosis and how does this relate to p53-mediated tumour suppression? *Cell Death Differ* 25: 104–113

Avcu N & Molina N (2016) Chromatin structure shapes the search process of transcription factors. 050146 doi:10.1101/050146 [PREPRINT]

Axpe E, Chan D, Offeddu GS, Chang Y, Merida D, Hernandez HL & Appel EA (2019) A Multiscale Model for Solute Diffusion in Hydrogels. *Macromolecules* 52: 6889–6897

Aymard F, Bugler B, Schmidt CK, Guillou E, Caron P, Briois S, Iacovoni JS, Daburon V, Miller KM, Jackson SP, *et al* (2014) Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. *Nat Struct Mol Biol* 21: 366–374

Baldi S, Korber P & Becker PB (2020) Beads on a string—nucleosome array arrangements and folding of the chromatin fiber. *Nat Struct Mol Biol* 27: 109–118

Bancaud A, Huet S, Daigle N, Mozziconacci J, Beaudouin J & Ellenberg J (2009) Molecular crowding affects diffusion and binding of nuclear proteins in heterochromatin and reveals the fractal organization of chromatin. *EMBO J* 28: 3785–3798

Bartek J, Bartkova J & Lukas J (2007) DNA damage signalling guards against activated oncogenes and tumour progression. *Oncogene* 26: 7773–7779

Bartman CR, Hsu SC, Hsiung CC-S, Raj A & Blobel GA (2016) Enhancer Regulation of Transcriptional Bursting Parameters Revealed by Forced Chromatin Looping. *Molecular Cell* 62: 237–247

Batta K & Kundu TK (2007) Activation of p53 Function by Human Transcriptional Coactivator PC4: Role of Protein-Protein Interaction, DNA Bending, and Posttranslational Modifications. *Molecular and Cellular Biology*

Baugh EH, Ke H, Levine AJ, Bonneau RA & Chan CS (2018) Why are there hotspot mutations in the TP53 gene in human cancers? *Cell Death Differ* 25: 154–160

Bénichou O, Chevalier C, Klafter J, Meyer B & Voituriez R (2010) Geometry-controlled kinetics. *Nature Chem* 2: 472–477

Berg OG, Winter RB & Von Hippel PH (1981) Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry* 20: 6929–6948

Bernardi R & Pandolfi PP (2007) Structure, dynamics and functions of promyelocytic leukaemia nuclear bodies. *Nat Rev Mol Cell Biol* 8: 1006–1016

Biderman L, Manley JL & Prives C (2012) Mdm2 and MdmX as Regulators of Gene Expression. *Genes & Cancer* 3: 264–273

Biggin MD (2011) Animal Transcription Networks as Highly Connected, Quantitative Continua. *Developmental Cell* 21: 611–626

Bilu Y & Barkai N (2005) The design of transcription-factor binding sites is affected by combinatorial regulation. *Genome Biology* 6: R103

Bizhanova A & Kaufman PD (2021) Close to the edge: Heterochromatin at the nucleolar and nuclear peripheries. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1864: 194666

Boettiger AN, Bintu B, Moffitt JR, Wang S, Beliveau BJ, Fudenberg G, Imakaev M, Mirny LA, Wu C & Zhuang X (2016) Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* 529: 418–422

Bohrmann B, Haider M & Kellenberger E (1993) Concentration evaluation of chromatin in unstained resin-embedded sections by means of low-dose ratio-contrast imaging in STEM. *Ultramicroscopy* 49: 235–251

Boija A, Klein IA, Sabari BR, Dall'Agnese A, Coffey EL, Zamudio AV, Li CH, Shrinivas K, Manteiga JC, Hannett NM, *et al* (2018) Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell* 175: 1842-1855.e16

Botcheva K, McCorkle SR, McCombie WR, Dunn JJ & Anderson CW (2011) Distinct p53 genomic binding patterns in normal and cancer-derived human cells. *Cell Cycle* 10: 4237–4249

Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, *et al* (2005) Core Transcriptional Regulatory Circuitry in Human Embryonic Stem Cells. *Cell* 122: 947–956

Boyle AP, Araya CL, Brdlik C, Cayting P, Cheng C, Cheng Y, Gardner K, Hillier LW, Janette J, Jiang L, *et al* (2014) Comparative analysis of regulatory information and circuits across distant species. *Nature* 512: 453–456

Candau R, Scolnick DM, Darpino P, Ying CY, Halazonetis TD & Berger SL (1997) Two tandem and independent sub-activation domains in the amino terminus of p53 require the adaptor complex for activity. *Oncogene* 15: 807–816

Chang J, Kim D-H, Lee SW, Choi KY & Sung YC (1995) Transactivation Ability of p53 Transcriptional Activation Domain Is Directly Related to the Binding Affinity to TATA-binding Protein (∗). *Journal of Biological Chemistry* 270: 25014–25019

Chen C, Lim HH, Shi J, Tamura S, Maeshima K, Surana U & Gan L (2016) Budding yeast chromatin is dispersed in a crowded nucleoplasm in vivo. *MBoC* 27: 3357–3368

Chen J, Zhang Z, Li L, Chen B-C, Revyakin A, Hajj B, Legant W, Dahan M, Lionnet T, Betzig E, *et al* (2014) Single-Molecule Dynamics of Enhanceosome Assembly in Embryonic Stem Cells. *Cell* 156: 1274–1285

Chen X, Ko LJ, Jayaraman L & Prives C (1996) p53 levels, functional domains, and DNA damage determine the extent of the apoptotic response of tumor cells. *Genes & Development* 10: 2438–2451

Chen Y, Cattoglio C, Dailey G, Zhu Q, Tjian R & Darzacq X (2021) Mechanisms Governing Target Search and Binding Dynamics of Hypoxia-Inducible Factors Molecular Biology

Cheng Q & Chen J (2010) Mechanism of p53 stabilization by ATM after DNA damage. *Cell Cycle* 9: 472–478

Cho W-K, Spille J-H, Hecht M, Lee C, Li C, Grube V & Cisse II (2018) Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* 361: 412–415

Chong S, Dugast-Darzacq C, Liu Z, Dong P, Dailey GM, Cattoglio C, Heckert A, Banala S, Lavis L, Darzacq X, *et al* (2018) Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* 361: eaar2555

Chong S, Graham TGW, Dugast-Darzacq C, Dailey GM, Darzacq X & Tjian R (2021) Tuning levels of low-complexity domain interactions to modulate endogenous oncogenic transcription. 2021.08.16.456551 doi:10.1101/2021.08.16.456551 [PREPRINT]

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, *et al* (2013) Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*

Cortini R & Filion GJ (2018) Theoretical principles of transcription factor traffic on folded chromatin. *Nat Commun* 9: 1740

Coulocheri SA, Pigis DG, Papavassiliou KA & Papavassiliou AG (2007) Hydrogen bonds in protein–DNA complexes: Where geometry meets plasticity. *Biochimie* 89: 1291–1303

Cramer P (2019) Organization and regulation of gene transcription. *Nature* 573: 45–54

Cravens SL, Schonhoft JD, Rowland MM, Rodriguez AA, Anderson BG & Stivers JT (2015) Molecular crowding enhances facilitated diffusion of two human DNA glycosylases. *Nucleic Acids Research* 43: 4087–4097

Cremer T, Cremer M, Hübner B, Silahtaroglu A, Hendzel M, Lanctôt C, Strickfaden H & Cremer C (2020) The Interchromatin Compartment Participates in the Structural and Functional Organization of the Cell Nucleus. *BioEssays* 42: 1900132

Cui F, Sirotin MV & Zhurkin VB (2011) Impact of Alu repeats on the evolution of human p53 binding sites. *Biol Direct* 6: 2

Daban J-R (2000) Physical Constraints in the Condensation of Eukaryotic Chromosomes. Local Concentration of DNA versus Linear Packing Ratio in Higher Order Chromatin Structures. *Biochemistry* 39: 3861–3866

Davey CA, Sargent DF, Luger K, Maeder AW & Richmond TJ (2002) Solvent Mediated Interactions in the Structure of the Nucleosome Core Particle at 1.9Å Resolution††We dedicate this paper to the memory of Max Perutz who was particularly inspirational and supportive to T.J.R. in the early stages of this study. *Journal of Molecular Biology* 319: 1097–1113

Davidson IF & Peters J-M (2021) Genome folding through loop extrusion by SMC complexes. *Nat Rev Mol Cell Biol* 22: 445–464

Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS & Ren B (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485: 376–380

Dorigo B, Schalch T, Kulangara A, Duda S, Schroeder RR & Richmond TJ (2004) Nucleosome Arrays Reveal the Two-Start Organization of the Chromatin Fiber. *Science* 306: 1571–1573

Eischen CM (2016) Genome Stability Requires p53. *Cold Spring Harb Perspect Med* 6: a026096

El-Deiry WS, Kern SE, Pietenpol JA, Kinzler KW & Vogelstein B (1992) Definition of a consensus binding site for p53. *Nat Genet* 1: 45–49

Elf J & Barkefors I (2019) Single-Molecule Kinetics in Living Cells. *Annu Rev Biochem* 88: 635–659

Elf J, Li G-W & Xie XS (2007) Probing Transcription Factor Dynamics at the Single-Molecule Level in a Living Cell. *Science*

Elmore S (2007) Apoptosis: A Review of Programmed Cell Death. *Toxicol Pathol* 35: 495–516

Eltsov M, MacLellan KM, Maeshima K, Frangakis AS & Dubochet J (2008) Analysis of cryo-electron microscopy images does not support the existence of 30-nm chromatin fibers in mitotic chromosomes in situ. *PNAS* 105: 19732–19737

England CG, Luo H & Cai W (2015) HaloTag Technology: A Versatile Platform for Biomedical Applications. *Bioconjugate Chem* 26: 975–986

Esadze A & Iwahara J (2014) Stopped-Flow Fluorescence Kinetic Study of Protein Sliding and Intersegment Transfer in the Target DNA Search Process. *Journal of Molecular Biology* 426: 230–244

Esadze A & Stivers JT (2018a) Facilitated Diffusion Mechanisms in DNA Base Excision Repair and Transcriptional Activation. *Chem Rev* 118: 11298–11323

Esadze A & Stivers JT (2018b) Facilitated Diffusion Mechanisms in DNA Base Excision Repair and Transcriptional Activation. *Chem Rev* 118: 11298–11323

Espinosa JM & Emerson BM (2001) Transcriptional Regulation by p53 through Intrinsic DNA/Chromatin Binding and Site-Directed Cofactor Recruitment. *Molecular Cell* 8: 57–69

Fields S & Jang SK (1990) Presence of a Potent Transcription Activating Sequence in the p53 Protein. *Science*

Fierz B & Poirier MG (2019) Biophysics of Chromatin Dynamics. *Annu Rev Biophys* 48: 321–345

Finn EH & Misteli T (2019) Molecular basis and biological function of variability in spatial genome organization. *Science* 365: eaaw9498

Fischer M (2017) Census and evaluation of p53 target genes. *Oncogene* 36: 3943–3956

Fischer M, Grossmann P, Padi M & DeCaprio JA (2016) Integration of TP53, DREAM, MMB-FOXM1 and RB-E2F target gene analyses identifies cell cycle gene regulatory networks. *Nucleic Acids Res* 44: 6070–6086

Fischer M, Steiner L & Engeland K (2014) The transcription factor p53: Not a repressor, solely an activator. *Cell Cycle* 13: 3037–3058

Fogal V (2000) Regulation of p53 activity in nuclear bodies by a specific PML isoform. *The EMBO Journal* 19: 6185–6195

Fraser P & Bickmore W (2007) Nuclear organization of the genome and the potential for gene regulation. *Nature* 447: 413–417

Fritsch CC & Langowski J (2010) Anomalous diffusion in the interphase cell nucleus: The effect of spatial correlations of chromatin. *J Chem Phys* 133: 025101

Funk WD, Pak DT, Karas RH, Wright WE & Shay JW (1992) A transcriptionally active DNA-binding site for human p53 protein complexes. *Mol Cell Biol* 12: 2866–2871

Fussner E, Strauss M, Djuric U, Li R, Ahmed K, Hart M, Ellis J & Bazett-Jones DP (2012) Open and closed domains in the mouse genome are configured as 10-nm chromatin fibres. *EMBO Rep* 13: 992–996

Galganski L, Urbanek MO & Krzyzosiak WJ (2017) Nuclear speckles: molecular organization, biological function and role in disease. *Nucleic Acids Research* 45: 10350–10368

Gan L, Ladinsky MS & Jensen GJ (2013) Chromatin in a marine picoeukaryote is a disordered assemblage of nucleosomes. *Chromosoma* 122: 377–386

Garavís M & Calvo O (2017) Sub1/PC4, a multifaceted factor: from transcription to genome stability. *Curr Genet* 63: 1023–1035

Garcia DA, Johnson TA, Presman DM, Fettweis G, Wagh K, Rinaldi L, Stavreva DA, Paakinaho V, Jensen RAM, Mandrup S, *et al* (2021a) An intrinsically disordered region-mediated confinement state contributes to the dynamics and function of transcription factors. *Molecular Cell* 81: 1484-1498.e6

Garcia DA, Johnson TA, Presman DM, Fettweis G, Wagh K, Rinaldi L, Stavreva DA, Paakinaho V, Jensen RAM, Mandrup S, *et al* (2021b) An intrinsically disordered region-mediated confinement state contributes to the dynamics and function of transcription factors. *Mol Cell* 81: 1484-1498.e6

Givaty O & Levy Y (2009) Protein Sliding along DNA: Dynamics and Structural Characterization. *Journal of Molecular Biology* 385: 1087–1097

Gomes NP & Espinosa JM (2010) Gene-specific repression of the p53 target gene PUMA via intragenic CTCF–Cohesin binding. *Genes Dev* 24: 1022–1034

Görisch SM, Wachsmuth M, Tóth KF, Lichter P & Rippe K (2005) Histone acetylation increases chromatin accessibility. *Journal of Cell Science* 118: 5825–5834

Granik N, Weiss LE, Nehme E, Levin M, Chein M, Perlson E, Roichman Y & Shechtman Y (2019) Single-Particle Diffusion Characterization by Deep Learning. *Biophysical Journal* 117: 185–192

Grimm JB, English BP, Choi H, Muthusamy AK, Mehl BP, Dong P, Brown TA, Lippincott-Schwartz J, Liu Z, Lionnet T, *et al* (2016) Bright photoactivatable fluorophores for single-molecule imaging. *Nat Methods* 13: 985–988

Grimm JB, Muthusamy AK, Liang Y, Brown TA, Lemon WC, Patel R, Lu R, Macklin JJ, Keller PJ, Ji N, *et al* (2017) A general method to fine-tune fluorophores for live-cell and in vivo imaging. *Nat Methods* 14: 987–994

Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A, Wessels L, de Laat W, *et al* (2008) Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* 453: 948–951

Hafner A, Bulyk ML, Jambhekar A & Lahav G (2019) The multiple mechanisms that regulate p53 activity and cell fate. *Nat Rev Mol Cell Biol* 20: 199–210

Hafner A, Kublo L, Tsabar M, Lahav G & Stewart-Ornstein J (2020) Identification of universal and cell-type specific p53 DNA binding. *BMC Mol and Cell Biol* 21: 5

Hafner A, Stewart-Ornstein J, Purvis JE, Forrester WC, Bulyk ML & Lahav G (2017) p53 pulses lead to distinct patterns of gene expression albeit similar DNA-binding dynamics. *Nat Struct Mol Biol* 24: 840–847

Hager KM & Gu W (2014) Understanding the non-canonical pathways involved in p53-mediated tumor suppression. *Carcinogenesis* 35: 740–746

Haigis KM & Sweet-Cordero A (2011) New insights into oncogenic stress. *Nat Genet* 43: 177–178

Hainaut P & Pfeifer GP (2016) Somatic TP53 Mutations in the Era of Genome Sequencing. *Cold Spring Harb Perspect Med* 6: a026179

Halford SE & Marko JF (2004) How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Research* 32: 3040–3052

Hamard P-J, Lukin DJ & Manfredi JJ (2012) p53 Basic C Terminus Regulates p53 Functions through DNA Binding Modulation of Subset of Target Genes*. *Journal of Biological Chemistry* 287: 22397–22407

Hammar P, Leroy P, Mahmutovic A, Marklund EG, Berg OG & Elf J (2012) The lac Repressor Displays Facilitated Diffusion in Living Cells. *Science*

Hansen AS, Amitai A, Cattoglio C, Tjian R & Darzacq X (2020) Guided nuclear exploration increases CTCF target search efficiency. *Nat Chem Biol* 16: 257–266

Hansen AS, Pustova I, Cattoglio C, Tjian R & Darzacq X (2017a) CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *eLife* 6: e25776

Hansen JC (2002) Conformational Dynamics of the Chromatin Fiber in Solution: Determinants, Mechanisms, and Functions. *Annu Rev Biophys Biomol Struct* 31: 361–392

Hansen JC, Connolly M, McDonald CJ, Pan A, Pryamkova A, Ray K, Seidel E, Tamura S, Rogge R & Maeshima K (2017b) The 10-nm chromatin fiber and its relationship to interphase chromosome organization. *Biochemical Society Transactions* 46: 67–76

Hauer MH & Gasser SM (2017) Chromatin and nucleosome dynamics in DNA damage and repair. *Genes Dev* 31: 2204–2221

Hauser K, Essuman B, He Y, Coutsias E, Garcia-Diaz M & Simmerling C (2016) A human transcription factor in search mode. *Nucleic Acids Research* 44: 63–74

Heintzmann R & Huser T (2017) Super-Resolution Structured Illumination Microscopy. *Chem Rev* 117: 13890–13908

Hnisz D, Shrinivas K, Young RA, Chakraborty AK & Sharp PA (2017) A Phase Separation Model for Transcriptional Control. *Cell* 169: 13–23

Hoffman WH, Biade S, Zilfou JT, Chen J & Murphy M (2002) Transcriptional Repression of the Anti-apoptoticsurvivin Gene by Wild Type p53. *Journal of Biological Chemistry* 277: 3247–3257

Hollenberg SM & Evans RM (1988) Multiple and cooperative trans-activation domains of the human glucocorticoid receptor. *Cell* 55: 899–906

Horvath MM, Wang X, Resnick MA & Bell DA (2007) Divergent Evolution of Human p53 Binding Sites: Cell Cycle Versus Apoptosis. *PLoS Genet* 3: e127

Hsieh T-HS, Weiner A, Lajoie B, Dekker J, Friedman N & Rando OJ (2015) Mapping Nucleosome Resolution Chromosome Folding in Yeast by Micro-C. *Cell* 162: 108–119

Inga A, Storici F, Darden TA & Resnick MA (2002) Differential Transactivation by the p53 Transcription Factor Is Highly Dependent on p53 Level and Promoter Target Sequence. *Mol Cell Biol* 22: 8612–8625

Itoh Y, Murata A, Sakamoto S, Nanatani K, Wada T, Takahashi S & Kamagata K (2016) Activation of p53 Facilitates the Target Search in DNA by Enhancing the Target Recognition Probability. *Journal of Molecular Biology* 428: 2916–2930

Ivanschitz L, Takahashi Y, Jollivet F, Ayrault O, Bras ML & Thé H de (2015) PML IV/ARF interaction enhances p53 SUMO-1 conjugation, activation, and senescence. *PNAS* 112: 14278–14283

Izeddin I, Récamier V, Bosanac L, Cissé II, Boudarene L, Dugast-Darzacq C, Proux F, Bénichou O, Voituriez R, Bensaude O, *et al* (2014a) Single-molecule tracking in live cells reveals distinct target-search strategies of transcription factors in the nucleus. *eLife* 3: e02230

Izeddin I, Récamier V, Bosanac L, Cissé II, Boudarene L, Dugast-Darzacq C, Proux F, Bénichou O, Voituriez R, Bensaude O, *et al* (2014b) Single-molecule tracking in live cells reveals distinct target-search strategies of transcription factors in the nucleus. *eLife* 3: e02230

Jain S, Shukla S, Yang C, Zhang M, Fatma Z, Lingamaneni M, Abesteh S, Lane ST, Xiong X, Wang Y, *et al* (2021) TALEN outperforms Cas9 in editing heterochromatin target sites. *Nat Commun* 12: 606

Jana T, Brodsky S & Barkai N (2021) Speed–Specificity Trade-Offs in the Transcription Factors Search for Their Genomic Binding Sites. *Trends in Genetics* 37: 421–432

Jaqaman K, Loerke D, Mettlen M, Kuwata H, Grinstein S, Schmid SL & Danuser G (2008) Robust single-particle tracking in live-cell time-lapse sequences. *Nat Methods* 5: 695–702

Jenkins LMM, Durell SR, Mazur SJ & Appella E (2012) p53 N-terminal phosphorylation: a defining layer of complex regulation. *Carcinogenesis* 33: 1441–1449

Joerger AC & Fersht AR (2008) Structural Biology of the Tumor Suppressor p53. *Annu Rev Biochem* 77: 557–582

Johnson RA, Ince TA & Scotto KW (2001) Transcriptional Repression by p53 through Direct Binding to a Novel DNA Element. *Journal of Biological Chemistry* 276: 27716–27720

Kaeser MD & Iggo RD (2002) Chromatin immunoprecipitation analysis fails to support the latency model for regulation of p53 DNA binding activity in vivo. *Proceedings of the National Academy of Sciences* 99: 95–100

Kamagata K, Kanbayashi S, Honda M, Itoh Y, Takahashi H, Kameda T, Nagatsugi F & Takahashi S (2020) Liquid-like droplet formation by tumor suppressor p53 induced by multivalent electrostatic interactions between two disordered domains. *Sci Rep* 10: 580

Kan P-Y, Caterino TL & Hayes JJ (2009) The H4 Tail Domain Participates in Intra- and Internucleosome Interactions with Protein and DNA during Folding and Oligomerization of Nucleosome Arrays. *Mol Cell Biol* 29: 538–546

Kanada R, Terakawa T, Kenzaki H & Takada S (2019) Nucleosome Crowding in Chromatin Slows the Diffusion but Can Promote Target Search of Proteins. *Biophysical Journal* 116: 2285–2295

Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, Xie M, Zhang Q, McMichael JF, Wyczalkowski MA, *et al* (2013) Mutational landscape and significance across 12 major cancer types. *Nature* 502: 333–339

Kent S, Brown K, Yang C, Alsaihati N, Tian C, Wang H & Ren X (2020) Phase-Separated Transcriptional Condensates Accelerate Target-Search Process Revealed by Live-Cell Single-Molecule Imaging. *Cell Reports* 33: 108248

Khazanov N & Levy Y (2011) Sliding of p53 along DNA Can Be Modulated by Its Oligomeric State and by Cross-Talks between Its Constituent Domains. *Journal of Molecular Biology* 408: 335–355

Kim H, Kim K, Choi J, Heo K, Baek HJ, Roeder RG & An W (2012) p53 Requires an Intact C-Terminal Domain for DNA Binding and Transactivation. *Journal of Molecular Biology* 415: 843–854

Kim JM, Visanpattanasin P, Jou V, Liu S, Tang X, Zheng Q, Li KY, Snedeker J, Lavis LD, Lionnet T, *et al* (2021) Single-molecule imaging of chromatin remodelers reveals role of ATPase in promoting fast kinetics of target search and dissociation from chromatin. *eLife* 10: e69387

Kiraz Y, Adan A, Kartal Yandim M & Baran Y (2016) Major apoptotic mechanisms and genes involved in apoptosis. *Tumor Biol* 37: 8471–8486

Kitayner M, Rozenberg H, Kessler N, Rabinovich D, Shaulov L, Haran TE & Shakked Z (2006) Structural Basis of DNA Recognition by p53 Tetramers. *Molecular Cell* 22: 741–753

Kitayner M, Rozenberg H, Rohs R, Suad O, Rabinovich D, Honig B & Shakked Z (2010) Diversity in DNA recognition by p53 revealed by crystal structures with Hoogsteen base pairs. *Nat Struct Mol Biol* 17: 423–429

Kracikova M, Akiri G, George A, Sachidanandam R & Aaronson SA (2013) A threshold mechanism mediates p53 cell fate decision between growth arrest and apoptosis. *Cell Death Differ* 20: 576–588

Krebs JE, Goldstein ES & Kilpatrick ST (2017) Lewin's GENES XII Jones & Bartlett Learning

Kryston TB, Georgiev AB, Pissis P & Georgakilas AG (2011) Role of oxidative stress and DNA damage in human carcinogenesis. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 711: 193–201

Lallemand-Breitenbach V & Thé H de (2010) PML Nuclear Bodies. *Cold Spring Harb Perspect Biol* 2: a000661

Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, Chen X, Taipale J, Hughes TR & Weirauch MT (2018) The Human Transcription Factors. *Cell* 172: 650–665

Laptenko O, Beckerman R, Freulich E & Prives C (2011) p53 binding to nucleosomes within the p21 promoter in vivo leads to nucleosome loss and transcriptional activation. *PNAS* 108: 10385–10390

Laptenko O & Prives C (2006) Transcriptional regulation by p53: one protein, many possibilities. *Cell Death Differ* 13: 951–961

Laptenko O, Shiff I, Freed-Pastor W, Zupnick A, Mattia M, Freulich E, Shamir I, Kadouri N, Kahan T, Manfredi J, *et al* (2015) The p53 C Terminus Controls Site-Specific DNA Binding and Promotes Structural Changes within the Central DNA Binding Domain. *Molecular Cell* 57: 1034–1046

Laurie NA, Donovan SL, Shih C-S, Zhang J, Mills N, Fuller C, Teunisse A, Lam S, Ramos Y, Mohan A, *et al* (2006) Inactivation of the p53 pathway in retinoblastoma. *Nature* 444: 61–66

Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, Meyerson M, Gabriel SB, Lander ES & Getz G (2014) Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* 505: 495–501

Lee CS, Friedman JR, Fulmer JT & Kaestner KH (2005) The initiation of liver development is dependent on Foxa transcription factors. *Nature* 435: 944–947

Leith JS, Tafvizi A, Huang F, Uspal WE, Doyle PS, Fersht AR, Mirny LA & Oijen AM van (2012) Sequence-dependent sliding kinetics of p53. *PNAS* 109: 16552–16557

Lerner J, Gomez-Garcia PA, McCarthy RL, Liu Z, Lakadamyali M & Zaret KS (2020) Two-Parameter Mobility Assessments Discriminate Diverse Regulatory Factor Behaviors in Chromatin. *Molecular Cell* 79: 677-688.e6

Leslie M (2021) Separation anxiety. *Science* 371: 336–338

Li G, Levitus M, Bustamante C & Widom J (2005) Rapid spontaneous accessibility of nucleosomal DNA. *Nat Struct Mol Biol* 12: 46–53

Li W, Notani D & Rosenfeld MG (2016) Enhancers as non-coding RNA transcription units: recent insights and future perspectives. *Nat Rev Genet* 17: 207–223

Link N, Kurtz P, O'Neal M, Garcia-Hughes G & Abrams JM (2013) A p53 enhancer region regulates target genes through chromatin conformations in cis and in trans. *Genes Dev* 27: 2433–2438

Lipski R, Lippincott DJ, Durden BC, Kaplan AR, Keiser HE, Park J-H & Levesque AA (2012) p53 Dimers Associate with a Head-to-Tail Response Element to Repress Cyclin B Transcription. *PLoS ONE* 7: e42615

Liu Y, Tavana O & Gu W (2019) p53 modifications: exquisite decorations of the powerful guardian. *Journal of Molecular Cell Biology* 11: 564–577

Loffreda A, Jacchetti E, Antunes S, Rainone P, Daniele T, Morisaki T, Bianchi ME, Tacchetti C & Mazza D (2017a) Live-cell p53 single-molecule binding is modulated by C-terminal acetylation and correlates with transcriptional activity. *Nat Commun* 8: 313

Loffreda A, Jacchetti E, Antunes S, Rainone P, Daniele T, Morisaki T, Bianchi ME, Tacchetti C & Mazza D (2017b) Live-cell p53 single-molecule binding is modulated by C-terminal acetylation and correlates with transcriptional activity. *Nat Commun* 8: 313

Lu F & Lionnet T (2021) Transcription Factor Dynamics. *Cold Spring Harb Perspect Biol* 13: a040949

Luger K, Mäder AW, Richmond RK, Sargent DF & Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389: 251–260

Maeshima K, Ide S, Hibino K & Sasai M (2016a) Liquid-like behavior of chromatin. *Current Opinion in Genetics & Development* 37: 36–45

Maeshima K, Imai R, Tamura S & Nozaki T (2014) Chromatin as dynamic 10-nm fibers. *Chromosoma* 123: 225–237

Maeshima K, Kaizu K, Tamura S, Nozaki T, Kokubo T & Takahashi K (2015) The physical size of transcription factors is key to transcriptional regulation in chromatin domains. *J Phys: Condens Matter* 27: 064116

Maeshima K, Rogge R, Tamura S, Joti Y, Hikima T, Szerlong H, Krause C, Herman J, Seidel E, DeLuca J, *et al* (2016b) Nucleosomal arrays self-assemble into supramolecular globular structures lacking 30-nm fibers. *EMBO J* 35: 1115–1132

Maeshima K, Tamura S, Hansen JC & Itoh Y (2020) Fluid-like chromatin: Toward understanding the real chromatin organization present in the cell. *Current Opinion in Cell Biology* 64: 77–89

Mardoukhi Y, Jeon J-H & Metzler R (2015) Geometry controlled anomalous diffusion in random fractal geometries: looking beyond the infinite cluster. *Phys Chem Chem Phys* 17: 30134–30147

Maréchal A & Zou L (2013) DNA Damage Sensing by the ATM and ATR Kinases. *Cold Spring Harb Perspect Biol* 5: a012716

Marney CB, Anderson ES, Adnan M, Peng K-L, Hu Y, Weinhold N & Schmitt AM (2021) p53-intact cancers escape tumor suppression through loss of long noncoding RNA Dino. *Cell Reports* 35: 109329

Mazza D, Abernathy A, Golob N, Morisaki T & McNally JG (2012a) A benchmark for chromatin binding measurements in live cells. *Nucleic Acids Research* 40: e119

Mazza D, Abernathy A, Golob N, Morisaki T & McNally JG (2012b) A benchmark for chromatin binding measurements in live cells. *Nucleic Acids Research* 40: e119

Mazzocca M, Colombo E, Callegari A & Mazza D (2021a) Transcription factor binding kinetics and transcriptional bursting: What do we really know? *Current Opinion in Structural Biology* 71: 239–248

Mazzocca M, Fillot T, Loffreda A, Gnani D & Mazza D (2021b) The needle and the haystack: single molecule tracking to probe the transcription factor search in eukaryotes. *Biochemical Society Transactions* 49: 1121–1132

McKinney K, Mattia M, Gottifredi V & Prives C (2004) p53 Linear Diffusion along DNA Requires Its C Terminus. *Molecular Cell* 16: 413–424

McSwiggen DT, Hansen AS, Teves SS, Marie-Nelly H, Hao Y, Heckert AB, Umemoto KK, Dugast-Darzacq C, Tjian R & Darzacq X (2019a) Evidence for DNA-mediated nuclear compartmentalization distinct from phase separation. *eLife* 8: e47098

McSwiggen DT, Mir M, Darzacq X & Tjian R (2019b) Evaluating phase separation in live cells: diagnosis, caveats, and functional consequences. *Genes Dev* 33: 1619–1634

Mehta GD, Ball DA, Eriksson PR, Chereji RV, Clark DJ, McNally JG & Karpova TS (2018) Single-Molecule Analysis Reveals Linked Cycles of RSC Chromatin Remodeling and Ace1p Transcription Factor Binding in Yeast. *Molecular Cell* 72: 875-887.e9

Melo CA, Drost J, Wijchers PJ, van de Werken H, de Wit E, Vrielink JAFO, Elkon R, Melo SA, Léveillé N, Kalluri R, *et al* (2013) eRNAs Are Required for p53-Dependent Enhancer Activity and Gene Transcription. *Molecular Cell* 49: 524–535

Meyer B, Bénichou O, Kafri Y & Voituriez R (2012) Geometry-Induced Bursting Dynamics in Gene Expression. *Biophysical Journal* 102: 2186–2191

Michael AK & Thomä NH (2021) Reading the chromatinized genome. *Cell* 184: 3599–3611

Mirny L, Slutsky M, Wunderlich Z, Tafvizi A, Leith J & Kosmrlj A (2009) How a protein searches for its site on DNA: the mechanism of facilitated diffusion. *J Phys A: Math Theor* 42: 434013

Miron E, Oldenkamp R, Brown JM, Pinto DMS, Xu CS, Faria AR, Shaban HA, Rhodes JDP, Innocent C, de Ornellas S, *et al* Chromatin arranges in chains of mesoscale domains with nanoscale functional topography independent of cohesin. *Science Advances* 6: eaba8811

Misteli T (2007) Beyond the Sequence: Cellular Organization of Genome Function. *Cell* 128: 787–800

Misteli T (2020) The Self-Organizing Genome: Principles of Genome Architecture and Function. *Cell* 183: 28–45

Momand J, Jung D, Wilczynski S & Niland J (1998) The MDM2 gene amplification database. *Nucleic Acids Research* 26: 3453–3459

Morris JP, Yashinskie JJ, Koche R, Chandwani R, Tian S, Chen C-C, Baslan T, Marinkovic ZS, Sánchez-Rivera FJ, Leach SD, *et al* (2019) α-Ketoglutarate links p53 to cell fate during tumour suppression. *Nature* 573: 595–599

Murata A, Itoh Y, Mano E, Kanbayashi S, Igarashi C, Takahashi H, Takahashi S & Kamagata K (2017) One-Dimensional Search Dynamics of Tumor Suppressor p53 Regulated by a Disordered C-Terminal Domain. *Biophysical Journal* 112: 2301–2314

Nagashima R, Hibino K, Ashwin SS, Babokhov M, Fujishiro S, Imai R, Nozaki T, Tamura S, Tani T, Kimura H, *et al* (2019) Single nucleosome imaging reveals loose genome chromatin networks via active RNA polymerase II. *Journal of Cell Biology* 218: 1511–1530

Narlikar GJ, Myong S, Larson D, Maeshima K, Francis N, Rippe K, Sabari B, Strader L & Tjian R (2021) Is transcriptional regulation just going through a phase? *Molecular Cell* 81: 1579–1585

Nguyen T-AT, Grimm SA, Bushel PR, Li J, Li Y, Bennett BD, Lavender CA, Ward JM, Fargo DC, Anderson CW, *et al* (2018) Revealing a human p53 universe. *Nucleic Acids Research* 46: 8153–8167

Nguyen VQ, Ranjan A, Liu S, Tang X, Ling YH, Wisniewski J, Mizuguchi G, Li KY, Jou V, Zheng Q, *et al* (2021) Spatiotemporal coordination of transcription preinitiation complex assembly in live cells. *Molecular Cell* 81: 3560-3575.e6

Nikulenkov F, Spinnler C, Li H, Tonelli C, Shi Y, Turunen M, Kivioja T, Ignatiev I, Kel A, Taipale J, *et al* (2012) Insights into p53 transcriptional function via genome-wide chromatin occupancy and gene expression analysis. *Cell Death Differ* 19: 1992–2002

Nili EL, Field Y, Lubling Y, Widom J, Oren M & Segal E (2010) p53 binds preferentially to genomic regions with high DNA-encoded nucleosome occupancy. *Genome Res* 20: 1361–1368

Nishino Y, Eltsov M, Joti Y, Ito K, Takata H, Takahashi Y, Hihara S, Frangakis AS, Imamoto N, Ishikawa T, *et al* (2012) Human mitotic chromosomes consist predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure: Human mitotic chromosomes consist of nucleosome fibres. *The EMBO Journal* 31: 1644–1653

Noureddine MA, Menendez D, Campbell MR, Bandele OJ, Horvath MM, Wang X, Pittman GS, Chorley BN, Resnick MA & Bell DA (2009) Probing the Functional Impact of Sequence Variation on p53-DNA Interactions Using a Novel Microsphere Assay for Protein-DNA Binding with Human Cell Extracts. *PLoS Genet* 5: e1000462

Nozaki T, Imai R, Tanbo M, Nagashima R, Tamura S, Tani T, Joti Y, Tomita M, Hibino K, Kanemaki MT, *et al* (2017) Dynamic Organization of Chromatin Domains Revealed by Super-Resolution Live-Cell Imaging. *Molecular Cell* 67: 282-293.e7

Ohno M, Ando T, Priest DG, Kumar V, Yoshida Y & Taniguchi Y (2019) Sub-nucleosomal Genome Structure Reveals Distinct Nucleosome Folding Motifs. *Cell* 176: 520-534.e25

Olins AL & Olins DE (1974) Spheroid Chromatin Units (ν Bodies). *Science* 183: 330–332

Osborne CS, Chakalova L, Brown KE, Carter D, Horton A, Debrand E, Goyenechea B, Mitchell JA, Lopes S, Reik W, *et al* (2004) Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat Genet* 36: 1065–1071

Ou HD, Phan S, Deerinck TJ, Thor A, Ellisman MH & O'Shea CC (2017) ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. *Science* 357: eaag0025

Papantonis A & Cook PR (2013) Transcription Factories: Genome Organization and Gene Regulation. *Chem Rev* 113: 8683–8705

Parmar JJ, Woringer M & Zimmer C (2019) How the Genome Folds: The Biophysics of Four-Dimensional Chromatin Organization. *Annu Rev Biophys* 48: 231–253

Pavletich NP, Chambers KA & Pabo CO (1993) The DNA-binding domain of p53 contains the four conserved regions and the major mutation hot spots. *Genes & Development* 7: 2556–2564

Pearson M, Carbone R, Sebastiani C, Cioce M, Fagioli M, Saito S, Higashimoto Y, Appella E, Minucci S, Pandolfi PP, *et al* (2000) PML regulates p53 acetylation and premature senescence induced by oncogenic Ras. *Nature* 406: 207–210

Pederson T (2000) Diffusional protein transport within the nucleus: a message in the medium. *Nat Cell Biol* 2: E73–E74

Persson F, Lindén M, Unoson C & Elf J (2013) Extracting intracellular diffusive states and transition rates from single-molecule tracking data. *Nat Methods* 10: 265–269

Pfeifer GP, Denissenko MF, Olivier M, Tretyakova N, Hecht SS & Hainaut P (2002) Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. *Oncogene* 21: 7435–7451

Piovesan A, Pelleri MC, Antonaros F, Strippoli P, Caracausi M & Vitale L (2019) On the length, weight and GC content of the human genome. *BMC Research Notes* 12: 106

Polach KJ & Widom J (1995) Mechanism of Protein Access to Specific DNA Sequences in Chromatin: A Dynamic Equilibrium Model for Gene Regulation. *Journal of Molecular Biology* 254: 130–149

Purvis JE, Karhohs KW, Mock C, Batchelor E, Loewer A & Lahav G (2012) p53 Dynamics Control Cell Fate. *Science*

Qian H, Wang T, Naumovski L, Lopez CD & Brachmann RK (2002) Groups of p53 target genes involved in specific p53 downstream effects cluster into different classes of DNA binding sites. *Oncogene* 21: 7901–7911

Quaedackers ME, Van Den Brink CE, Wissink S, Schreurs RHMM, Gustafsson J-Å, Van Der Saag PT & Van Der Burg B (2001) 4-Hydroxytamoxifen Trans-Represses Nuclear Factor-κB Activity in Human Osteoblastic U2-OS Cells through Estrogen Receptor (ER)α, and Not through ERβ*. *Endocrinology* 142: 1156–1166

Quinodoz SA, Jachowicz JW, Bhat P, Ollikainen N, Banerjee AK, Goronzy IN, Blanco MR, Chovanec P, Chow A, Markaki Y, *et al* (2021) RNA promotes the formation of spatial compartments in the nucleus. *Cell* 184: 5775-5790.e30

Raj A, Peskin CS, Tranchina D, Vargas DY & Tyagi S (2006) Stochastic mRNA Synthesis in Mammalian Cells. *PLOS Biology* 4: e309

Raj N & Attardi LD (2017) The Transactivation Domains of the p53 Protein. *Cold Spring Harb Perspect Med* 7: a026047

Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, *et al* (2014) A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell* 159: 1665–1680

Rashi-Elkeles S, Warnatz H-J, Elkon R, Kupershtein A, Chobod Y, Paz A, Amstislavskiy V, Sultan M, Safer H, Nietfeld W, *et al* (2014) Parallel Profiling of the Transcriptome, Cistrome, and Epigenome in the Cellular Response to Ionizing Radiation. *Sci Signal* 7

Récamier V, Izeddin I, Bosanac L, Dahan M, Proux F & Darzacq X (2014) Single cell correlation fractal dimension of chromatin: A framework to interpret 3D single molecule super-resolution. *Nucleus* 5: 75–84

Redman-Rivera LN, Shaver TM, Jin H, Marshall CB, Schafer JM, Sheng Q, Hongo RA, Beckermann KE, Wheeler FC, Lehmann BD, *et al* (2021) Acquisition of aneuploidy drives mutant p53-associated gain-of-function phenotypes. *Nat Commun* 12: 5184

Reed S & Quelle D (2014) p53 Acetylation: Regulation and Consequences. *Cancers* 7: 30–69

Ricci MA, Manzo C, García-Parajo MF, Lakadamyali M & Cosma MP (2015) Chromatin Fibers Are Formed by Heterogeneous Groups of Nucleosomes In Vivo. *Cell* 160: 1145–1158

Richmond TJ & Davey CA (2003) The structure of DNA in the nucleosome core. *Nature* 423: 145–150

Riggs AD, Bourgeois S & Cohn M (1970) The lac represser-operator interaction. *Journal of Molecular Biology* 53: 401–417

Riley KJ-L & Maher LJ (2007) p53–RNA interactions: New clues in an old mystery. *RNA* 13: 1825–1833

Rokudai S, Laptenko O, Arnal SM, Taya Y, Kitabayashi I & Prives C (2013) MOZ increases p53 acetylation and premature senescence through its complex formation with PML. *PNAS* 110: 3895–3900

Rowell JP, Simpson KL, Stott K, Watson M & Thomas JO (2012) HMGB1-Facilitated p53 DNA Binding Occurs via HMG-Box/p53 Transactivation Domain Interaction, Regulated by the Acidic Tail. *Structure* 20: 2014–2024

Sabari BR, Dall'Agnese A, Boija A, Klein IA, Coffey EL, Shrinivas K, Abraham BJ, Hannett NM, Zamudio AV, Manteiga JC, *et al* (2018) Coactivator condensation at super-enhancers links phase separation and gene control. *Science* 361: eaar3958

Sahu G, Wang D, Chen CB, Zhurkin VB, Harrington RE, Appella E, Hager GL & Nagaich AK (2010) p53 Binding to Nucleosomal DNA Depends on the Rotational Positioning of DNA Response Element*. *Journal of Biological Chemistry* 285: 1321–1332

Sammons MA, Zhu J, Drake AM & Berger SL (2015) TP53 engagement with the genome occurs in distinct local chromatin environments via pioneer factor activity. *Genome Res* 25: 179–188

Schalch T, Duda S, Sargent DF & Richmond TJ (2005) X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* 436: 138–141

Schlereth K, Beinoraviciute-Kellner R, Zeitlinger MK, Bretz AC, Sauer M, Charles JP, Vogiatzi F, Leich E, Samans B, Eilers M, *et al* (2010) DNA Binding Cooperativity of p53 Modulates the Decision between Cell-Cycle Arrest and Apoptosis. *Molecular Cell* 38: 356–368

Schlereth K, Heyl C, Krampitz A-M, Mernberger M, Finkernagel F, Scharfe M, Jarek M, Leich E, Rosenwald A & Stiewe T (2013) Characterization of the p53 Cistrome – DNA Binding Cooperativity Dissects p53's Tumor Suppressor Functions. *PLoS Genet* 9: e1003726

Schmitt AM, Garcia JT, Hung T, Flynn RA, Shen Y, Qu K, Payumo AY, Peres-da-Silva A, Broz DK, Baum R, *et al* (2016) An inducible long noncoding RNA amplifies DNA damage signaling. *Nat Genet* 48: 1370–1376

Schoenfelder S & Fraser P (2019) Long-range enhancer–promoter contacts in gene expression control. *Nat Rev Genet* 20: 437–455

Schröder M, Yusein-Myashkova S, Todorova J, Pasheva E & Ugrinova I (2019) High mobility group box 1 protein (HMGB1) stimulates the nuclear accumulation of p53. *Biotechnology & Biotechnological Equipment* 33: 645–650

Shlesinger MF (2007) First encounters. *Nature* 450: 40–41

Silverstein TD, Gibb B & Greene EC (2014) Visualizing protein movement on DNA at the single-molecule level using DNA curtains. *DNA Repair* 20: 94–109

Sinha RP & Häder D-P (2002) UV-induced DNA damage and repair: a review. *Photochem Photobiol Sci* 1: 225–236

Smeenk L, van Heeringen SJ, Koeppel M, Gilbert B, Janssen-Megens E, Stunnenberg HG & Lohrum M (2011) Role of p53 Serine 46 in p53 Target Gene Regulation. *PLoS ONE* 6: e17574

Song F, Chen P, Sun D, Wang M, Dong L, Liang D, Xu R-M, Zhu P & Li G (2014) Cryo-EM Study of the Chromatin Fiber Reveals a Double Helix Twisted by Tetranucleosomal Units. *Science* 344: 376–380

Spector DL & Lamond AI (2011) Nuclear Speckles. *Cold Spring Harb Perspect Biol* 3: a000646

Spolar RS & M. Thomas Record J (1994) Coupling of Local Folding to Site-Specific Binding of Proteins to DNA. *Science*

Statello L, Guo C-J, Chen L-L & Huarte M (2021) Gene regulation by long non-coding RNAs and its biological functions. *Nat Rev Mol Cell Biol* 22: 96–118

van Steensel B & Furlong EEM (2019) The role of transcription in shaping the spatial organization of the genome. *Nat Rev Mol Cell Biol* 20: 327–337

Stone MB & Veatch SL (2015) Steady-state cross-correlations for live two-colour super-resolution localization data sets. *Nat Commun* 6: 7347

Stormo GD & Fields DS (1998) Specificity, free energy and information content in protein–DNA interactions. *Trends in Biochemical Sciences* 23: 109–113

Strickfaden H, Tolsma TO, Sharma A, Underhill DA, Hansen JC & Hendzel MJ (2020) Condensed Chromatin Behaves like a Solid on the Mesoscale In Vitro and in Living Cells. *Cell* 183: 1772-1784.e13

Štros M (2010) HMGB proteins: Interactions with DNA and chromatin. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1799: 101–113

Štros M, Muselíková-Polanská E, Pospíšilová Š & Strauss F (2004) High-Affinity Binding of Tumor-Suppressor Protein p53 and HMGB1 to Hemicatenated DNA Loops. *Biochemistry* 43: 7215–7225

Su D, Wang X, Campbell MR, Song L, Safi A, Crawford GE & Bell DA (2015) Interactions of Chromatin Context, Binding Site Sequence Content, and Sequence Evolution in Stress-Induced p53 Occupancy and Transactivation. *PLoS Genet* 11: e1004885

Subekti DRG, Murata A, Itoh Y, Takahashi S & Kamagata K (2020) Transient binding and jumping dynamics of p53 along DNA revealed by sub-millisecond resolved single-molecule fluorescence tracking. *Sci Rep* 10: 13697

Sullivan KD, Galbraith MD, Andrysik Z & Espinosa JM (2018) Mechanisms of transcriptional regulation by p53. *Cell Death Differ* 25: 133–143

Suter DM (2020) Transcription Factors and DNA Play Hide and Seek. *Trends in Cell Biology* 30: 491–500

Sutherland H & Bickmore WA (2009) Transcription factories: gene expression in unions? *Nat Rev Genet* 10: 457–466

Szabo Q, Bantignies F & Cavalli G (2019) Principles of genome folding into topologically associating domains. *Science Advances*

Tafvizi A, Huang F, Fersht AR, Mirny LA & Oijen AM van (2011) A single-molecule characterization of p53 search on DNA. *PNAS* 108: 563–568

Tafvizi A, Huang F, Leith JS, Fersht AR, Mirny LA & van Oijen AM (2008) Tumor Suppressor p53 Slides on DNA with Low Friction and High Stability. *Biophysical Journal* 95: L01–L03

Tang Y, Zhao W, Chen Y, Zhao Y & Gu W (2008) Acetylation Is Indispensable for p53 Activation. *Cell* 133: 612–626

Teilum K, Olsen JG & Kragelund BB (2021) On the specificity of protein–protein interactions in the context of disorder. *Biochemical Journal* 478: 2035–2050

Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, Sheffield NC, Stergachis AB, Wang H, Vernot B, *et al* (2012) The accessible chromatin landscape of the human genome. *Nature* 489: 75–82

Tims HS, Gurunathan K, Levitus M & Widom J (2011) Dynamics of Nucleosome Invasion by DNA Binding Proteins. *Journal of Molecular Biology* 411: 430–448

Tinevez J-Y, Perry N, Schindelin J, Hoopes GM, Reynolds GD, Laplantine E, Bednarek SY, Shorte SL & Eliceiri KW (2017) TrackMate: An open and extensible platform for single-particle tracking. *Methods* 115: 80–90

Tokunaga M, Imamoto N & Sakata-Sogawa K (2008) Highly inclined thin illumination enables clear single-molecule imaging in cells. *Nat Methods* 5: 159–161

Tonelli C, Morelli MJ, Bianchi S, Rotta L, Capra T, Sabò A, Campaner S & Amati B (2015) Genome-wide analysis of p53 transcriptional programs in B cells upon exposure to genotoxic stress *in vivo*. *Oncotarget* 6: 24611–24626

Trojanowski J, Frank L, Rademacher A, Grigaitis P & Rippe K (2021) Transcription activation is enhanced by multivalent interactions independent of phase separation. 2021.01.27.428421 doi:10.1101/2021.01.27.428421 [PREPRINT]

Tumbar T, Sudlow G & Belmont AS (1999) Large-Scale Chromatin Unfolding and Remodeling Induced by VP16 Acidic Activation Domain. *Journal of Cell Biology* 145: 1341–1354

Tweddle DA, Pearson ADJ, Haber M, Norris MD, Xue C, Flemming C & Lunec J (2003) The p53 pathway and its inactivation in neuroblastoma. *Cancer Letters* 197: 93–98

Vega AR, Freeman SA, Grinstein S & Jaqaman K (2018) Multistep Track Segmentation and Motion Classification for Transient Mobility Analysis. *Biophysical Journal* 114: 1018–1025

Vendramin R, Litchfield K & Swanton C (2021) Cancer evolution: Darwin and beyond. *The EMBO Journal* 40

Verfaillie A, Svetlichnyy D, Imrichova H, Davie K, Fiers M, Kalender Atak Z, Hulselmans G, Christiaens V & Aerts S (2016) Multiplex enhancer-reporter assays uncover unsophisticated TP53 enhancer logic. *Genome Res* 26: 882–895

Visvanathan A, Ahmed K, Even-Faitelson L, Lleres D, Bazett-Jones DP & Lamond AI (2013) Modulation of Higher Order Chromatin Conformation in Mammalian Cell Nuclei Can Be Mediated by Polyamines and Divalent Cations. *PLOS ONE* 8: e67689

Wei C-L, Wu Q, Vega VB, Chiu KP, Ng P, Zhang T, Shahab A, Yong HC, Fu Y, Weng Z, *et al* (2006) A Global Map of p53 Transcription-Factor Binding Sites in the Human Genome. *Cell* 124: 207–219

Wei M-T, Chang Y-C, Shimobayashi SF, Shin Y, Strom AR & Brangwynne CP (2020) Nucleated transcriptional condensates amplify gene expression. *Nat Cell Biol* 22: 1187–1196

Wei S, Falk SJ, Black BE & Lee T-H (2015) A novel hybrid single molecule approach reveals spontaneous DNA motion in the nucleosome. *Nucleic Acids Research* 43: e111

Weidtkamp-Peters S, Lenser T, Negorev D, Gerstner N, Hofmann TG, Schwanitz G, Hoischen C, Maul G, Dittrich P & Hemmerich P (2008) Dynamics of component exchange at PML nuclear bodies. *Journal of Cell Science* 121: 2731–2743

Weinberg RL, Freund SMV, Veprintsev DB, Bycroft M & Fersht AR (2004a) Regulation of DNA Binding of p53 by its C-terminal Domain. *Journal of Molecular Biology* 342: 801–811

Weinberg RL, Veprintsev DB, Bycroft M & Fersht AR (2005) Comparative Binding of p53 to its Promoter and DNA Recognition Elements. *Journal of Molecular Biology* 348: 589–596

Weinberg RL, Veprintsev DB & Fersht AR (2004b) Cooperative Binding of Tetrameric p53 to DNA. *Journal of Molecular Biology* 341: 1145–1159

Woodcock CL, Frado LL & Rattner JB (1984) The higher-order structure of chromatin: evidence for a helical ribbon arrangement. *J Cell Biol* 99: 42–52

Woringer M & Darzacq X (2018a) Protein motion in the nucleus: from anomalous diffusion to weak interactions. *Biochemical Society Transactions* 46: 945–956

Woringer M & Darzacq X (2018b) Protein motion in the nucleus: from anomalous diffusion to weak interactions. *Biochem Soc Trans* 46: 945–956

Woringer M, Darzacq X & Izeddin I (2014) Geometry of the nucleus: a perspective on gene expression regulation. *Current Opinion in Chemical Biology* 20: 112–119

York AG, Parekh SH, Nogare DD, Fischer RS, Temprine K, Mione M, Chitnis AB, Combs CA & Shroff H (2012) Resolution doubling in live, multicellular organisms via multifocal structured illumination microscopy. *Nat Methods* 9: 749–754

Younger ST & Rinn JL (2017) p53 regulates enhancer accessibility and activity in response to DNA damage. *Nucleic Acids Research* 45: 9889–9900

Yu X & Buck MJ (2019) Defining TP53 pioneering capabilities with competitive nucleosome binding assays. *Genome Res* 29: 107–115

Zandarashvili L, Vuzman D, Esadze A, Takayama Y, Sahu D, Levy Y & Iwahara J (2012) Asymmetrical roles of zinc fingers in dynamic DNA-scanning process by the inducible transcription factor Egr-1. *PNAS* 109: E1724–E1732

Zaret KS & Mango SE (2016) Pioneer transcription factors, chromatin dynamics, and cell fate control. *Current Opinion in Genetics & Development* 37: 76–81

Zheng C, Lu X, Hansen JC & Hayes JJ (2005) Salt-dependent Intra- and Internucleosomal Interactions of the H3 Tail Domain in a Model Oligonucleosomal Array. *Journal of Biological Chemistry* 280: 33552–33557

Appendix

Movie availability:
A PPTx file with the movies cited in this thesis can be found at:
https://www.dropbox.com/s/sd8vzzi5m5dn6zy/MAzzocca_MoviesForPhDthesis.pptx?dl=1

Sequence of the repair vector used to generate the p53-HaloTag KI cells:
Here we report the sequence of the repair vector used to generate p53-HaloTag cells. The vector is composed by 2 regions of homology with TP53, the HaloTag insert to be add to the Tp53locus and an antibiotic resistance to facilitate selection of cells correctly modified:

Homology Arm 1 (sequence homologous to Tp53locus)
GAACATCCAAACCCAGGGACGAGTGTGGATACTTCTTTGCCATTCTCCGCAACTCCCAGCC
CAGAGCTGGAGGGTCTCAAGGAGGGGCCTAATAATTGTGTAATACTGAATACAGCCAGAG
TTTCAGGTCATATACTCAGCCCTGCCATGCACCGGCAGGTCCTAGGTGACCCCCGTCAAAC
TCAGTTTCCTTATATATAAAATGGGGTAAGGGGGCCGGGCGCAGTGGCTCACGAATCCCA
CACTCTGGGAGGCCAAGGCGAGTGGATCACCTGAGGTCGGGAGTTTGAGCCCAGCCTGA
CCAACATGGAGAAACCCCATCTCTACTAAAAATACAAAAGTAGCCGGGCGTGGTGATGCAT
GCCTGTAATCCCAGCTACCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGA
GGCAGAGGTTGCGGTGAGCTGAGATCTCACCATTACACTCCAGCCTGGGCAACAAGAGTG
AAACTCCGTCTCAAAAAAGATAAATAAAGTAAAATGGGGTAAGGGAAGATTACGAGACTAA
TACACACTAATACTCTGAGGTGCTCAGTAAACATATTTGCATGGGGTGTGGCCACCATCTT
GATTTGAATTCCCGTTGTCCCAGCCTTAGGCCCTTCAAAGCATTGGTCAGGGAAAAGGGGC
ACAGACCCTCTCACTCATGTGATGTCATCTCTCCTCCCTGCTTCTGTCTCCTACAGTCATCT
CAAATCGAAGAAAGGTCAGTCTACCTCCCGCCATAAAAAACTCATGTTCAAGACAGAAGGG
CCTGACTCAGACTCTAGA

HaloTag
ATGGCAGAAATCGGTACTGGCTTTCCATTCGACCCCCATTATGTGGAAGTCCTGGGCGAGC
GCATGCACTACGTCGATGTTGGTCCGCGCGATGGCACCCCTGTGCTGTTCCTGCACGGTA
ACCCGACCTCCTCCTACGTGTGGCGCAACATCATCCCGCATGTTGCACCGACCCATCGCTG
CATTGCTCCAGACCTGATCGGTATGGGCAAATCCGACAAACCAGACCTGGGTTATTTCTTC
GACGACCACGTCCGCTTCATGGATGCCTTCATCGAAGCCCTGGGTCTGGAAGAGGTCGTC
CTGGTCATTCACGACTGGGGCTCCGCTCTGGGTTTCCACTGGGCCAAGCGCAATCCAGAG
CGCGTCAAAGGTATTGCATTTATGGAGTTCATCCGCCCTATCCCGACCTGGGACGAATGGC
CAGAATTTGCCCGCGAGACCTTCCAGGCCTTCCGCACCACCGACGTCGGCCGCAAGCTGA
TCATCGATCAGAACGTTTTTATCGAGGGTACGCTGCCGATGGGTGTCGTCCGCCCGCTGAC
TGAAGTCGAGATGGACCATTACCGCGAGCCGTTCCTGAATCCTGTTGACCGCGAGCCACT
GTGGCGCTTCCCAAACGAGCTGCCAATCGCCGGTGAGCCAGCGAACATCGTCGCGCTGGT
CGAAGAATACATGGACTGGCTGCACCAGTCCCCTGTCCCGAAGCTGCTGTTCTGGGGCAC
CCCAGGCGTTCTGATCCCACCGGCCGAAGCCGCTCGCCTGGCCAAAAGCCTGCCTAACTG
CAAGGCTGTGGACATCGGCCCGGGTCTGAATCTGCTGCAAGAAGACAACCCGGACCTGAT
CGGCAGCGAGATCGCGCGCTGGCTGTCGACGCTTGAGATTTCCGGTcCGTA

Geneticin resistance
ATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGC
TATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCG
CAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAG
GACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCT
CGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGG
ATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCG
GCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATC
GAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAG
CATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGG
CGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGG
CCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACAT
AGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCT
CGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGAC
GAGTTCTTCTGA

Homology Arm 2 (sequence homologous to Tp53locus)
CATTCTCCACTTCTTGTTCCCCACTGACAGCCTCCCACCCCCATCTCTCCCTCCCCTGCCAT
TTTGGGTTTTGGGTCTTTGAACCCTTGCTTGCAATAGGTGTGCGTCAGAAGCACCCAGGAC
TTCCATTTGCTTTGTCCCGGGGCTCCACTGAACAAGTTGGCCTGCACTGGTGTTTTGTTGT
GGGGAGGAGGATGGGGAGTAGGACATACCAGCTTAGATTTTAAGGTTTTTACTGTGAGGG
ATGTTTGGGAGATGTAAGAAATGTTCTTGCAGTTAAGGGTTAGTTTACAATCAGCCACATTC
TAGGTAGGGGCCCACTTCACCGTACTAACCAGGGAAGCTGTCCCTCACTGTTGAATTTTCT
CTAACTTCAAGGCCCATATCTGTGAAATGCTGGCATTTGCACCTACCTCACAGAGTGCATT
GTGAGGGTTAATGAAATAATGTACATCTGGCCTTGAAACCACCTTTTATTACATGGGGTCG
AGAACTTGACCCCCTTGAGGGTGCTTGTTCCCTCTCCCTGTTGGTCGGTGGGTTGGTAGTT
TCTACAGTTGGGCAGCTGGTTAGGTAGAGGGAGTTGTCAAGTCTCTGCTGGCCCAGCCAA
ACCCTGTCTGACAACCTCTTGGTGAACCTTAGTACCTAAAAGGAAATCTCACCCCATCCCAC
ACCCTGGAGGATTTCATCTCTTGTATATGATGATCTGGATCCACCAAGACTTGTTTTATGCT
CAGGGTCAATTTCTTTTTTCTTTTTTTTTTTTTTTTTCTTTTTCTTTGAGACTGGGTCTCGCT
TT

Primers/Oligos List:

qPCR primers for mRNA targets	
CDKN1A Fw:	5’-CTGGAGACTCTCAGGGTCGAAA-3’,
CDKN1A Rev:	5’-GATTAGGGCTTCCTCTTGGAGAA-3’;
MDM2 Fw:	5’-GTGAATCTACAGGGACGCCATC-3’
MDM2 Rev:	5’-CTGATCCAACCAATCACCTGAA-3’
BBC3 Fw:	5’-GAAGAGCAAATGAGCCAAACG-3’
BBC3 Rev:	5’-GGAGCAACCGGCAAACG-3’

smRNA FISH probes for *BBC3* (*HuluFISH,* PixelBiotech GmbH)	
Name	Sequence
BBC3_132	CCGCTCTAACTGCAGTGG
BBC3_434	TGTGCGTTGAGGTCGTCC
BBC3_452	TCTTGTCTCCGCCGCTCGTAC
BBC3_502	TGAGATTGTACAGGACCCTCC
BBC3_529	TGGGTAAGGGCAGGAGTC
BBC3_559	AATTGGGCTCCATCTCGGGGG
BBC3_76	TGCAGAGAAAGTCCCCCG
BBC3_94	GGAGTCCAGTATGCTACATGG
BBC3_170	ATCTCCGTCAGTGCACCCCAG
BBC3_197	TTCTTGGCCAGGGACCCAGGA
BBC3_221	ATCAGCCGTCCCTCTCCT
BBC3_239	ACCTTCCGATGCTGAGTCC
BBC3_300	TTCCTGAGATGGTGGTGGGCG
BBC3_321	CACCAGCACAACAGCCT
BBC3_341	TGTCACCCCTGCAGCTGGAAC
BBC3_388	GCCCAGAGTGAAGGAGCA
BBC3_446	TTTGCTCTTCACGGGCCCCCT
BBC3_467	AGTGGTCACGTTTGGCTCA
BBC3_539	TCAGGGAAGATGGCTGGG
BBC3_577	TGAAGGTGAGGCAGGCATG
BBC3_609	TTCACAGTCTGGGCCCCTCCT
BBC3_654	AATGGGATTGATGGGGCGGGG
BBC3_675	GTGCTCTCTCTAAACCTATGC
BBC3_697	ATGAATGCCAGTGGTCACA
BBC3_719	AGCCAAAATCTCCCACCCCC
BBC3_802	TCTTCCGAGATTTCCCCC
BBC3_826	ATCCCTCCCCACTCCCAGACT
BBC3_848	ACAGTATCTTACAGGCTGGGC
BBC3_869	GGTATCTACAGCAGCGCATAT
smRNA FISH probes for *CDKN1A* (*HuluFISH,* PixelBiotech GmbH)	
Name	Sequence
CDKN1A_211	CTGCCGCAGAAACACCTGT
CDKN1A_0	CCCAGCCGGTTCTGACAT
CDKN1A_18	GCATGGGTTCTGACGGACATC
CDKN1A_85	ATCACAGTCGCGGCTCAGCT
CDKN1A_160	TCCAGTGGTGTCTCGGTGA
CDKN1A_219	CGTGGGAAGGTAGAGCTTGGG
CDKN1A_255	TCCTCCCAACTCATCCCG
CDKN1A_305	TCTTCCTCTGCTGTCCCCTGC
CDKN1A_345	GCGAGGCACAAGGGTACAAGA
CDKN1A_366	TTCAGCCTGCTCCCCTGA
CDKN1A_393	TGAGAGTCTCCAGGTCCACC

CDKN1A_427	TCTGTCATGCTGGTCTGCC
CDKN1A_447	CCGGCGTTTGGAGTGGTAGAA
CDKN1A_11	TTCCAGGACTGCAGGCTTCC
CDKN1A_41	ATGTAGAGCGGGCCTTTGAG
CDKN1A_121	GCCAGGGTATGTACATGAGGA
CDKN1A_187	CTCTCATTCAACCGCCTAGTT
CDKN1A_208	TGCCCAGCACTCTTAGGAAC
CDKN1A_251	ACACGGGATGAGGAGGCTTTA
CDKN1A_313	GGAGGAGGAAGTAGCTGGCAT
CDKN1A_335	TACCACCCAGCGGACAAGTG
CDKN1A_367	AGCGATGGGAAGGAGCCACAC
CDKN1A_393	GGGTGAATTTCATAACCGCCT
CDKN1A_414	GGTCTGAGTGTCCAGGAAAGG
CDKN1A_460	CCCTTCAAAGTGCCATCTGTT
CDKN1A_483	ATGATGCCCCCACTCGGTGAG
CDKN1A_530	CACCCTGCCCAACCTTAGAG
CDKN1A_550	GCTGTGCTCACTTCAGGGT
CDKN1A_576	TACCAGGTCCCCAGCTCA
CDKN1A_594	GGGTATCAAGAGCCAGGAGGG
CDKN1A_620	CCCCTGCCTTCACAAGACA
CDKN1A_676	TGCAGGTCAGAGGGGCCATGA
smRNA FISH probes for *GAPDH* (*HuluFISH,* PixelBiotech GmbH)	
GAPDH_35	CGGCTGGCGACGCAAAAGAAG
GAPDH_57	TGGTGTCTGAGCGATGTGGC
GAPDH_77	CGACCTTCACCTTCCCCA
GAPDH_102	GCCCAATACGACCAAATCCG
GAPDH_132	CCAGAGTTAAAAGCAGCCCTG
GAPDH_235	CGGTGCCATGGAATTTGCC
GAPDH_254	CTTCCCGTTCTCAGCCTTGA
GAPDH_291	TCCTGGAAGATGGTGATGGGA
GAPDH_317	CCCACTTGATTTTGGAGGGAT
GAPDH_345	TCCACGACGTACTCAGCGCCA
GAPDH_384	CCAGCCTTCTCCATGGTG
GAPDH_424	GGCAGAGATGATGACCCTTTT
GAPDH_452	TGACGAACATGGGGGCATCAG
GAPDH_499	GCTGATGATCTTGAGGCTGTT
GAPDH_530	TGCTAAGCAGTTGGTGGTGC
GAPDH_575	GTCCTTCCACGATACCAAAGT
GAPDH_601	GTGATGGCATGGACTGTGGT
GAPDH_629	GGGCCATCCACAGTCTTCT
GAPDH_648	TCACGCCACAGTTTCCCGGAG
GAPDH_674	TGATGTTCTGGAGAGCCCCGC
GAPDH_739	CTTCCCGTTCAGCTCAGG

GAPDH_786	CCACCACTGACACGTTGGCA
GAPDH_811	AGGTTTTTCTAGACGGCAGGT
GAPDH_841	CACCACCTTCTTGATGTCATC
GAPDH_888	TCAGTGTAGCCCAGGATGCCC
GAPDH_925	GGTGTCGCTGTTGAAGTCAGA
GAPDH_947	CAGCGTCAAAGGTGGAGGAGT
GAPDH_978	ACAAAGTGGTCGTTGAGGGCA
GAPDH_1018	GCTGTAGCCAAATTCGTTGTC
GAPDH_1061	ACTCCTTGGAGGCCATGTGGG
GAPDH_1107	TCTCTTCCTCTTGTGCTCTTG
GAPDH_1145	ACTGAGTGTGGCAGGGACTCC
smRNA FISH probes for *MDM2* (*Design Ready Stellaris,* Biosearch Technologies)	
MDM1	TGTTGGTATTGCACATTTGC
MDM2	TACAGCACCATCAGTAGGTA
MDM3	CGAAGCTGGAATCTGTGAGG
MDM4	TAAGTGTCTTTTTGTGCACC
MDM5	TAGTCATAATATACTGGCCA
MDM6	ATGTTGTTGCTTCTCATCAT
MDM7	TGGCACGCCAAACAAATCTC
MDM8	CTGTGCTCTTTCACAGAGAA
MDM9	TTCCTGTAGATCATGGTATA
MDM10	CTGCTGATTGACTACTACCA
MDM11	TGATCACTCCCACCTTCAAG
MDM12	GGTAGATGGTCTAGAAACCA
MDM13	CTAATTGCTCTCCTTCTAGA
MDM14	GTCGTTCACCAGATAATTCA
MDM15	CTATCAGATTTGTGGCGTTT
MDM16	TATTACACACAGAGCCAGGC
MDM17	TTCTTTCACAACATATCTCC
MDM18	CCTGTAGATTCACTGCTACT
MDM19	CACTTACACCAGCATCAAGA
MDM20	TCCAACCAATCACCTGAATG
MDM21	CTGATCTGAAACTGAATCCT
MDM22	CCTTCTTCACTAAGGCTATA
MDM23	CTGCCTGATACACAGTAACT
MDM24	ATTGCATGAAGTGCATTTCC
MDM25	AATGTGATGGAAGGGGGGGA
MDM26	GTGTTGAGTTTTCCAGTTTG
MDM27	ACTCTCTGGAATCATTCACT
MDM28	TCATCATTTTCCTCAACACA
MDM29	TGTGATTGTGAAGCTTGTGT

MDM30	GGCTGAGAATAGTCTTCACT
MDM31	ACTCTTTCACATCTTCTTGG
MDM32	AAGGGGCAAACTAGATTCCA
MDM33	TCACACAAGGTTCAATGGCA
MDM34	TAAGATGTCCTGTTTTGCCA
MDM35	TTTGCACATGTAAAGCAGGC
MDM36	TTGGTTGTCTACATACTGGG
MDM37	AGGGGAAATAAGTTAGCACA
MDM38	ATTCTCTTATAGACAGGTCA